Amino Acids (2015) 47:1589–1603
DOI 10.1007/s00726-015-2002-5

ORIGINAL ARTICLE

Rapid and sensitive analysis of 27 underivatized free amino acids, dipeptides, and tripeptides in fruits of *Siraitia grosvenorii* Swingle using HILIC-UHPLC-QTRAP®/MS^2 combined with chemometrics methods

Guisheng Zhou[1] · Mengyue Wang[1] · Yang Li[1] · Ying Peng[1] · Xiaobo Li[1]

Received: 6 February 2015 / Accepted: 1 May 2015 / Published online: 15 May 2015

Abstract In the present study, a new strategy based on chemical analysis and chemometrics methods was proposed for the comprehensive analysis and profiling of underivatized free amino acids (FAAs) and small peptides among various *Luo-Han-Guo* (LHG) samples. Firstly, the ultrasound-assisted extraction (UAE) parameters were optimized using Plackett–Burman (PB) screening and Box–Behnken designs (BBD), and the following optimal UAE conditions were obtained: ultrasound power of 280 W, extraction time of 43 min, and the solid–liquid ratio of 302 mL/g. Secondly, a rapid and sensitive analytical method was developed for simultaneous quantification of 24 FAAs and 3 active small peptides in LHG at trace levels using hydrophilic interaction ultra-performance liquid chromatography coupled with triple–quadrupole linear ion-trap tandem mass spectrometry (HILIC-UHPLC-QTRAP®/MS^2). The analytical method was validated by matrix effects, linearity, LODs, LOQs, precision, repeatability, stability, and recovery. Thirdly, the proposed optimal UAE conditions and analytical methods were applied to measurement of LHG samples. It was shown that LHG was rich in essential amino acids, which were beneficial nutrient substances for human health. Finally, based on the contents of the 27 analytes, the chemometrics methods of unsupervised principal component analysis (PCA) and supervised counter propagation artificial neural network (CP-ANN) were applied to differentiate and classify the 40 batches of LHG samples from different cultivated forms, regions, and varieties. As a result, these samples were mainly clustered into three clusters, which illustrated the cultivating disparity among the samples. In summary, the presented strategy had potential for the investigation of edible plants and agricultural products containing FAAs and small peptides.

Keywords Siraitiae Fructus · Free amino acids · Peptides · HILIC-UHPLC-QTRAP®/MS^2 · Chemometrics · Artificial neural network

Introduction

Siraitiae Fructus are the fruits of *Siraitia grosvenorii* Swingle, a perennial vine of the Cucurbitaceae family. The Chinese name of Siraitiae Fructus was *Luo-Han-Guo* (LHG), and these plants were primarily cultivated in Guangxi Province of China (Zhang et al. 2012). Recently, the potential medicinal and nutritional effects of LHG have attracted considerable attention, as it is considered to be rich in free amino acids (FAAs), polypeptides, proteins, polysaccharide, vitamins, and microelements. In addition, mogrosides, the active components of high-intensity sweetness from LHG, were reported to have anti-tumor, anti-obesity, and insulin-secretion stimulating activities (Chen et al. 2011; Li et al. 2014; Takasaki et al. 2003). There were also many studies focusing on the pharmaceutical active components of mogrosides in LHG (Li et al. 2014). The active pharmaceutical and nutritional components were key factors of functional foods evaluation (Bernal et al. 2011);

Handling Editor: D. Tsikas.

Electronic supplementary material The online version of this article (doi:10.1007/s00726-015-2002-5) contains supplementary material, which is available to authorized users.

✉ Xiaobo Li
xbli@sjtu.edu.cn

[1] School of Pharmacy, Shanghai Jiao Tong University, Shanghai 200240, China

therefore, it would be necessary to investigate the ingredients when developing and utilizing LHG resources.

Functional foods like LHG were often recommended because of their high level of FAAs, polypeptides, and proteins (Wu 2013; Lahrichi et al. 2013). Proteins from food were traditionally regarded as the source of energy and essential amino acids, which were needed for body growth and maintenance of physiological functions (Sarmadi and Ismail 2010). Additionally, many bioactive polypeptides were at the size of 2–20 amino acids and possessed interesting functions such as immunoregulation and antioxidation (Meisel and FitzGerald 2003; Sarmadi and Ismail 2010). The proteins and polypeptides from LHG will be investigated in our following research. In this study, the contents of FAAs and small peptides (less than 3 amino acids) in LHG were firstly evaluated. As previously reported, FAAs have many biological activities such as regulating gene expression, preventing tumorigenesis, suppressing obesity, and reducing blood pressure (Zhou et al. 2013; Wu 2013). Moreover, nutritionally relevant small food peptides such as glutathione (GSH), cystine (Cyst), and alanyl–glutamine (Ala–Gln) have arisen common interest in their potential positive effects, such as cellular homeostasis, antioxidant, and detoxifier (Bakirdere et al. 2010; Kuśmierek and Bald 2008). Hence, considering their benefits to human health, many FAAs and small peptides have been selected as quality control markers for functional foods such as royal jelly, mushroom, and Chinese teas (Akamatsu and Mitsuhashi 2013; Kıvrak et al. 2014; Yan et al. 2014). Therefore, it would be a straightforward way to assess the quality for LHG through the detection of contained FAAs and small peptides, which can potentially promote its value as a functional food.

