Journal of Chromatography A, 1217 (2010) 1414–1418

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Short communication

Simple quality assessment approach for herbal extracts using high performance liquid chromatography-UV based metabolomics platform

Ke Lan*, Yan Zhang, Junyi Yang, Liang Xu

Key Laboratory of Drug Targeting and Drug Delivery Systems, Ministry of Education, No. 17, Section 3, Renminnan Road, West China School of Pharmacy, Sichuan University, Chengdu 610041, Sichuan, China

ARTICLE INFO

Article history:
Received 26 October 2009
Received in revised form 3 December 2009
Accepted 9 December 2009
Available online 16 December 2009

Keywords:
Metabolomics
Herbal extract
Quality control
HPLC-UV

ABSTRACT

A lack of adequate or accepted research methodology has been a major obstacle to study herbal medicines. In this study, instead of the prevalent hyphenated chromatographies, common high performance liquid chromatography equipped with ultraviolet detector (HPLC-UV) and multivariate statistical analysis were utilized to assess the qualities of total flavones of sea buckthorn (TFS), an 85% ethanol extract of the sea buckthorn berries. Two complementary HPLC-UV methods were developed, validated and combined to comprehensively determine the ingredients in TFS. Principal component analysis (PCA) and partial least square-discriminant analysis (PLS-DA) of the combined analytical data showed that the six batches of TFS could be well differentiated. Hierarchical cluster analysis (HCA) using Ward's minimum variance method of the PLS-DA loading matrix demonstrated the known ingredients (quercetin, kaempferol, isorhamnetin, oleanolic acid and ursolic acid) and three unknown ingredients in TFS significantly contributed to the quality differences. A PLS regression model indicated that the results of the present method correlated well with the content of total flavones, which is now the quality control approach of TFS. Results from this study indicated that the proposed method is reliable for the quality reassessment of some widely used herbal extracts.

1. Introduction

Quality control of herbal medicines (HMs) is a big headache because of the high complexity and sequential unknown mechanism on disease treatment [1]. Leaving the facts alone, some herbal products have been commercially used world wide, especially in Asian nations for decades. These products are commonly quality-controlled by determining a total content of a group of ingredients (such as total flavones, total alkaloids and total saccharides), or the contents of several pharmacologically active ingredients. These approaches are of course far from adequate. Lack of adequate or accepted research methodology for evaluating HMs has been a major obstacle to the globalization of HMs.

Chromatographic fingerprint has been an acceptable technique for quality control of HMs both in theory and practice [1,2]. Metabolomics platform has also been recently introduced into the quality control of HMs [3–6]. These two techniques have different terms, but are nearly the same analytical process. Metabolomics platform exhibits more powerful capability in multivariate statistical analysis than fingerprint technique does. Hyphenated chromatographies, such as UPLC-MS, LC-MS and GC-MS are commonly recommended for the purpose of quality control of HMs, since they may provide information about the chemical structures of the analytes [2]. However, their applications are extremely hindered by the expensive instruments and highly specialized techniques. May commonly available analytical strategies, such as high performance liquid chromatography equipped with ultraviolet detector (HPLC-UV), be used for metabolomics based quality assessment of HMs? This paper aimed to answer this question.

Total flavone of sea buckthorn (TFS) is an 85% ethanol extract of the sea buckthorn berries. TFS has been clinically used in China since 1980, for treatments of cardiovascular disorders. TFS is now quality-controlled by determining the content of total flavones with spectrophotometry at 430 nm (isorhamnetin as reference) after chromogenesis with aluminum trichloride. Our previous study has identified several pharmacologically active ingredients in TFS, such as flavonoids (quercetin, kaempferol and isorhamnetin) and triterpenoids (oleanolic acid and ursolic acid) (published in Chinese). How different are the qualities of batches of TFS? How do these known and other unknown ingredients contribute to the quality differences? In the present study, different batches of TFS are employed to deliver the strategies and performance of HPLC-UV based metabolomics approach in the quality assessment of herbal extracts.

* Corresponding author. Tel.: +86 028 85503722.
E-mail address: lanwoco@scu.edu.cn (K. Lan).

