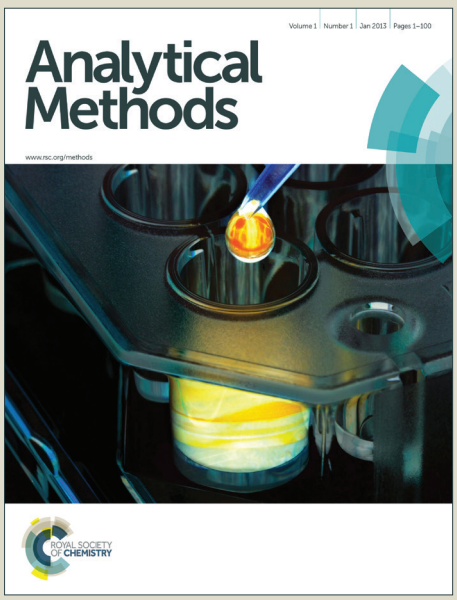
Analytical Methods

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/methods

Chemical Fingerprint and Simultaneous Determination of Flavonoids in Flos Sophorae Immaturus by HPLC-DAD and HPLC-DAD-ESI-MS/MS Combined with Chemometrics Analysis

Zhisheng Xie[1, 2], Shingchung Lam[1, 2], Junwei Wu[1, 2], Depo Yang[1, 2], Xinjun Xu[1, 2*]

[1]*School of Pharmaceutical Sciences, Sun Yat-sen University, Guangzhou, 510006 , China*

[2]*Guangdong Technology Research Centre for Advanced Chinese Medicine, Guangzhou, 510006 , China*

[*] *Correspondence details: Xinjun Xu

Address: No. 132, East Waihuan Rd.,Guangzhou Higher Education Mega Centre, Guangzhou, 510006 China

Tel.: +86 20 39943041; Fax: +86 20 39943000

E-mail: xxj2702@sina.com

Abstract

A simple and efficient HPLC method was developed for the fingerprint analysis of Flos Sophorae Immaturus and for the simultaneous determination of three flavonoids in Flos Sophorae Immaturus, namely rutin, narcissin and quercetin. The separation of analytes was conducted on a Dikma Diamonsil C_{18} column (250 mm×4.6 mm, 5 μm) at 35 ℃; the wavelength of UV detector was set at 254 nm; the mobile phases were composed of acetonitrile and aqueous acetic acid (0.5%, v/v). A gradient elution was carried out at a flow rate of 1.0 mL/min. A common chromatographic fingerprint consisting of 9 characteristic peaks was established among 14 batches of Flos Sophorae Immaturus samples, which were collected from different source areas of China. The constituents in Flos Sophorae Immaturus were further identified by HPLC-DAD-ESI-MS/MS and most of them were flavonoids. Multiple chemometrics analysis, including similarity analysis (SA), hierarchical cluster analysis (HCA) and principal component analysis (PCA) were performed to classify samples. In quantitative analysis, all of the calibration curves showed good linear regression ($R^2 \geq 0.9991$) within the tested ranges, and the mean recoveries ranged from 100.28% to 101.08%. Three flavonoid compounds in Flos Sophorae Immaturus were simultaneously quantified by the established method. The results demonstrated that the developed method was accurate and effective, which could be readily utilized for the comprehensive quality control of Flos Sophorae Immaturus.

Keywords: Chemometrics analysis; Fingerprint; Flos Sophorae Immaturus; High-performance liquid chromatography; Quantification

1 Introduction

Flos Sophorae Immaturus ("Huaimi" in Chinese), the dried flower buds of *Sophora japonica* L. (Leguminosae) produced in many provinces of China, is clinically used as a traditional Chinese medicine (TCM) to cure diseases like hemafecia, hemorrhoids blood, blood flux, uterine bleeding and hematemesis, etc..[1-2] Empirical results had shown that flavonoids were the major bioactive compounds of Flos Sophorae Immaturus, which exhibited anti-allergic, anti-inflammatory, anti-oxidative, free radical scavenging and anti-mutagenic activities.[3-5] As the most abundant flavonoid in Flos Sophorae Immaturus, rutin is documented in Chinese pharmacopoeia (2010 Version) as the index component for the quality control of Flos Sophorae Immaturus.[6] However, only a single component or a single class of component does not totally stand for the bioactivity of the medicinal plant; it is the synergy of all components that accounts.[7] Most of the researches focused on the quantification of rutin or quercetin in Flos Sophorae Immaturus, but few methods for quantifying multi-component were reported.[8-12] In that case, most of the constituents of Flos Sophorae Immaturus remained unclear. In order to comprehensively perform good quality control, it is necessary to develop an efficient method to analyze the constituents of Flos Sophorae Immaturus.

Fingerprint has been internationally regarded as a good approach for the quality control of TCMs and their derived products.[13] The combination of fingerprint with the quantification of multi-component was proved to be a feasible and comprehensive approach when controlling the quality of TCMs, which not only quantified active

components but also gave an overview of all the constituents in TCM.[14-16]

In this work, an HPLC-DAD method was developed to acquire the fingerprinting profile of Flos Sophorae Immaturus and also to quantify three major flavonoids in Flos Sophorae Immaturus, which were rutin, narcissin and quercetin (Figure 1). Besides, HPLC-DAD-ESI-MS was used for the identification of chemical constituents in Flos Sophorae Immaturus. Furthermore, chemometrics analysis including similarity analysis (SA), hierarchical clustering analysis (HCA) and principal components analysis (PCA) were successfully applied to demonstrate the variability between the fingerprints of 14 batches of Flos Sophorae Immaturus collected from different localities.

2 Materials and Methods

2.1 Chemicals and reagents

Reference substance of rutin was purchased from the National Institutes for Food and Drug Control of China (Batch number: 10080-200707). Reference substance of narcissin and quercetin were provided by Guangdong Technology Research Center for Advanced Chinese Medicine. Acetonitrile was of chromatographic grade, which was purchased from SK Chemicals (Seoul, Korea). Ultrapure water was obtained from a Milli-Q RG purification unit (ELGA Lab Water, UK). Other solvents were of analytical grade and were purchased from Tianjin Damao Chemical Reagent Factory (Tianjin, China).