Several analytical methods were reported for the analysis of FAAs or small peptides in different matrixes. Generally, these methods could be divided into chemical derivatization (Megías et al. 2015) and underivatized strategies. The chemical derivatization reactions were developed to deal with FAAs or small peptides with high polarity, low volatility, or weak chromophore groups, which were difficult to separate and determine. Approaches that could be applied include increasing the volatility of the analytes for analysis by gas chromatography (Mudiam and Ratnasekhar 2013), introducing chromophores or fluorophores groups in their structure for analysis by liquid chromatography (LC)/capillary electrophoresis (CE) with ultraviolet (UV) (Rezazadeh et al. 2015; Tian et al. 2014), or fluorescence (Tuberoso et al. 2015). The most frequently used derivatization reagents were ninhydrin, o-phthalaldehyde (OPA), 5-dimethylaminonaphthalene-1-sulfonyl chloride (DANSYLCl) and so on (Friedman 2004; Ilisz et al. 2008; Tuberoso et al. 2015). However, several drawbacks for the above reagents need to be considered: limitation to post-column derivatization (ninhydrin); long derivatization (DANSYL-Cl); the inability to derivatize the secondary amino group (OPA) (Fiechter and Mayer 2011; Kıvrak et al. 2014). To overcome the disadvantage of derivatization, underivatized method would be a better choice for determining the target compounds with regard to its shorter analysis time, higher separation efficiency, and minimal solvent consumption when using suitable chromatographic and detected techniques. LC (Gökmen et al. 2012; Troise et al. 2015) and CE (Schultz and Moini 2003) were the two major chromatographic techniques described in literatures to quantify underivatized amino acids and small peptides, and these chromatographic techniques were frequently coupled with mass spectrometric (MS) and/or evaporative light scattering detection (Guo et al. 2013). Among these methods, it was noteworthy that high-resolution MS (HRMS) methods could avoid the risk of inaccurate measurements caused by unresolved background interferences in complex matrices (Gökmen et al. 2012; Troise et al. 2015). Unfortunately, the isomers peaks of leucine and isoleucine could not be well separated by HRMS methods (Guo et al. 2013). In summary, most of the existing methods harbored one or more of the following disadvantages, multiple and time-consuming sample preparations, long analyzing time, poor separation of isomers, low sensitivity and/or bad resolution, and limited quantities and types of analytes that can be detected (only FAAs and small peptides). For this reason, in this study, a rapid, specific, and sensitive method of hydrophilic interaction ultra-performance liquid chromatography coupled with triple–quadrupole linear ion-trap tandem mass spectrometry (HILIC-UHPLC-QTRAP®/MS^2) was developed for simultaneous quantitation of 27 underivatized FAAs and small peptides. Specific and sensitive QTRAP®/MS^2 was selected because of its software switchable ability to trigger a sensitive enhanced product ion (EPI), and to scan in the sensitive linear ion-trapping (LIT) mode whenever an MRM survey scan signal exceeded a predefined independent data acquisition (IDA) criteria (Guna and Londry 2011).

During recent years, there had been increasing interest in illuminating the intrinsic relationship of multivariate data which were composed of numerous variables measured from many samples by validated analytical method. Chemometrics methods had been widely used for depicting the intrinsic similarities and differences of functional foods by comparing particular chemical parameters in different samples (Duan et al. 2014). The central idea of unsupervised principal component analysis (PCA) was to reduce the dimensionality of a data set consisting of a large number of interrelated variables. PCA was a very useful classification technique and widely used in the field of analytical chemistry (Kumar et al. 2014). Counter propagation artificial neural network (CP-ANN) was one of the most

popular ANNs and a common supervised learning method. CP-ANN was increasing its uses related to different chemical issues and nowadays could be considered as a popular tool in multivariate analysis. One of the reasons of its success could be related to the ability of calculating different multivariate models, such as clustering, classification, and regression models (Ballabio et al. 2011).