0021-9673/$ – see front matter.
doi:10.1016/j.chroma.2009.12.031

2. Materials and methods

2.1. Chemicals, reagents, and standards

Methanol was purchased from Burdick & Jackson (SK Chemicals, Ulsan, Korea). Formic acid (≥96% purity) and ammonium acetate (≥97% purity) were purchased from TEDIA (Fairfield, OH, USA). Quercetin (97.3%, CAS Registry No. 117-39-5) was obtained from National Institute for the Control of Pharmaceutical & Biological Products (Beijing, China). Isorhamnetin (99%, CAS Registry No. 480-19-3), kaempferol (99%, CAS Registry No. 520-18-3), ursolic acid (>98%, CAS Registry No. 77-52-1) and oleanolic acid (>95%, CAS Registry No. 508-02-1) were obtained from TAUTO Biotech (Shanghai, China).

2.2. Sample preparations

Six batches of finely powdered TFS (batch A–F) were kindly gifted by Sichuan Chuanda West China Pharmaceutical Co. Ltd. The content of total flavones was determined by the manufacturer (11.38%, 10.34%, 13.68%, 10.58%, 11.50% and 12.49% for batch A–F, respectively). 20 ± 0.5 mg of the TFS powder was extracted with 10 mL of methanol using an ultrasonic bath for 20 min at 50 °C. One volume of the solution was diluted with one volume of water, and filtered with a 0.22-μm-porosity nylon syringe filter as sample I. Sample I was mainly used for analysis the flavonoids in TFS. The solution was directly filtered without dilution as sample II. Sample II was mainly used for analysis the triterpenoids and phytosterols. 100 μL aliquots from every sample I or sample II were combined in a representative pool sample as QC sample I or QC sample II, respectively [7–10]. An aliquot of 20 μL of each sample solution was injected into the column for analysis.

2.3. Chromatographic analysis

High performance liquid chromatography (HPLC) analyses were performed on an Agilent 1200 series HPLC-UV system. An Agilent ZORBAX Eclipse XDB C18 column (4.6 mm i.d. × 150 mm, particle size 5 μm) was maintained at 35 °C. The aliquots of sample I were gradient-eluted at 1.000 mL/min using (A) 5 mM ammonium acetate/0.2% formic acid in water and (B) methanol (35–45% B for 10 min, 45–50% B in 20 min, 35% B in 20 min), and detected at 254 nm. The aliquots of sample II were isocratic-eluted at 1.000 mL/min using 89% B for 15 min, and detected at 210 nm. In the analytical run of sample I or sample II, QC sample I or QC sample II was injected for three times at the beginning to pre-condition the system, and inserted every 1–3 samples to monitor the potential drift of chromatographic performance. The sequences of samples injected in the analytical run of sample I and sample II were the same. The peaks of quercetin, kaempferol, isorhamnetin, ursolic acid and oleanolic acid were identified with the corresponding reference standards.

2.4. Data processing and multivariate analysis

The chromatographic data was analyzed by the Agilent ChemStation for LC systems (Revision B.04.01). Peak 10 (unknown) and peak 27 (isorhanmentin) were selected as references to align the retention time (RT) of chromatograms of sample I. Peak 32 (unknown) and peak 40 (ursolic acid) were used to align the chromatograms of sample II. The peaks integration threshold of the sample II was set much higher than that of sample I, because of a high baseline noise. The parameters used for peaks integration of the sample I and sample II were RT 3–30 min and 2–15 min, slope sensitivity 1, width 0.01 min and 0.1 min, minimum area 5 and 100, minimum height 0.3 and 10, advanced baseline calibration mode, vertical shoulder peak mode. No specific peak was excluded. 28 and 12 peaks were identified in sample I and sample II and included in the dataset I and dataset II, which were summarized in Table 1. Provided that the analytical run was validated by monitoring the QC samples, dataset I and dataset II could be combined in sequence as the combined dataset. The resulting datasets containing arbitrarily assigned peak index, retention time, and peak area were further exported to SIMCA-P+ software 12.0 (Umetrics, Umeå, Sweden) for multivariate statistical analysis according to a recent report based on ultra-performance LC/TOF MS analysis [3].