2.2 Preparation of reference and sample solutions

The standard stock solutions of the three compounds, rutin, narcissin and quercetin,

were prepared by respectively dissolving each reference substance with methanol-water (80: 20, v/v), and then were stored in the refrigerator at 4 ℃. The concentration of each stock solution was 0.6 mg/mL (rutin), 0.13 mg/mL (narcissin) and 0.04 mg/mL (quercetin). The working solutions were prepared by appropriately diluting the stock solutions with methanol-water (80: 20, v/v) to yield another six concentrations. Fourteen batches of Flos Sophorae Immaturus were collected from different regions of China and authenticated as the dried flower buds of *Sophora japonica* L. by associate professor Xinjun Xu (School of Pharmaceutical Sciences, Sun Yat-sen University) (Table 1). All samples were pulverized to fine powder before passing through 60-mesh griddle. An accurately 0.3 g weighed powder sample was extracted with 50 mL methanol-water (80:20, v/v) in an ultrasonic water bath (600 W, 40 kHz) at 30 ℃ for 30 min. After cooling, the extracted solution was added with methanol-water (80:20, v/v) to the original weight. The sample solution was filtered; 5 mL of the filtrate was diluted with methanol-water (80:20, v/v) and transferred to a 50-mL volumetric flask. The solution was filtered through a 0.45-μm nylon millipore membrane and then 10 μL was injected into the chromatography.

2.3 HPLC conditions

The HPLC analysis was carried out on an SSI HPLC system (SSI, USA), consisting of a UV 6000 detector, a 1500 pump, an AS1000 autosampler and a thermostatic column compartment. The separation was performed on a Dikma Diamonsil C_{18} column (4.6 mm × 250 mm, 5 μm) preceded by Diamonsil C_{18} guard column (10 mm × 4.6 mm, 5 μm). The separation was conducted at 35 ℃ with a flow rate of 1.0

mL/min. Acetonitrile (A)-0.5% aqueous solution of acetic acid (B) system was used as the mobile phase in gradient elution mode. The elution gradient was set as follows: 15%→34% A for 40 min, 34%→38% A for 5 min, and 15% A kept for 15 min to equilibrate the system at last. The detection wavelength for fingerprint was 254 nm and on-line UV spectra were recorded within 200~400 nm.

2.4 HPLC-DAD-ESI-MS/MS conditions

The HPLC-DAD-MS/MS analysis for identifying the constituents in Flos Sophorae Immaturus was carried out on a Finnigan Liquid Chromatography coupled with a DAD detector and a TSQ quantum mass spectrometer (Thermo, USA). The mass spectrometer was operated in electrospray ionization (ESI) resource. The chromatographic conditions were described above, except that the column was maintained at ambient temperature. A union tee was used for split-flow to maintain the flow rate at 0.3 mL/min for MS. Other instrument parameters were set as follows: negative-ion mode; ESI needle voltage 3000 V; capillary temperature 270 ℃; sheath gas nitrogen (> 99% purity), 40 arbitrary units and auxiliary gas nitrogen (> 99% purity), 20 arbitrary units. The full scan mode covered the mass range from *m/z* 150 to 1000. The MS data were simultaneously acquired for the selected precursor ion. The collision-induced decomposition MS-MS experiments were performed using argon as the collision gas; the collision energy was 40 eV. Instrumental control and data acquisition was implemented on the Xcalibur 2.0 data system.

2.5 Method validation for quantitative analysis

According to the guideline of International Conference on Harmonization (ICH), the

developed method was validated for its linearity, limits of detection (LOD), limits of quantitation (LOQ), precision (inter-day and intra-day precision), repeatability, stability, and accuracy.

2.6 Chemometric analysis

Similarity analysis (SA) was performed with the help of Similarity Evaluation System for Chromatographic Fingerprint of Traditional Chinese Medicine (Version 2004A), which was recommended by the State Food and Drug Administration of China. Principal components analysis (PCA) and hierarchical clustering analysis (HCA) were performed on the SIMCA-P+ and SPSS (Version 19.0) software, respectively, to demonstrate the variability of the chromatographic fingerprinting analysis among 14 batches of Flos Sophorae Immaturus samples. Mahalanobis distance of radius on the PCA plots was calculated using MATLAB 7.0 (The MathWork Inc.). The data set was organized in a matrix with 14 lines corresponding to samples and 9 columns corresponding to the relative peak areas of the common peaks.

3 Results and discussion

3.1 Optimization of HPLC condition

HPLC conditions were optimized to obtain desired resolution. Different mobile phases, including water-acetonitrile, water-methanol, aqueous acetic acid (0.5%, v/v)-acetonitrile were tested. Good resolution and symmetric peak shape were obtained when aqueous acetic acid (0.5%, v/v)-acetonitrile was used; the final gradient was confirmed accordingly. The wavelength of 254 nm was set because all analytes had maximum absorption at this wavelength, and the baseline was well

improved on the chromatographic profiles. Column temperature was kept at 35 ℃ and the flow rate was set at 1.0 mL/min for optimal separation. The typical chromatograms of standard solution and sample solutions were shown in Figure 2.

3.2 Optimization of extraction method

In order to obtain optimum extraction efficiency, variables involved in the whole sample preparation procedure, i.e., extraction solvent, extraction method, solvent volume, and extraction time, were investigated. Extraction efficiencies were evaluated by comparing the peak area of target compounds when samples were fixed at the same weight. Different solvents (methanol, water, ethanol and ethyl acetate) were tested for good extraction efficacy and the results (Figure S1A) indicated that methanol was the best extracting solvent. Further, a series of methanol aqueous solution at different concentration (100%, 80% and 60%) was compared; 80% aqueous solution of methanol obtained the highest yield (Figure S1B). Extraction methods were also compared; ultrasonic extraction proved to be a simpler method than refluxing extraction when almost the same yields were achieved. Solvent volumes (30, 40, 50 and 60 mL) together with extraction duration (20, 30, 40 min) were optimized for the sake of environment-friendly operation and resources saving; the results were depicted in Figure S1C and D. Finally, the optimal extraction procedure, i.e. extracted with 50 mL of 80% methanol with the assistance of ultrasound for 30 min was confirmed.

3.3 Method validation of quantitative analysis

The method was validated in terms of linearity, LOD, LOQ, precision, repeatability,

stability and recovery test. Calibration was performed by analyzing three standard solutions in duplicate at six concentration levels; the calibration curves were constructed by plotting the peak areas versus the concentration of each compound. The LODs and LOQs were evaluated as S/N of 3 and 10, respectively. Regression data, LODs, and LOQs for three standard substances were given in Table 2. The precision was performed by six replicate determinations of a sample solution and the repeatability was examined by six replications of a batch. To evaluate the stability, the sample solution was injected at 0, 1, 2, 4, 8, 12, and 24 h after preparation. Variations were expressed as relative standard deviations (RSDs). The recovery test was determined by the standard addition method. Samples were prepared at three concentration levels in triplicate by spiking known quantities of each of the three standards into the Flos Sophorae Immaturus sample, and then were extracted and analyzed according to the described procedures. The validation data were shown in Table 3.