In this study, a new strategy based on chemical analysis and chemometrics methods was proposed for the comprehensive analysis and profiling of FAAs and small peptides in different LHG samples. Firstly, a simple, rapid, accurate, and repeatable HILIC-UHPLC-QTRAP®/MS^2 method was validated for determination of 24 FAAs and 3 active small peptides. Furthermore, the above validated method was employed for the analysis of 27 target compounds in different LHG samples. Finally, PCA, the common unsupervised method, was used to classify LHG samples based on the 27 chemical markers. The supervised learning method of CP-ANN was utilized to further separate clusters and validate the classify model. The present strategy might be employed to investigate and profile the relevant LHG products containing FAAs and small peptides.

Experimental

Materials and samples

The acetonitrile and formic acid were of HPLC grade from Merck (Darmstadt, Germany). Deionized water was prepared from distilled water through a Milli–Q water purification system (Millipore, Bedford, MA, USA). Other reagents and chemicals were of analytical grade. 27 reference standards of tryptophan, phenylalanine, leucine, isoleucine, methionine, γ-aminobutyric acid, valine, proline, tyrosine, cysteine, alanine, hydroxyproline, threonine, glycine, glutamic acid, glutamine, serine, glutathione (reduction), asparagine, alanyl–glutamine, citrulline, aspartic acid, arginine, histidine, lysine, ornithine, and cystine were purchased from Sigma-Aldrich (St. Louis, MO, USA). Some information of these compounds is shown in Table S1 and Fig. S1.

Standard stock solutions (20,000 ng/mL) were prepared by dissolving 2.0 mg individual standards in 100 mL 50 % (v/v) acetonitrile. These solutions were diluted with 50 % (v/v) acetonitrile to a series of appropriate concentrations and used to construct calibration curves. All the standard solutions were stored at 4 °C (Table S2) until further use and were filtered through a 0.22 μm nylon membrane before analysis.

Forty batches of LHG samples (**S1**–**40**) were collected and detailed information on these samples is summarized in Table 1. The samples were authenticated as the fruits of *Siraitia grosvenorii* Swingle by Prof. Xiaobo Li from Shanghai Jiao Tong University. Voucher specimens were deposited at the School of Pharmacy, Shanghai Jiao Tong University (Shanghai, China). After collection, the samples were dried by air, which was better consistent with the current production facts of LHG. The dried samples were ground using a mortar and pestle, and then sieved through a 40 mesh stainless-steel sieve before extraction.

Extraction procedures

Based on previous reports, ultrasound-assisted extraction (UAE) was an effective method for extracting FAAs from foods (Carrera et al. 2015). In this study, UAE was employed to extract FAAs and small peptides from LHG, and the extraction conditions were optimized using response surface methodology (RSM). The extraction process was performed with an ultrasonic device (SY-5200T, Shanghai Shenyuan Ultrasonic Instrument Co. Shanghai, China). Test powder of 1.0 g (**S11**) was weighed accurately and extracted in several designed UAE conditions. After UAE, the sample was dissolved by equal volume of acetonitrile (namely the final system was 50 % acetonitrile). The dissolved sample was shaken for 3 min and then centrifuged at 13000 rpm for 10 min to collect the supernatant for analysis.

Box–Behnken design (BBD) is a spherical, rotatable, or nearly rotatable second order design based on three-level incomplete factorial designs (Deshmukh et al. 2014). In preliminary experiments, ultrasonic power (W, χ_1), extraction time (min, χ_2), and the solid–liquid ratio (mL/g, χ_3) were the three effective independent variables according to Plackett–Burman (PB) screening experiments. By employing the PB and BBD, ultrasonic power of 100–400 W, extraction time of 30–50 min, and solid–liquid ratio of 200–400 mL/g were optimized to obtain the maximum yield of the total contents of 27 target compounds from LHG.

A three-factor, three-level BBD was applied to determine the optimal conditions for UAE in this study. The BBD in the experimental design consisted of 17 treatments including five replicates of the central point (Table S3). Data from BBD were analyzed by multiple regressions to fit the following quadratic equation:

$$Y = \varphi_0 + \sum_{i=1}^{n} \varphi_i x_i + \sum_{i=1}^{n} \varphi_{ii} x_i^2 + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \varphi_{ij} x_i x_j, \quad (1)$$

where Y, φ_0, φ_i, φ_{ii}, and φ_{ij} indicated the predicted response, intercept term, linear coefficient, squared coefficient, and interaction coefficient, respectively. Model adequacy was evaluated using F ratio and coefficient of determination (R^2) represented at 5 % level of significance, accordingly (Zhou et al. 2014). BBD results were analyzed using the

Table 1 Summary of the tested samples of the fruit of *Siraitia grosvenorii* Swingle