Previous to multivariate statistical analysis, procedures that can be used for data pretreatment are scaling, centering and transformations. Centering converts all the variables to fluctuations around zero instead of around their means, and thus focuses on the differences and not the similarities in the data [11]. Pareto scaling uses square root of the standard deviation as the scaling factor. With Pareto scaling, principal component analysis (PCA) amplifies the contribution of lower concentration variables but not to such an extent where noise produces a large contribution [3,11]. Centering and Pareto scaling were performed to pre-treat the data sets resulting from the above samples. No specific data transformation was used. PCA was initially used to visualize general clustering, trends, and outliers among the observations. For further identifying the peaks accountable for the separation between six batches of TFS, partial least square-discriminant analysis (PLS-DA) and hierarchical cluster analysis (HCA) using Ward's minimum variance method were sequentially carried out.

3. Results and discussion

3.1. HPLC method development

Flavonoids exhibit strong UV absorbance at about 254 nm and 350–370 nm [12], while triterpenoids and phytosterols only exhibit weak UV absorbance at 200–210 nm [13]. It is difficult to separate them all in a single HPLC-UV method, due to the dramatic baseline shift at low wavelength. Therefore, a gradient elution was developed to separate the flavonoids in TFS at 254 nm, and another isocratic elution was optimized to separate the triterpenoids and phytosterols at 210 nm. The two methods were complementary and their chromatograms may be combined. Fig. 1 clearly shows that the chromatographic shift of the gradient elution of sample I was more significant than the one of the isocratic elution of sample II, indicating that the methods should be carefully validated with multivariate analysis of the data from QC samples. The developed HPLC methods may also be applied to LC-MS analysis of TFS.

TFS is insoluble in water and soluble in methanol. In TFS, flavonoids are more polar than triterpenoids and phytosterols. Flavonoids were separated by gradient elution, with which samples in pure methanol produced significant leading peaks in the preliminary test, while samples in 50% methanol demonstrated good peak shapes. However, only small amount of flavonoids was dissolved and detected if TFS was directly extracted with 50% methanol using an ultrasonic bath for 20 min at 50 °C. Therefore, TFS was firstly extracted with methanol and subsequently diluted with 1 volume of water. Triterpenoids and phytosterols were separated at 210 nm, at which a big baseline noise was detected. To ensure a higher signal/noise ratio, the methanol solution of TFS was analyzed without dilution. Both sample preparations were validated to be efficient to extract the analytes from TFS (data was no shown).

3.2. Validation of analytical run

The robustness of analytical method plays a key role in metabolomics study. The results of multivariate statistical analysis

Table 1
Summary of the 40 detected peaks in the six batches of total flavones of sea buckthorn ($n = 12$).

Peak index	Retention time (mean ± sd, min)	Peak area		Substances	Peak index	Retention time (mean ± sd, min)	Peak area		Substances
		Mean ± sd	RSD				Mean ± sd	RSD	
1	3.28 ± 0.06	9 ± 2	26%	Unknown	21	18.90 ± 0.02	8 ± 2	21%	Unknown
2	3.49 ± 0.06	17 ± 4	24%	Unknown	22	19.70 ± 0.03	10 ± 2	17%	Unknown
3	3.75 ± 0.06	35 ± 8	23%	Unknown	23	20.44 ± 0.04	9 ± 2	20%	Unknown
4	4.09 ± 0.03	6 ± 0	7%	Unknown	24	21.88 ± 0.02	13 ± 5	38%	Unknown
5	4.66 ± 0.04	30 ± 14	45%	Unknown	25	23.50 ± 0.02	129 ± 28	21%	Unknown
6	5.67 ± 0.03	40 ± 18	44%	Unknown	26	24.38 ± 0.01	509 ± 65	13%	Kaempferol
7	7.61 ± 0.01	14 ± 6	42%	Unknown	27	26.95 ± 0.00	6013 ± 386	6%	Isorhamnetin
8	8.21 ± 0.01	13 ± 5	40%	Unknown	28	28.10 ± 0.01	178 ± 19	11%	Unknown
9	8.91 ± 0.01	24 ± 7	28%	Unknown	29	2.25 ± 0.00	291 ± 31	11%	Unknown
10	9.33 ± 0.00	434 ± 345	79%	Unknown	30	2.57 ± 0.00	684 ± 73	11%	Unknown
11	9.81 ± 0.03	36 ± 3	7%	Unknown	31	2.80 ± 0.00	951 ± 158	17%	Unknown
12	10.45 ± 0.02	88 ± 11	13%	Unknown	32	4.29 ± 0.00	353 ± 25	7%	Unknown
13	11.18 ± 0.03	7 ± 2	23%	Unknown	33	4.56 ± 0.00	702 ± 73	10%	Unknown
14	11.53 ± 0.03	48 ± 16	32%	Unknown	34	5.05 ± 0.00	203 ± 18	9%	Unknown
15	11.89 ± 0.03	11 ± 5	44%	Unknown	35	5.45 ± 0.00	172 ± 20	12%	Unknown
16	13.14 ± 0.19	67 ± 31	46%	Unknown	36	5.92 ± 0.02	232 ± 71	30%	Unknown
17	13.51 ± 0.01	10 ± 4	44%	Unknown	37	6.38 ± 0.00	205 ± 42	21%	Unknown
18	14.75 ± 0.03	20 ± 6	30%	Unknown	38	7.44 ± 0.00	650 ± 72	11%	Oleanolic acid
19	16.26 ± 0.02	2019 ± 210	10%	Quercetin	39	7.76 ± 0.00	2829 ± 155	5%	Ursolic acid
20	17.99 ± 0.02	94 ± 8	9%	Unknown	40	9.32 ± 0.01	371 ± 117	31%	Unknown