The described method possessed lower LODs as compared to those reported previously.[17,18] It was reported that HPCL-CL (chemiluminescence) was capable of quantifying trace amounts of rutin and quercetin but it required a derivation with the help of luminol and potassium ferricyanide solution.[11] In comparison, the HPLC-DAD method established in this work was simpler in analysis procedure, even though the LOD was higher.

3.4 Validation of fingerprint analysis

The proposed method for fingerprint analysis was validated in terms of precision,

repeatability, and stability of common peaks. The validation was performed on one of the 14 samples (S14) based on the relative retention times (RRTs) and relative peak areas (RPAs). The precision was assessed by analyzing samples in five replicate; the results of RRTs and RPAs expressed as RSD were less than 0.43% and 4.33%, respectively. The repeatability and stability (RSDs) were less than 1.73% and 0.73% (n = 6) for RRTs and less than 4.97% and 4.68% for RPAs. The results indicated that the fingerprinting method was reliable.

3.5 Fingerprint analysis

The chromatographic fingerprints of 14 batches of Flos Sophorae Immaturus were shown in Figure 3. The reference fingerprint was developed with the median of 14 chromatograms, as given in Figure 3 (marked with R). Nine common peaks were observed in all fingerprints, which were favorably separated under the given chromatographic condition; they totally accounted for over 95% of the total peak area in any individual chromatogram, which might be regarded as characteristic peaks. Among the 9 peaks, peaks 4, 6, and 8 were three of the most abundant components in all samples. The RRTs and RPAs of characteristic peaks related to the reference peak were calculated for the quantitative parameters of chemical properties in the chromatographic pattern analysis.

Similarity analysis was performed for evaluating the varieties of the Flos Sophorae Immaturus samples. The similarity values of chromatograms were calculated by the similarity evaluation system, which was based on correlative coefficient calculations and recommended by the State Food and Drug Administration of China. As detailed in

Table 1, the similarity values of different samples ranged from 0.927 to 0.983, which indicated that different samples shared similar chromatographic patterns (similarity values ≥ 0.900). On the other hand, even the samples from the same province, especially Henan province, were different from each other. The similarity values of S4 and S14, i.e., 0.927 and 0.938, respectively, were lower than that of others (similarity values ≥ 0.950), suggesting that these two samples might be different from others. From the chromatographic profiles, it was considered that S4 had lower peak signal overall than that of other samples while S14 had more prominent peak signal. The similarity analysis demonstrated that dissimilarities in the average components ratios and the total integrated areas resulted in low similarity values, even if all samples presented similar chromatographic fingerprint profile.

3.6 Identification of chromatographic peaks in Flos Sophorae Immaturus by HPLC-DAD-ESI-MS/MS

In order to gain more information about constituents in Flos Sophorae Immaturus, it is significant to identify their structures, especially the characteristic peaks in the fingerprint. S14, originated from Henan Province, a famous cultivation area, was chosen for the identification study. The technique used was HPLC-DAD-ESI-MS/MS, which was reported the first use for the identification of constituents in Flos Sophorae Immaturus. The MS and UV data of constituents in Flos Sophorae Immaturus extract were summarized in Table 4. The retention time, UV and mass spectral data of the components were compared to those of available references.[19-23] Most constituents in Flos Sophorae Immaturus were successfully identified. Quercetin and its derivatives

accounted for most constituents in Flos Sophorae Immaturus, followed by isorhamnetin and kaempferol derivatives. Based on the fragment information of flavonoids, the typical fragmentation patterns of quercetin and isorhamnetin glycosides were depicted in Figure S2A and Figure S2B, respectively. On the whole, 14 flavonoids were identified in Flos Sophorae Immaturus by HPLC-DAD-ESI-MS/MS. The successful identification of those components was of great importance in establishing a fingerprinting method for the quality control of this medicine.

3.7 Principal component analysis (PCA)

Principal component analysis (PCA), a multivariate analysis technique, could visualize similarities or differences within multivariate data.[24] It was employed to analyze the fingerprints of Flos Sophorae Immaturus. The RPAs of nine characteristic peaks were set as variables, while 14 batches of samples were set as observations. PC1 explained 42.9% of the total variance in the data set while PC2 explained 25.8%. The first two PCs reduced the multidimensional dataset to a two-dimensional dataset. The scores plot of PCA generated from the 14 samples, as well as the loading plot, was shown in Figure 4. According to the loading plot, PC1 showed a positive correlation with peak 1(quercetin-3-*O*-*β*-glucosyl-(1→2) [*α*-rhamnosyl-(1→6)]-*β*-glucoside), 3(saluenin), 4(rutin), 6(narcissin) and 7(unknown compound) while PC2 showed a positive correlation with peak 8(quercetin) and 9(isorhamnetin). Then the score matrix was typed in MATLAB 7.0 and the Mahalanobis distance of the score points was calculated by the given command

(Figure S3).[25,26] The output was listed in Figure S4. The first column represented the Mahalanobis distance between each sample point and the origin (radius), and the field included by the blue box represented the Mahalanobis distance between each sample point. The radius of S4 and S14 were both of long distance (over 2, marked in red on Figure 4A), indicating they were discrete. It could be confirmed by the lower similarity values listed in Table 1. Additionally, they were far away from each other in PCA score plot (Mahalanobis distance 3.10), indicating differences in chemical characteristics even if they resemble each other from similarity analysis. The radius of S1, 3, 5, 8, 11, 13 (marked in blue on Figure 4A) was less than 1. The radius of the rest samples (marked in black on Figure 4A) was between 1 and 2. Also, S4 and S14 were significantly far away from others in respect of the Mahalanobis distance between each sample point, ($p < 0.05$). Thus, it was notable that all samples were clearly classified into three groups, including group I (S4 from Henan province), group II (S14 from Henan province), and group III. Group III, the largest one, was composed of the other 12 batches of samples. In particular, samples from Henan province were clustered in different groups, which implied that PCA enabled reliable discrimination and quality control of Flos Sophorae Immaturus samples even if they were growing under the same geography condition. Therefore, PCA could be a useful tool in discriminating samples which resemble each other in similarity value. However, the relationship of growing environment with corresponding quality of Flos Sophorae Immaturus samples could not be verified in this study for insufficient widely-distributed samples and it still needed to be further investigated with more

samples from different localities.