Sample number	Cultivation region	Cultivated form	Cultivated varieties
S.1	Yongfu, Guangxi	Cuttage	Qingpiguo
S.2	Yongfu, Guangxi	Tissue culture	Qingpiguo
S.3	Yongfu, Guangxi	Tissue culture	Qingpiguo
S.4	Yongfu, Guangxi	Tissue culture	Qingpiguo
S.5	Lingui, Guangxi	Tissue culture	Qingpiguo
S.6	Lingui, Guangxi	Tissue culture	Qingpiguo
S.7	Lingui, Guangxi	Tissue culture	Qingpiguo
S.8	Longsheng, Guangxi	Cultivate	Qingpiguo
S.9	Longsheng, Guangxi	Cultivate	Qingpiguo
S.10	Longsheng, Guangxi	Cultivate	Qingpiguo
S.11	Longsheng, Guangxi	Cultivate	Qingpiguo
S.12	Longsheng, Guangxi	Cultivate	Qingpiguo
S.13	Yongfu, Guangxi	Cultivate	Qingpiguo
S.14	Yongfu, Guangxi	Cultivate	Qingpiguo
S.15	Yongfu, Guangxi	Cultivate	Qingpiguo
S.16	Yongfu, Guangxi	Cultivate	Qingpiguo
S.17	Yongfu, Guangxi	Cultivate	Qingpiguo
S.18	Lingui, Guangxi	Cultivate	Qingpiguo
S.19	Lingui, Guangxi	Cultivate	Qingpiguo
S.20	Lingui, Guangxi	Cultivate	Qingpiguo
S.21	Lingui, Guangxi	Cultivate	Qingpiguo
S.22	Guilin, Guangxi	Cultivate	Qingpiguo
S.23	Guilin, Guangxi	Cultivate	Qingpiguo
S.24	Guilin, Guangxi	Cultivate	Qingpiguo
S.25	Guilin, Guangxi	Cultivate	Qingpiguo
S.26	Guilin, Guangxi	Cultivate	Qingpiguo
S.27	Longsheng, Guangxi	Tissue culture	Hongmaoguo
S.28	Yongfu, Guangxi	Tissue culture	Hongmaoguo
S.29	Lingui, Guangxi	Tissue culture	Hongmaoguo
S.30	Lingui, Guangxi	Tissue culture	Hongmaoguo
S.31	Lingui, Guangxi	Tissue culture	Hongmaoguo
S.32	Lingui, Guangxi	Tissue culture	Hongmaoguo
S.33	Yongfu, Guangxi	Cultivate	Hongmaoguo
S.34	Yongfu, Guangxi	Cultivate	Hongmaoguo
S.35	Lingui, Guangxi	Cultivate	Hongmaoguo
S.36	Lingui, Guangxi	Cultivate	Hongmaoguo
S.37	Longsheng, Guangxi	Tissue culture	Dongguaguo
S.38	Lingui, Guangxi	Tissue culture	Dongguaguo
S.39	Lingui, Guangxi	Tissue culture	Dongguaguo
S.40	Lingui, Guangxi	Cultivate	Dongguaguo

Design–Expert 8.5 software (Trial version, State–Ease Inc., Minneapolis, MN).

HILIC-UHPLC-QTRAP®/MS² analysis conditions

An ACQUITY UPLC system (Waters, Milford, MA, USA) coupled to a hybrid triple–quadrupole linear ion trap mass spectrometer (5500 QTRAP®) with a turbo ion spray source from AB SCIEX (Foster City, CA, USA) were employed in this study.

UHPLC was equipped with a binary solvent delivery system and an auto-sampler. An Acquity UPLC BEH amide column (2.1 mm × 100 mm, 1.7 μm) equipped with an ACQUITY UPLC BEH Amide 1.7 μm VanGuard

Pre-column. The mobile phase was A (water, 10 mmol/L ammonium acetate and 0.5 % acetic acid) and B (acetonitrile, 1 mmol/L ammonium acetate, and 0.1 % acetic acid) at a flow rate of 0.4 mL/min. The linear gradient conditions were 0–3 min, 15–17.5 % A; 3–5.5 min, 17.5–30 % A; and 5.5–7 min, 30–46 %; 7–7.5 min, 46 % A ("Effect of gradient elution" in Supporting Information). The column temperature was maintained at 35 °C and the injection volume was 3 μL. Two cycles of weak (90 % acetonitrile) and strong (20 % acetonitrile) solvent washing of the injecting system were carried out between injections.

The operating parameters of QTRAP®/MS^2 performed in positive ionization mode were set as follows: ion spray voltage 5000 V, curtain gas 20 (arbitrary units), nebulizer gas 40 psi (GS1), auxiliary gas 60 psi (GS2), and probe temperature 550 °C. Nitrogen served as nebulizer and collision gas. Detection was performed in multiple reaction monitoring (MRM) mode of selected ions at the first ($Q1$) and third quadrupole ($Q3$). To choose the fragmentation patterns of m/z ($Q1$) → m/z ($Q3$) for the analytes in MRM mode, direct infusion into the MS of single standard solution was performed and the product ion scan mass spectra were recorded. Once the fragment ions were chosen, the MRM conditions were further optimized for the analytes to obtain maximum sensitivity. All ESI and MS parameters were optimized individually for each target compound and listed in Table S1. All data were acquired and processed using the Analyst® 1.5.1 software package.