Peaks 1–28 were detected in sample I by a gradient elution within 50 min run time, and peaks 29–40 were detected in sample II by an isocratic elution with a run time of 15 min.

may demonstrate the instrumental shifts rather than real differences of samples, if the analytical method was not well validated [7,14]. PCA scores plot of the dataset I and dataset II were illustrated in Fig. 2a and b. $Q2$ (cumulative) of the PCA model of dataset I and dataset II were 0.905 and 0.326, respectively. The PCA model of dataset II was more robust if the dataset was only centered but not Pareto scaled ($Q2 = 0.791$), while the clustering of samples did not differ much (PCA scores plot was not shown). Both PCA scores plot demonstrated a clear separation of the six batches of TFS. However, corresponding to the previous visual inspection of the chromatographic shifts, it was clearly observed in Fig. 2a that QC09, QC10 and QC11 progressively left the clustering of the previous QC samples. All the QC samples in dataset II clustered well. Accordingly, only the first two samples of each batch of TFS, which were analyzed before QC09 in both analytical runs, were included in the final combined dataset.

3.3. Quality assessment of TFS

PCA scores plot of the final combined dataset was illustrated in Fig. 3a, in which clear separation of the six batches of TFS was observed. $R2$ (cumulative) and $Q2$ (cumulative) of the PCA model were 0.992 and 0.950, respectively. PLS-DA was then carried out to cluster six batches of TFS ($Q2 = 0.980$). The six batches of TFS were clearly clustered as four classes in the PLS-DA scores plot (Fig. 3b). The corresponding PLS-DA loadings plot was illustrated in Fig. 3c.

Instead of commonly used visual inspections of the loadings plot coupled with subsequent validation with representative intensity plots [3,4,15], HCA using Ward's minimum variance method was applied to the loadings matrix to identify the differential variables. With HCA, variables accountable for the separation among groups may be automatically clustered according to their contributory power to the principal components. The HCA results were illustrated in the dendrogram (Fig. 3d), in which when the vertical lines are tall the clusters are far apart (different), and when they are short the clusters are close together (similar). The variables were clustered at the loading level of 0.2 as five groups, whose clustering patterns could be simultaneously seen with the same colors on the PLS-DA loadings plot (Fig. 3c). It was clearly shown that variables from group 1 (quercetin, kaempferol, isorhamnetin, unknown ingredient of peaks 31 and 36), group 2 (unknown ingredient of peak 10) and group 4 (oleanolic acid and ursolic acid) significantly