3.8 Hierarchical cluster analysis (HCA)

Hierarchical cluster analysis (HCA), one of the most commonly used unsupervised pattern recognition methods, was a useful multivariate statistic technique to assign a data set into groups by creating a cluster tree or dendrogram, according to similarity.[15] In order to assess the resemblance and differences of these samples, HCA of Flos Sophorae Immaturus samples was performed based on the RPAs of all the nine characteristics chromatographic peaks by SPSS software. The Ward's method was applied as the amalgamation rule and the squared Euclidean distance was selected to measure the resemblance and classify the 14 samples. The results obtained following HCA were shown in a dendrogram (Figure 5), in which three well-defined clusters were visible. S14 was categorized into cluster I; S3, 8, 10, 1, 11, 13 and 4 were categorized into in cluster II while the rest in cluster III. It seemed that sample clustering was not coincident with their localities. However, the fact that cluster I was only made up of S14 was in accordance with the PCA result, in which S14 was far away from others. Similarly, it was noteworthy that S4 was separated on its own when cluster II was further divided into two sub-clusters. This was quite consistent with the results of SA and PCA. Additionally, S4 and S14 were significantly distinct from each other in PCA score plot but a bit close in the dendrogram. It was mainly due to that PCA and HCA were based on different principles and objectives.

3.9 Quantitative analysis of investigated compounds in Flos Sophorae Immaturus

As can be seen from the PCA results, rutin, narcissin and quercetin were three of the responsible variables in sample classification and discrimination. Notably, they possessed high abundance in Flos Sophorae Immaturus samples from the chromatogram profile. Therefore, it was significant to quantify them. The established method was applied to the simultaneous determination of these three flavonoids in 14 samples of Flos Sophorae Immaturus. Each sample was analyzed in triplicate to determine the mean contents (mg/g) of three selected constituents. The results were shown in Table 5 and Figure 6. Rutin was found to be predominant among the three determined analytes, ranging from 158.51 to 219.35 mg/g. The content of narcissin and quercetin ranged from 15.85 to 28.41 mg/g and from 3.00 to 6.28 mg/g, respectively. The content of each analyte varied at some level among the different samples, which could give some explanation quantitatively for the quality variance of the medicine.

4 Conclusions

In this study, a powerful and reliable HPLC-DAD method was developed for the comprehensive quality evaluation of Flos Sophorae Immaturus for the first time. The proposed method combined fingerprint profiling of the herb and quantitative analyzing of three major flavonoid components in the herb, i.e. rutin, narcissin and quercetin; HPLC-DAD-ESI-MS/MS was further used to rapidly identify the characteristic peaks and other constituents in the herb. The results indicated that Flos Sophorae Immaturus from different regions shared a similar HPLC pattern; all of them contained 9 characteristic peaks and possessed high concentration of the above

three flavonoids but differed in contents. Based on the fingerprints, 14 batches of Flos Sophorae Immaturus samples were classified or discriminated by chemometric tools (SA, PCA and HCA) objectively and successfully. The study demonstrated that the developed method was efficient and reliable, which could be readily utilized as a more significant tool than current one for the comprehensive quality control of Flos Sophorae Immaturus. The operability and the analytical capacity offered by the developed HPLC-DAD and HPLC-DAD-ESI-MS/MS methods enabled their adoption as powerful analytical fingerprint techniques, especially in combination with chemometric tools. In conclusion, the fingerprint analysis combined with multi-components determination was proved to be an efficient and comprehensive tool for the quality control of TCM.

Acknowledgements

This work was supported by the planning subject of 'The Twelfth Five-Year-Plan' in National Science and Technology for the Rural Development in China: The development and application of functional plants extract (No: 2013BAD10B04).

References

1 Ha, D. T., Trung, T. N., Phuong, T. T., Yim, N., Chen, Q. C., Bae, K. *Bioorg. Med. Chem. Lett.* 2010, **20**, 6076-6081.

2 Kim, J., Yun-Choi, H. *Arch. Pharm. Res.*, 2008, **31**, 886-890.

3 Tang, Y. P., Li, Y. F., Hu, J., Lou, F. C. *J. Asian. Nat. Prod. Res.* 2002, **4**, 123-128.

4 Lo, Y. H., Lin, R. D., Lin, Y. P., Liu, Y. L., Lee, M. H. *J. Ethnopharmacol.*, 2009, **124**, 625-629.

5 Liu, R., Qi, Y., Sun, A., Xie, H. *J. Sep. Sci.,* 2007, **30**, 1870-1874.

6 Commission, C. P., *Chinese Pharmacopoeia*, China Medical Science Press, Beijing 2010.

7 Wu, Q. Y., Zhou, Y., Jin, X., Guan, Y., Xu, M., Liu, L. F. *Int. J. Mol. Sci.,* 2011, **12**, 8740-8749.

8 Gan, Z., Chen, Q., Fu, Y., Chen, G. *Food Chem.,* 2012, **130**, 1122-1126.

9 Huang, J., Xie, W. *Anal. Sci.,* 2010, **26**, 383-386.

10 Li, F. J., Ning, S. L., Li, Y., Yu, Y. J., Shen, C. D., Duan, G. L. *Phytochem. Anal.,* 2012, **23**, 292-298.

11 Wu, H., Chen, M., Fan, Y., Elsebaei, F., Zhu, Y. *Talanta,* 2012, **88**, 222-229.

12 Xu, J., Zhang, H., Chen, G. *Talanta,* 2007, **73**, 932-937.

13 Zhou, X., Zhao, Y., Lei, P., Cai, Z., Liu, H. *J. Sep. Sci.,* 2010, **33**, 2258-2265.

14 Bian, Q., Yang, H., Chan, C. O., Jin, D., Mok, D. K., Chen, S. *Chem. Pharm. Bull.,* 2013, **61**, 802-808.

15 Liao, H. J., Lai, Z. Q., Su, J. Y., Yi, Y. Y., Li, Y. C., Lai, X. P., Su, Z. R., Lin, Z. X. *J. Sep. Sci.,* 2012, **35**, 2193-2202.