Method performance

The slope of the standard addition plot was compared with the slope of standard calibration plot to calculate the signal suppression/enhancement, which was used for quantitative assessment of the matrix effect. According to our previous report, the detail process of matrix effect was described in Supporting Information (Zhou et al. 2013). The working standard mixture of 20 μg/mL containing 27 target compounds was prepared and diluted to appropriate concentrations for the establishment of calibration curves. At least 6 concentrations of the 27 analytes solution were analyzed in triplicate, and then the calibration curves were constructed by plotting the peak areas versus the concentrations of each analyte. The limit of detection (LOD) and limit of quantification (LOQ) were determined using diluted standard solution with the signal-to-noise ratios (S/N) of 3 and 10, respectively. The precision of the method was evaluated by analyzing the 27 standard compounds. For intra-day variability test, the mixed standards solution was analyzed for six replicates ($n = 6$) within 1 day, while for inter-day variability test, the solution was examined in duplicates for consecutive 3 days ($n = 6$). To evaluate repeatability of the proposed method, six samples from the same origin were extracted and analyzed in parallel. A sample solution was stored at room temperature and analyzed at 0, 2, 4, 6, 12, and 24 h, respectively, to evaluate sample stability. The RSDs were used as the indicator of stability. A recovery test was used to evaluate the accuracy of the above method. It was performed by adding the corresponding standard compounds with high (1.0 mg/g) and low (0.1 mg/g) levels to 0.5 g LHG powder of sample **S11** which had previously been analyzed and known amount of the 27 analytes. The spiked samples were then extracted and analyzed using the aforementioned method, and triplicate experiments were performed at each level.

Chemometric analysis

The LHG data set consisted of a 40 × 27 matrix, in which rows represented the 40 LHG samples from different sources, and columns for the 27 chemical target compounds (variables) including 24 FAAs and 3 active small peptides. In this study, multivariate classification was completed using the commonly used chemometric techniques of unsupervised PCA combined with supervised CP-ANN analysis.

The procedure of PCA established a linear model which allowed converting original and correlated variables into uncorrelated variables called principal components (PCs). In this study, PCs were calculated as the following:

$$Y = PV^{\mathrm{P}}, \tag{2}$$

where Y is the original analytical data matrix with m LHG samples and n variables. The P and V^{P} were the score matrix with a new set of ($m \times j$) vectors (PCs) and a loadings matrix with a set of ($j \times n$) vectors (Vs), respectively, (Ni et al. 2014) as shown in the formula below:

$$P = \left[P_1, P_2, P_3, \ldots P_j\right] \tag{3}$$

$$V = \left[V_1, V_2, V_3, \ldots V_j\right]. \tag{4}$$

The distribution of the samples on the PCs, known as the scores plots, was a powerful strategy for classifying samples according to the data measured. These graphs revealed patterns and other features that may be correlated to sample characteristics (Dago et al. 2014). The analysis of the distribution of variables (the loadings plot) provided information about the correlations among 24 FAAs and 3 active small peptides concentrations. In this study, PCA was performed on the SPSS 16.0 software (SPSS, Chicago, IL, USA), to demonstrate the variability of the 24 FAAs and 3 active small peptides among 40 batches of LHG samples.

CP-ANN was a modeling method which combined features from both supervised and unsupervised learning. CP-ANN consisted of two layers, a Kohonen layer, and an output layer, in which the neurons possessed as many weights as the number of classes to be modeled. Kohonen layer in

CP-ANN performed the mapping of the multidimensional input data into lower dimensional array (most often two-dimensional, since the exploration and visualization of more than two-dimensional self-organizing maps were not easy for human perception) (Ballabio and Vasighi 2012). The output layer had the same topological set-up as the input layer, and was needed to perform supervised classification. At the end of the network training, each neuron of the Kohonen layer could be assigned to a class on the basis of the output weights and all the samples placed in that neuron were automatically assigned to the corresponding classes. A given model was trained to obtain a cluster and an unknown cluster was similarly compared to the training set (Ni et al. 2014). In this study, CP-ANN was used to confirm the results of the unsupervised PCA for classification of LHG samples using 27 target markers, and the toolbox of CP-ANN was used for supervised classification purposes under MATLAB 6.5 (Mathworks Natick, MA, USA).