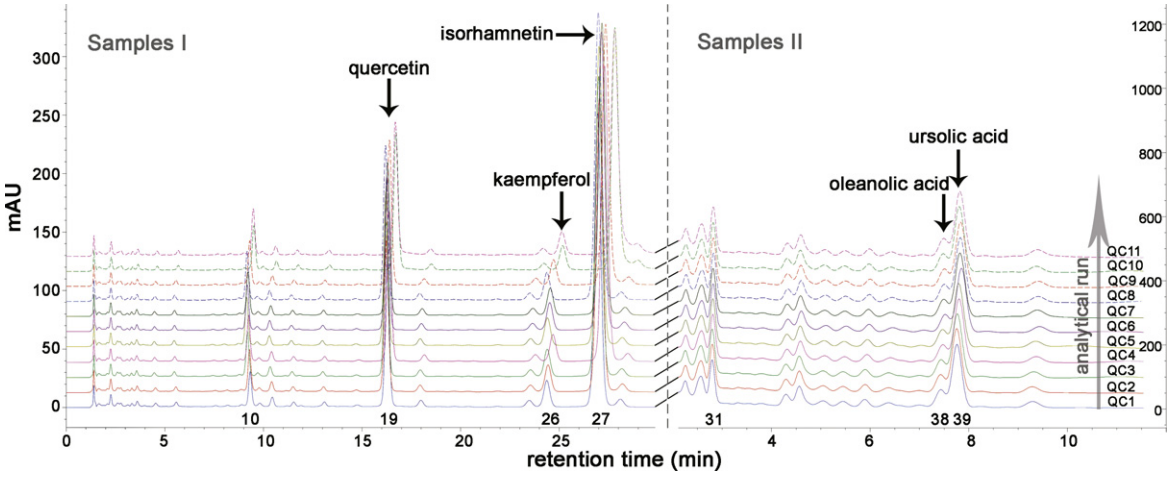

Fig. 1. Combined chromatograms of the QC sample I and QC sample II with the injection sequence in the analytical run. The peaks within the beginning 2 min of sample II were cut.

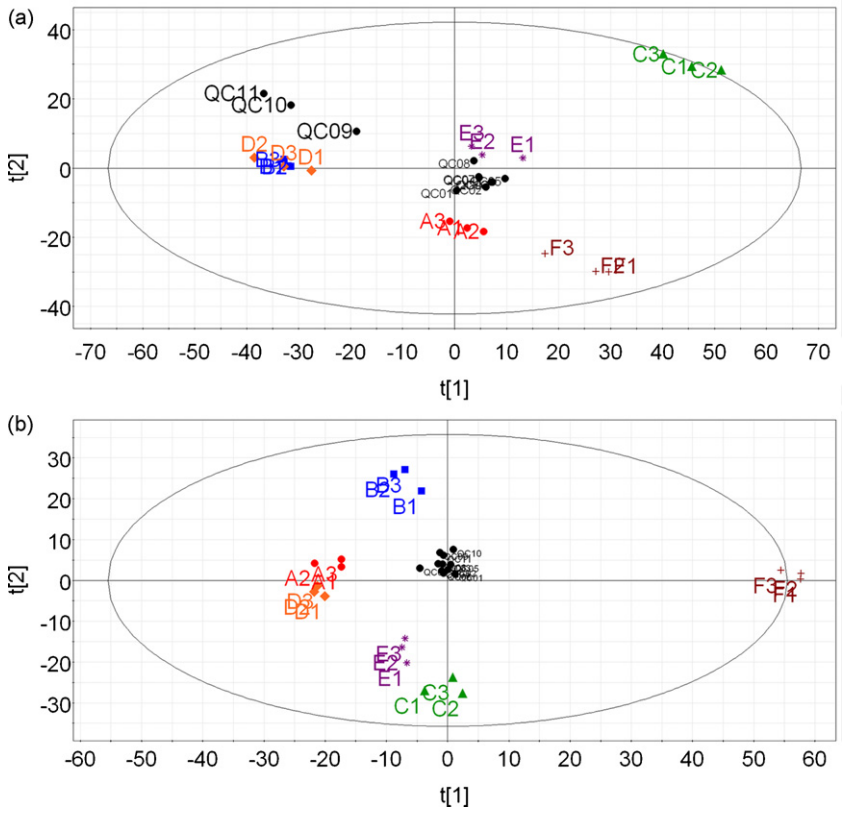

Fig. 2. PCA scores plot of the samples in dataset I (a) and PCA scores plot of the samples in dataset II (b).

contributed to the classification of six batches of TFS. The chemical structures of the identified unknown ingredients require further characterizations.