16 Avula, B., Wang, Y. H., Ali, Z., Smillie, T. J., Khan, I. A. *Biomed. Chromatogr.,* 2013, DOI: 10.1002/bmc.3019.

17 Li, X., Zhang, Y., Yuan, Z. *Chromatographia,* 2002, **55**, 767-767.

18 Deng, F., Chen, B., Chen, G., Liang, S., Zhou, P., Xia, L. *Bull Hunan Med Univ,* 2000, **25**, 597-598.

19 Atoui, A. K., Mansouri, A., Boskou, G., Kefalas, P. *Food Chem.,* 2005, **89**, 27-36.

20 Inbaraj, B. S., Lu, H., Kao, T. H., Chen, B. H. *J. Pharm. Biomed.,* 2010, **51**, 549-556.

21 Xie, Z., Zhao, Y., Chen, P., Jing, P., Yue, J., Yu, L. *J. Agr. Food Chem.,* 2011, **59**, 3042-3049.

22 Kite, G. C., Veitch, N. C., Boalch, M. E., Lewis, G. P., Leon, C. J., Simmonds, M. S. J. *Phytochemistry.,* 2009, **70**, 785-794.

23 Zhao, H. Y., Sun, J. H., Fan, M. X., Fan, L., Zhou, L., Li, Z., Han, J., Wang, B. R., Guo, D. A. *J. Chromatogr. A,* 2008, **1190**, 157-181.

24 Niu, C., Zhang, J. *J. Sep. Sci.,* 2011, **34**, 2989-2996.

25 Mahalanobis, P. C. *Proc. Nat. Inst. Sci. India,* 1936, **2**, 49-55.

26 De Maesschalck, R., Jouan-Rimbaud, D., Massart, D. L. *Chemometr. Intell. Lab,* 2000, **50**, 1-18.

Figure captions

Graphical abstract: A fingerprinting and quantifying method was developed for the quality evaluation of Flos Sophorae Immaturus by HPLC-DAD and HPLC-DAD-ESI-MS/MS.

Figure 1 The chemical structures of investigated compounds.

Figure 2 Typical HPLC Chromatograms of the mixture of standards (A) and sample solution (B), 1: rutin; 2: narcissin; 3: quercetin.

Figure 3 HPLC fingerprints of 14 batches of samples (R: Reference).

Figure 4 Scores (A) and loading plots (B) of PCA. Numbers in the loading plots represents the peak number depicted in Figure 3.

Figure 5 The dendrogram of 14 batches of Flos Sophorae Immaturus samples by HCA

Figure 6 Contents of three flavonoid constituents in Flos Sophorae Immaturus from different localities.

Tables

Table 1 Localities and similarity values of Flos Sophorae Immaturus collected from different geographic regions in China.

Batch No.	Supplier	Locality (Province)	Similarity value
S1	Qingping pharmacy	Hebei	0.968
S2	Baicao Pharmacy	Shandong	0.954
S3	Jianmin Pharmacy	Hunan	0.978
S4	Jianmin Pharmacy	Henan	0.927
S5	Dashenlin Pharmacy	Hebei	0.983
S6	Baozhilin Pharmacy	Henan	0.962
S7	Laobaixing Pharmacy	Hunan	0.960
S8	Qingping pharmacy	Shanxi	0.969
S9	Zikuan Pharmacy	Henan	0.961
S10	Fuzhou Pharmacy	Henan	0.957
S11	Tongyi Pharmacy	Henan	0.953
S12	Yizhou Pharmacy	Henan	0.961
S13	Hongyuan Pharmacy	Henan	0.976
S14	Shunfa Pharmacy	Henan	0.938

Table 2 Linear regression data, LODs and LOQs of the investigated compounds.

Compound	Calibration Equation y = ax + b [a]	R^b	Linear Range (μg/mL)	LOQ (μg/mL)	LOD (μg/mL)
Rutin	y = 84383030.5x+356120.6	0.9995	10.9~271.5	0.50	0.25
Narcissin	y = 51661816.3x+19581.1	0.9998	2.6~130.0	0.65	0.43
Quercetin	y = 163903723.4x-46335.8	0.9999	0.8~40.0	0.28	0.13

(a) y and x are, respectively, the peak areas and concentrations (μg/mL) of the analytes.

(b) R = correlation coefficient, n = 6.

Table 3 Precision, repeatability, stability and recovery of three reference substances.

Compound	Precision RSD% (*n*= 6)		Repeatability RSD% (*n*=6)	Stability RSD% (*n*=7)	Recovery (%)[a] (*n*=9)	
	Inter-day RSD%	Intra-day RSD%			Mean	RSD%
Rutin	0.63	1.44	0.54	1.26	100.72	2.03
Narcissin	0.42	1.46	0.70	1.73	100.28	1.96
Quercetin	0.66	1.87	1.70	1.74	101.08	1.53

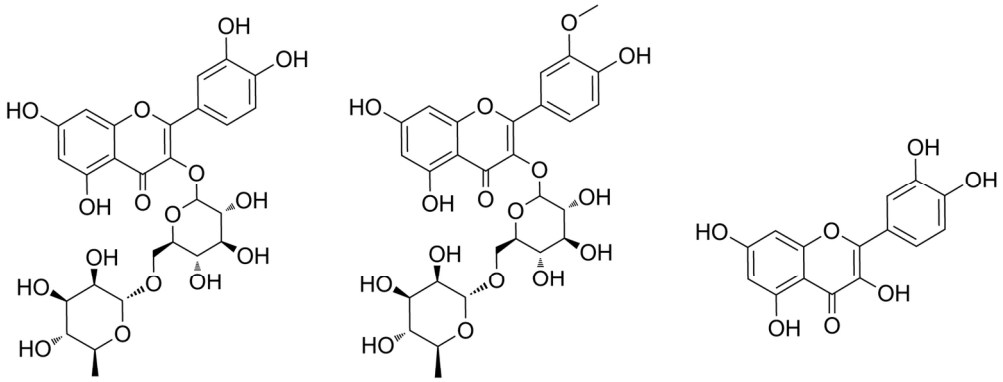
(a) Recovery (%) = 100 × (amount found-original amount)/amount spiked.

Table 4 HPLC-DAD-ESI-MS/MS data of identified and tentatively characterized compounds from Flos Sophorae Immaturus.