Results and discussion

Optimization of UAE experimental conditions for the maximum extraction of target compounds

An internal standard (IS) could be used to correct variability in dilutions, evaporation, degradation, recovery, adsorption, derivatization, and MS parameters such as injection volume and MS response (Pirman et al. 2013; Stokvis et al. 2005). Either a stable isotope labeled (SIL) form of the analyte or a structural analog was used as IS in different studies. Nowadays, SIL, one of the most powerful techniques used in analytical chemistry field, has been used for the quantitative analysis of chemicals in biological, environmental, and food samples. SIL provided many reproducible results for sample extraction process to investigate multiple extraction parameters (Zhang et al. 2014). However, expense was frequently raised as an objection to the use of SIL peptides (Gillette and Carr 2013). Instead, a simple and practical RSM could be used for evaluating multiple extraction parameters as well as their interactions, and its feature could effectively overcome the drawbacks of conventional 'one-factor-at-a-time' or 'full-factors' approaches (Zhou et al. 2015).

Based on several previous reports, UAE and water were suitable extraction tool and solvent for extracting FAAs from functional foods, respectively (Carrera et al. 2015; Zhou et al. 2013). However, the effects of other UAE parameters were usually neglected when the simultaneously maximum yield of both FAAs and peptides was aimed for. To optimize UAE parameters, a screening design of PB was built to identify the main factors affecting the responses (total content of 24 FAAs and 3 active small peptides from LHG) among 6 variables (such as ultrasonic power, frequency, extraction time, times, temperature, and the solid–liquid ratio). The screening experiments indicated that ultrasonic power (χ_1), extraction time (χ_2), and solid–liquid ratio (χ_3) were the most relevant parameters to the yield of target compounds.

After the screening design, a BBD for the three most significant factors was employed to evaluate the optimal UAE conditions. Analysis of variance (ANOVA) was performed to identify the statistically significant parameters ($p < 0.05$), the result suggested that χ_1, χ_2, χ_1^2, χ_2^2, and χ_3^2 were significant. Consequently, a simplified model can thus be expressed as:

$$Y(\mathrm{Tc}) = 5616.9 + 326.9\chi_1 + 233.4\chi_2 - 946.6\chi_1^2 - 469.6\chi_2^2 - 442.3\chi_3^2, \quad (5)$$

where Y(Tc) represented the yield of total content of FAAs and peptides. The model's F value of 13.13 implied the model to be significant ($p < 0.05$). The "fitness" of the model ($p > 0.05$) suggested that it was accurate enough to predict the response variable. The regression coefficient $R^2 = 0.9441$ also indicated that the resulting model was a good fit for FAAs and peptides extraction.

3D response surface plots (Fig. 1a–c) and 2D contour plots (Fig. 1a–c) illustrated interacting effects of factors on the responses. According to the results calculated from the desirability function (DF), the final optimum working conditions were selected: ultrasound power of 280 W, extraction time of 43 min, and the solid–liquid ratio of 302 mL/g. Under the optimum conditions, the experimental yield of total content of target compounds was 5590.3 ± 76.2 μg/g ($n = 3$), which was close to the predicted values (5682.64 μg/g), indicating that the model was adequate for the extraction process. More detailed explanation could be found at "Optimization and verification procedures", Table S3, and Fig. S2 in Supporting Information.

Optimization of HILIC-UHPLC-QTRAP®/MS² conditions

To optimize the QTRAP®/MS² conditions, Q1 full scans were conducted under both positive and negative electrospray ionization (ESI) modes. The results revealed that not only 24 FAAs but also 3 active small peptides exhibited higher sensitivity and clearer mass spectra (preferable signal-to-noise ratios) in the positive ion mode, which made it easier to detect amino acids of lower content in LHG, and easier to confirm molecular ions or quasi-molecular ions in the identification of each peak. Thus, the ESI^+ mode was selected in the following studies ("Optimization of ESI modes" in Supporting Information). To choose the fragmentation patterns of m/z ($Q1$) → m/z ($Q3$) for the analytes