The result became very interesting while comparing the content of total flavones of samples to the results of PLS-DA classifications, as illustrated in Fig. 4a. It seemed the classification correlated well to the content of total flavones. This deduction was further validated by a PLS regression model. R2X, R2Y and *Q*2 of the PLS model were 0.823, 0.991 and 0.974, respectively. The comparisons between PLS predicted data (YPred) and the determined data (YVar) were illustrated in Fig. 4b. The present study based on two complementary HPLC-UV analyses could well predict the content of total flavones in TFS. On the other hand, this finding also prompted the determination of the content of total flavones potentially be a valid way to preliminarily control the quality of TFS.

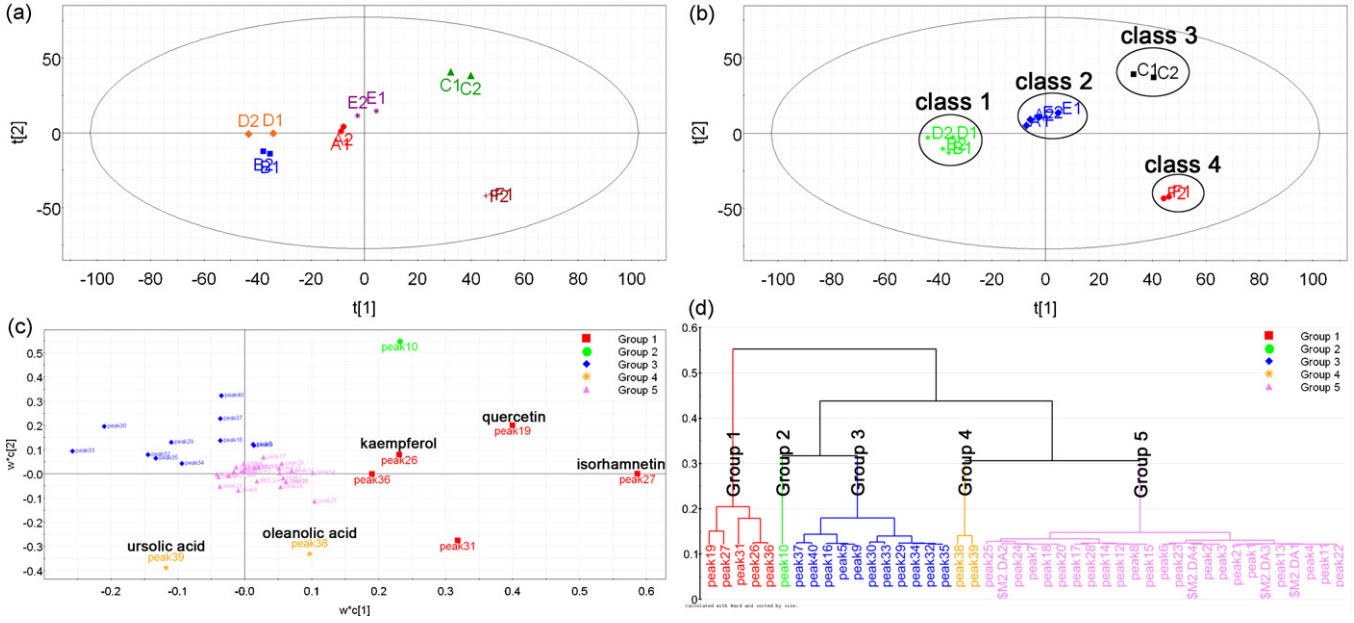

Fig. 3. PCA scores plot (a), PLS-DA scores plot (b), PLS-DA loadings plot (c) of the combined dataset and dendrogram of hierarchical cluster analysis (HCA) using Ward's minimum variance method for the loadings matrix (d). Variables in the loadings plot (c) were colored according to the dendrogram of HCA (d).