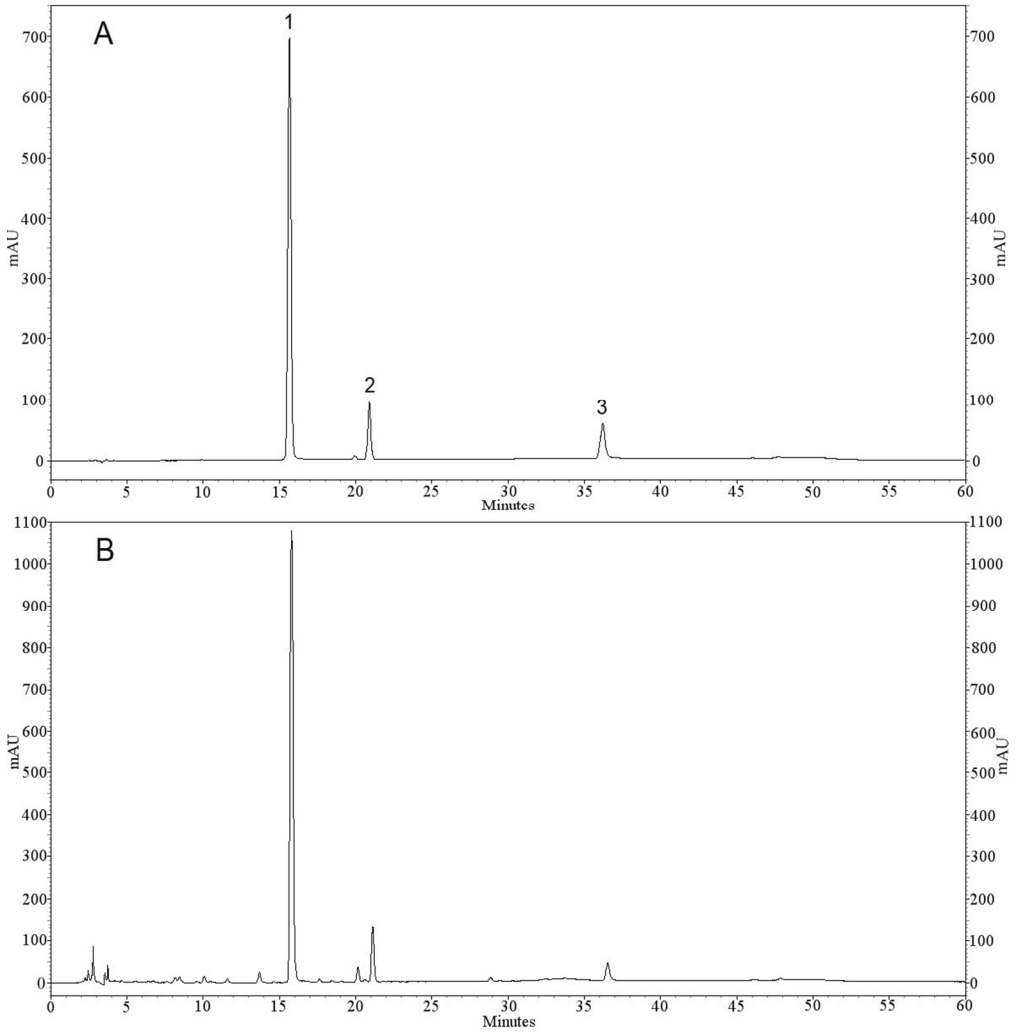
No.	Rt (min)	UV (nm)	$[M-H]^-$ *(m/z)*	MS/MS *(m/z)*	Identity
1	3.06	264	193.19	179.19, 161.01, 113.12, 101.33, 86.36, 73.27, 59.40, 45.17	2-*O*-methyl-inositol
2	3.47	260, 348	332.88	289.24	Unknown
3	4.42	236	191.15	110.98, 87.26, 85.16, 67.11, 39.38	Citric acid
4	5.13	270	251.35		Unknown
5	8.23	254, 356	771.46	590.87, 299.86	Quercetin-3-*O*-glucosyl-(1→2)-glucoside-7-*O*-rhamnoside
6	13.04	254, 344	608.84	300.34	Quercetin 3-*O*-rhamnosylglucoside
7	13.32	254, 354	771.46	609.20, 300.07, 271.02, 255.00, 243.10, 179.02	Quercetin-3-*O*-*β*-glucosyl-(1→2)[*α*-rhamnosyl-(1→ 6)]-*β*-glucoside
8	14.54	254, 354	755.47	609.02, 446.16, 298.99, 270.92, 228.61	Quercetin-3-*O*-rutinoside-7-*O*-rhamnoside
9	16.16	250, 354	785.44	337.76, 314.94, 299.09, 270.70, 255.17	Isorhamnetin-3-*O*-β-glucosyl-(1→2)[α-rhamnoxyl-(1→6)]-*β*-glucoside
10	18.52	252, 354	741.43	299.93, 270.88, 254.90, 243.12, 210.99, 201.66, 151.10	Quercetin-3-*O*-xylosyl-(1→2)-rhamnosyl-(1→6)-glucoside (Saluenin)
11	20.66	254, 354	609.34	299.99, 270.96, 254.96, 242.98, 227.09, 178.94, 151.02	Quercetin-3-*O*-rutinoside
12	22.75	250, 354	755.45	315.04, 299.77, 271.14, 243.32	Isorhamnetin-3-*O*-xylosyl-(1→2)-glucoside-7-rhamnoside
13	22.75		463.19	301.09, 270.78, 179.02	Quercetin-3-*O*-glucoside
14	24.73	264, 346	593.35	284.04, 255.10, 226.91	Kaempferol-3-*O*-rutinoside
15	25.37	254, 354	623.35	314.60, 298.96, 284.7, 270.99, 243.04, 226.85, 199.01	Isorhamnetin-3-*O*-rutinoside
16	28.96	318	964.04	903.90, 676.63, 451.57, 338.29, 225.15	Unknown
17	32.76	242,314	599.27	475.52, 433.35, 159.08, 145.07, 118.85	Unknown
18	36.71	242, 310	431.17		Sophoricoside
19	42.71	254, 368	301.16	254.71, 227.27, 151.01, 120.79	Quercetin
20	49.78	254, 360	315.11	300.01, 271.40, 243. 30	isorhamnetin

Table 5 Contents of three flavonoid compounds in Flos Sophorae Immaturus collected from different localities (n=3).