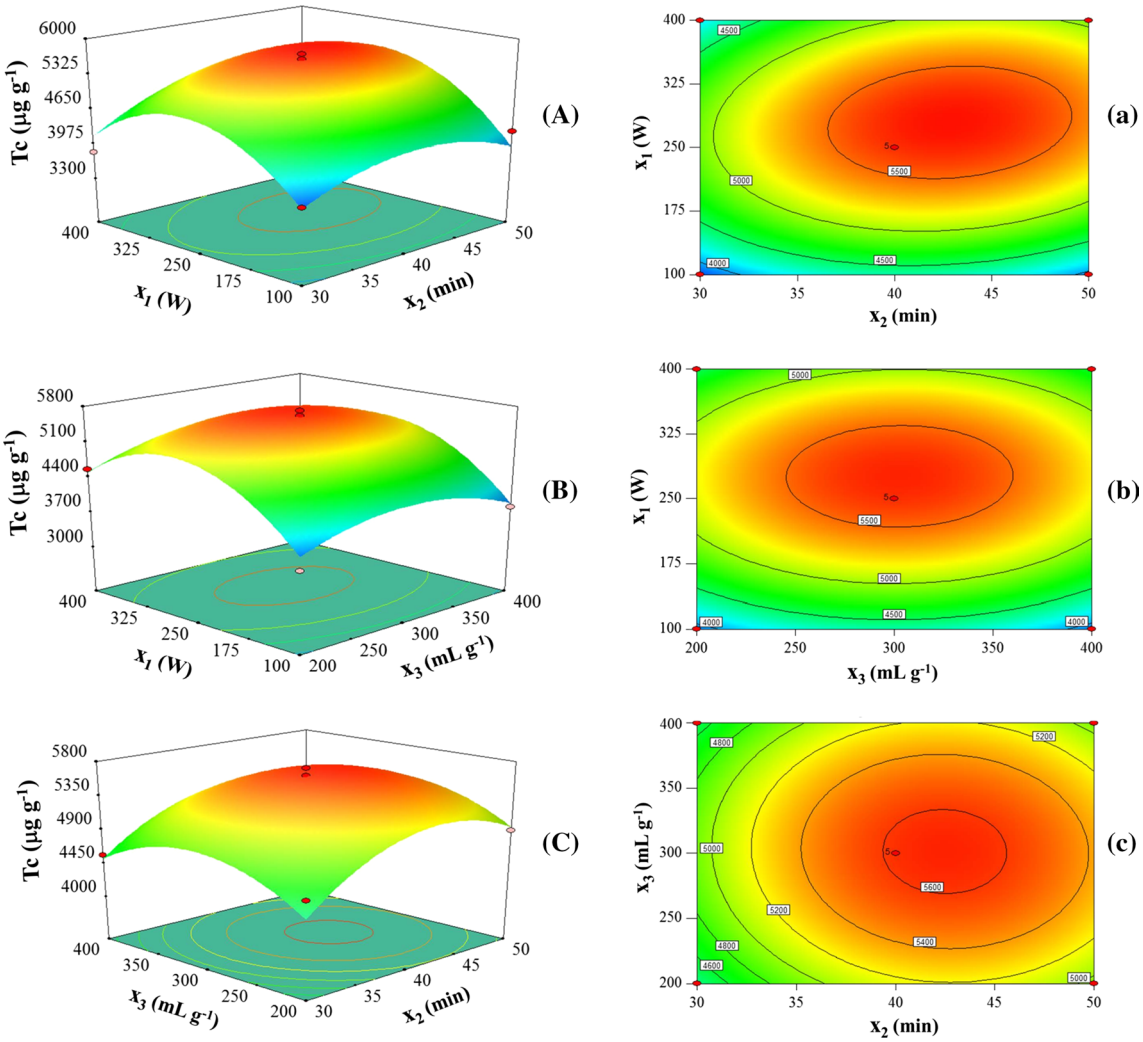

Fig. 1 The developed RSM model for the yield of total content of 27 target compounds from LHG, 3D response surface plots (**A**–**C**) and 2D contour plots (**a**–**c**)

in MRM mode, direct infusion into the MS of each standard solution (100 ng/mL solution at 10 μL/min) was performed and the product ion scan mass spectra were recorded. Once the fragment ions were confirmed, the MRM conditions would also be optimized for the analytes (Table S1, "Optimization of CE values" in Supporting Information).

For all FAAs analyzed ("Optimization of MRM" in Supporting Information), the protonated molecules $[M+ H]^+$ were detected with the abundant intensities, and selected as the precursor ions ($Q1$). There were five main forming pathways for the product ions ($Q3$) of target compounds. (1) Phe, Leu, Ile, Met, Val, Pro, Tyr, Ala, Thr, Gly, Ser, His, and Asp (involved the carboxyl and the α-amino group) showed an abundant product ion at $[M + H - 46]^+$, which corresponded to the neutral loss of formic acid by a rearrangement (Guo et al. 2013; Zhou et al. 2013); (2) Trp, GABA, Cit, and Orn had a common neutral loss of *m/z* 17 due to the loss of NH_3 (Guo et al. 2013); (3) Cys, Gln, and Lys had an abundant product ion at $[M + H - 46 - 17]^+$ (Guo et al. 2013; Zhou et al. 2013); (4) Hpro and Glu had the Q3 of $[M + H - 46 - 18]^+$ which were formed by the neutral loss of 18 Da (H_2O) from the fragments of $[M + H - 46]^+$ (Zhou et al. 2013); (5) The ions at *m/z* 74 and 70 were chosen as the optimal products of Asn and Arg, respectively (Zhou et al. 2013).