(a)

Batch	Content of total flavones	Classification by PLS-DA
B	10.34%	Class 1
D	10.58%	Class 1
A	11.38%	Class 2
E	11.50%	Class 2
C	13.68%	Class 3
F	12.49%	Class 4

(b)

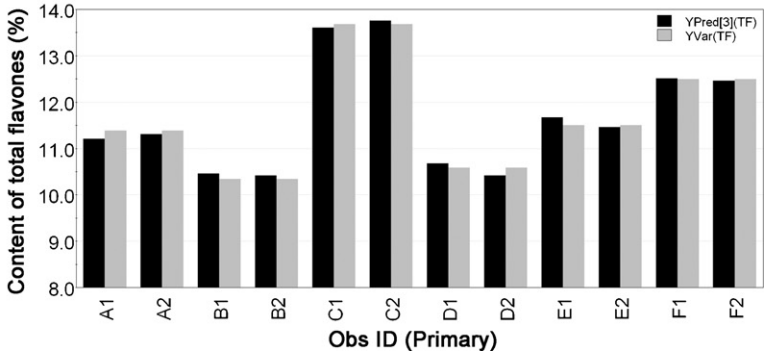

Fig. 4. Comparisons between the content of total flavones and the classifications by PLS-DA of six batches of TFS (a), comparisons between PLS predicted data (YPred) and the determined data (YVar) of the content of total flavones (b).

4. Conclusion

This study presented a good example on quality assessment of herbal extracts using HPLC-UV coupled with multivariate statistical analysis. Even though such a HPLC-UV based metabolomics platform has clear disadvantages that HPLC-UV may not provide the information of chemical structures of the ingredients determined, lack of sensitivity and fail to be a universal detector, as well as evolve a much longer run time compared to UPLC technique, the results from this study indicate that the proposed method could serve as a preliminary and cheap way to re-evaluate the quality of some widely used herbal extracts.

Acknowledgment

This work was supported by a grant from the Ph.D. Programs Foundation of Ministry of Education of China (No. 200806101089).

References

[1] Z.D. Zeng, Y.Z. Liang, F.T. Chau, S. Chen, M.K. Daniel, C.O. Chan, Anal. Chim. Acta 604 (2007) 89.
[2] Y.Z. Liang, P. Xie, K. Chan, J. Chromatogr. B: Analyt. Technol. Biomed. Life Sci. 812 (2004) 53.
[3] G. Xie, R. Plumb, M. Su, Z. Xu, A. Zhao, M. Qiu, X. Long, Z. Liu, W. Jia, J. Sep. Sci. 31 (2008) 1015.
[4] J. Kang, S. Lee, S. Kang, H.N. Kwon, J.H. Park, S.W. Kwon, S. Park, Arch. Pharm. Res. 31 (2008) 330.
[5] F. van der Kooy, F. Maltese, Y.H. Choi, H.K. Kim, R. Verpoorte, Planta Med. 75 (2009) 763.
[6] H. Wen, S. Kang, Y. Song, S.H. Sung, S. Park, Phytochem. Anal. (2009).
[7] T. Sangster, H. Major, R. Plumb, A.J. Wilson, I.D. Wilson, Analyst 131 (2006) 1075.
[8] H.G. Gika, E. Macpherson, G.A. Theodoridis, I.D. Wilson, J. Chromatogr. B: Analyt. Technol. Biomed. Life Sci. 871 (2008) 299.
[9] F. Michopoulos, L. Lai, H. Gika, G. Theodoridis, I. Wilson, J. Proteome Res. 8 (2009) 2114.
[10] E. Zelena, W.B. Dunn, D. Broadhurst, S. Francis-McIntyre, K.M. Carroll, P. Begley, S. O'Hagan, J.D. Knowles, A. Halsall, I.D. Wilson, D.B. Kell, Anal. Chem. 81 (2009) 1357.
[11] R.A. van den Berg, H.C. Hoefsloot, J.A. Westerhuis, A.K. Smilde, M.J. van der Werf, BMC Genomics 7 (2006) 142.
[12] C. Chen, H. Zhang, W. Xiao, Z.P. Yong, N. Bai, J. Chromatogr. A 1154 (2007) 250.
[13] M. Martelanc, I. Vovk, B. Simonovska, J. Chromatogr. A 1216 (2009) 6662.
[14] H.G. Gika, G.A. Theodoridis, J.E. Wingate, I.D. Wilson, J. Proteome Res. 6 (2007) 3291.
[15] E.C. Chan, S.L. Yap, A.J. Lau, P.C. Leow, D.F. Toh, H.L. Koh, Rapid Commun. Mass Spectrom. 21 (2007) 519.