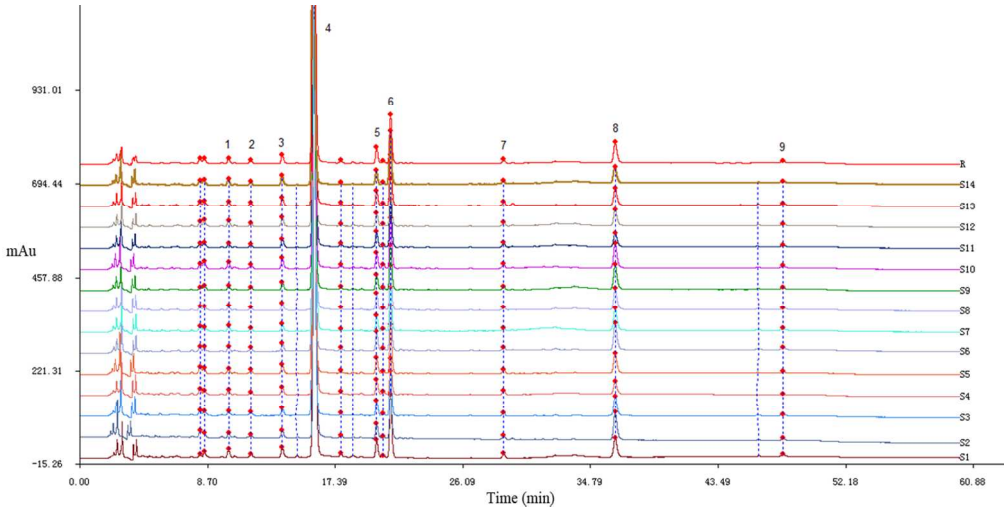
Batch	Locality	Content of investigated compounds (mg/g, mean ± SD)		
		Rutin	Narcissin	Quercetin
S1	Hebei	160.11 ±2.44	19.26 ±0.17	6.38 ±0.09
S2	Shandong	190.60 ±0.71	20.76 ±0.79	4.59 ±0.08
S3	Hunan	189.84 ±1.92	15.85 ±0.10	3.00 ±0.05
S4	Henan	169.06 ±1.84	20.17 ±0.13	4.35 ±0.08
S5	Hebei	158.51 ±0.31	18.43 ±0.16	5.53 ±0.05
S6	Henan	164.07 ±1.02	20.28 ±0.11	5.34 ±0.05
S7	Hunan	179.12 ±1.17	19.35 ±0.27	4.81 ±0.06
S8	Shanxi	173.76 ±0.89	19.84 ±0.29	5.53 ±0.05
S9	Henan	180.01 ±0.56	21.60 ±0.18	5.30 ±0.06
S10	Henan	187.27 ±0.59	20.95 ±0.11	3.33 ±0.06
S11	Henan	180.50 ±0.14	23.37 ±0.08	4.32 ±0.03
S12	Henan	198.62 ±1.85	23.21 ±0.10	3.85 ±0.05
S13	Henan	191.11 ±0.58	23.99 ±0.15	4.19 ±0.10
S14	Henan	219.35 ±2.33	28.41 ±0.31	4.30 ±0.04



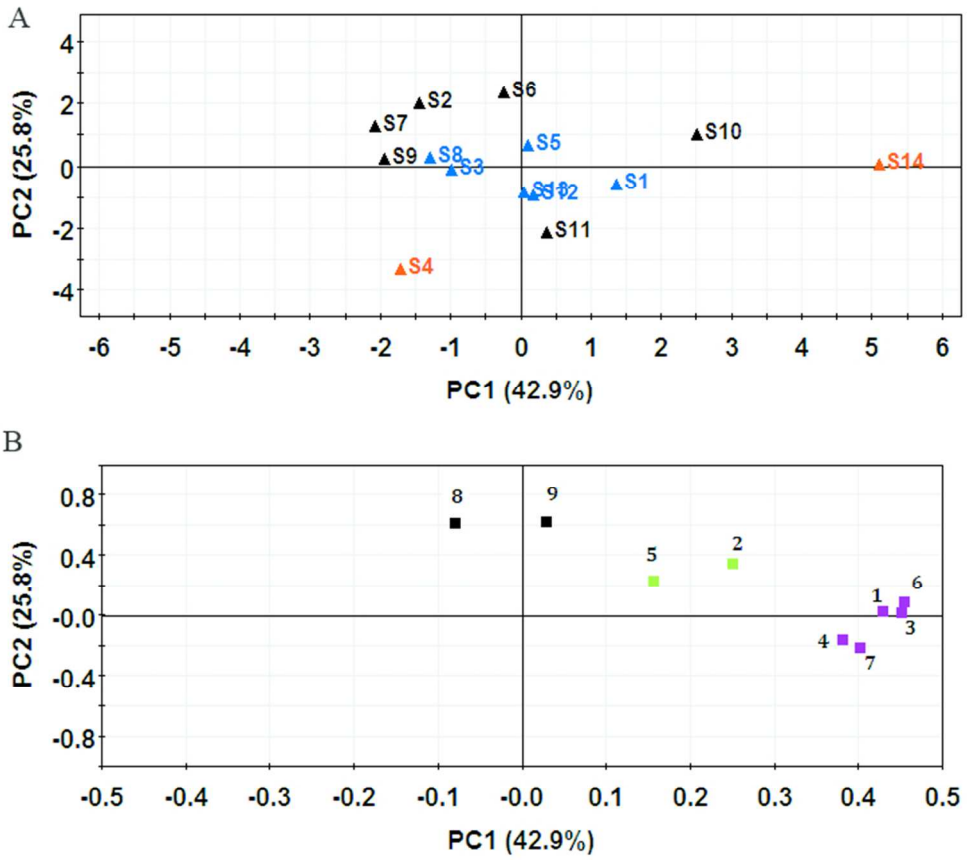
The chemical structures of investigated compounds.
72x33mm (600 x 600 DPI)



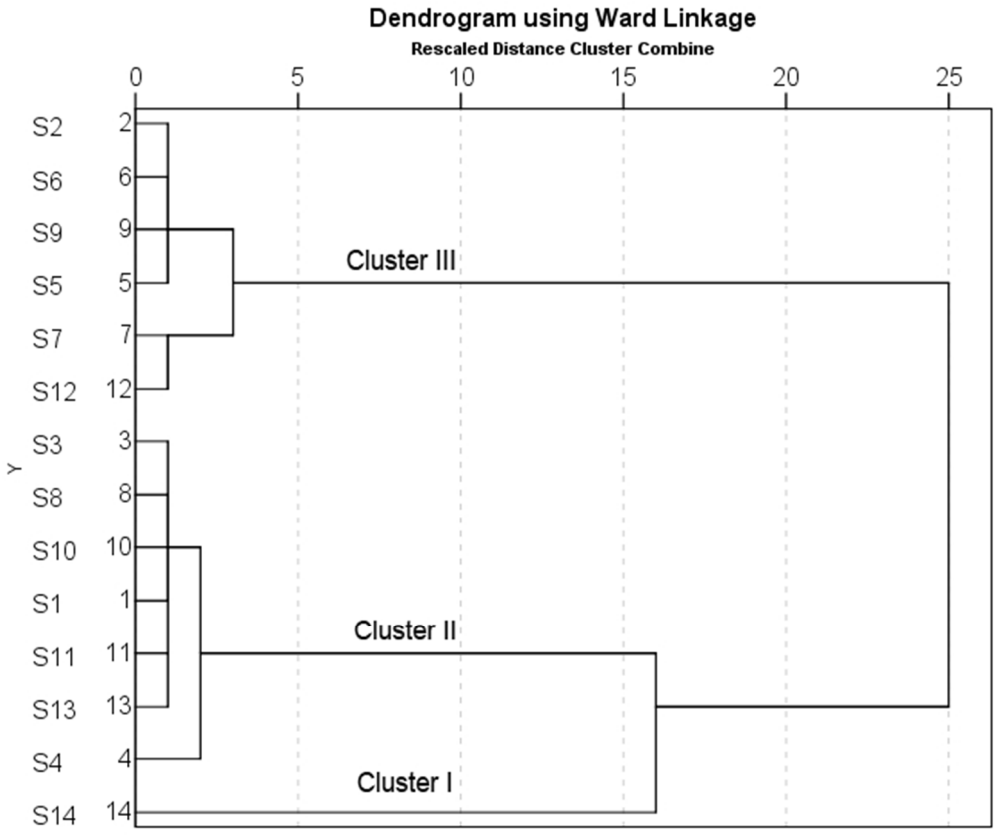
Typical HPLC Chromatograms of the mixture of standards (A) and sample solution (B), 1: rutin; 2: narcissin; 3: quercetin
100x105mm (300 x 300 DPI)



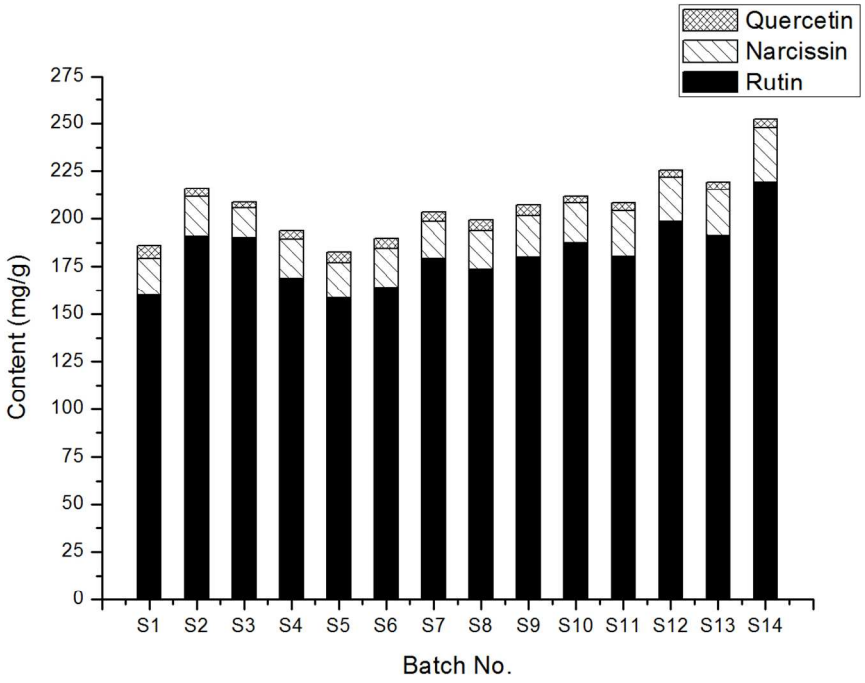
HPLC fingerprints of 14 batches of samples (R: Reference).
110x55mm (300 x 300 DPI)



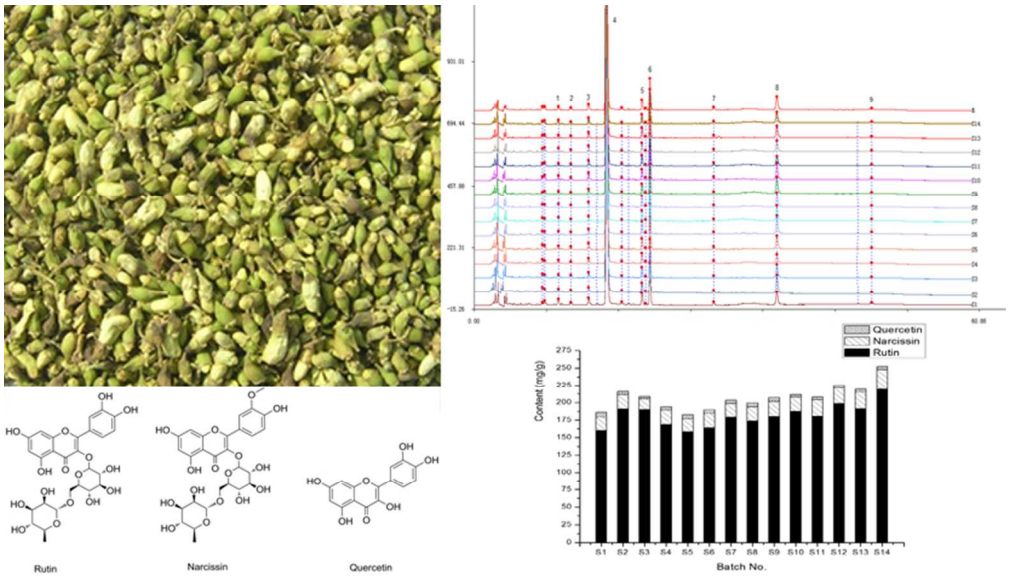
Scores (A) and loading plots (B) of PCA. Numbers in the loading plots represents the peak number depicted in Figure 3.
63x56mm (300 x 300 DPI)



The dendrogram of 14 batches of Flos Sophorae Immaturus samples by HCA
229x194mm (72 x 72 DPI)



Contents of three flavonoid constituents in Flos Sophorae Immaturus from different localities.
119x84mm (300 x 300 DPI)



A fingerprinting and quantifying method was developed for the quality evaluation of Flos Sophorae Immaturus by HPLC-DAD and HPLC-DAD-ESI-MS/MS.
70x39mm (300 x 300 DPI)