Three active small peptides glutathione (GSH), alanyl–glutamine (Ala–Gln), and cystine (Cyst), generated predominant protonated $[M + H]^+$ precursor ions at *m/z* 308.1, *m/z* 218.0, and *m/z* 240.9, respectively, in $Q1$ MS full scan spectra. The fragmentation of GSH produced major signals at *m/z* 179.0 with a neutral loss of 129 Da ($C_5H_7O_3N$, pyroglutamic acid) from the ion at *m/z* 308.1 which was confirmed by the reports (Castro-Perez et al. 2005). The major signal at *m/z* 83.8 of Ala–Gln was generated by a

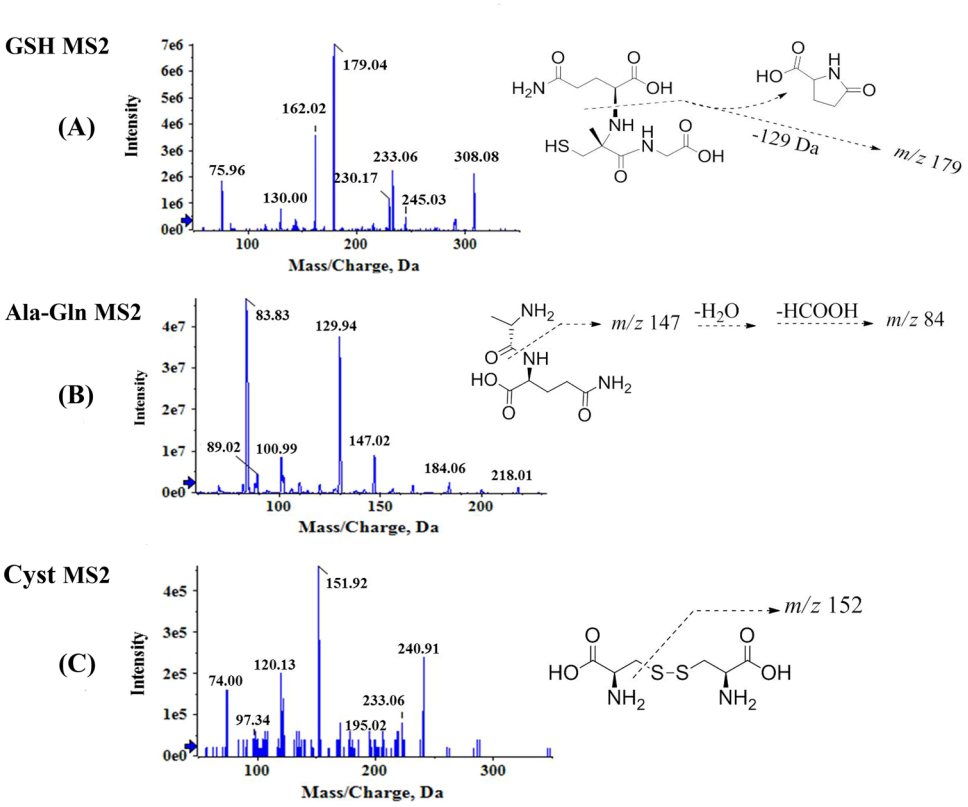

Fig. 2 Product ion mass spectra (ESI^+) of GSH (**a**), Ala–Gln (**b**) and Cyst (**c**)

series of neutral loss of 18 Da (H_2O) and 46 Da (HCOOH) from the identical ion at *m/z* 147.0 which corresponded to the neutral loss of an alanine moiety. Finally, the fragments of $[M + H - 89]^+$ was selected as the product ion for the Cyst. The MS^2 spectrometric fragmentation patterns of three active small peptides are shown in Fig. 2.

According to our previous study, the HILIC column was suitable for separation of these polar FAAs with poor retention property on RP-LC, especially for high polar ornithine, lysine, histidine, arginine, and aspartic acid. The HILIC column has formerly been optimized for the separation of FAAs (Zhou et al. 2013). However, there were no reports yet on whether this column was appropriate for the simultaneous separation of both highly polar FAAs and small peptides. Two types of HILIC columns, an Acquity UPLC BEH HILIC (2.1 × 100 mm, 1.7 μm), and an Acquity UPLC BEH Amide column (2.1 mm × 100 mm, 1.7 μm) were then investigated in this study. The results indicated that Acquity UPLC BEH Amide column provided better resolution and could be used for both 24 FAAs and 3 mall peptides.

In order to get sharper peaks with less broadening and tailing, two strategies were applied in this study. The first one was the use of organic solvent acetonitrile which could reduce ionization of high polar compounds. Both the standards and samples were solved in 50 % acetonitrile at the suitable concentration (<10 μg/mL for each compound) at which all the target compounds could dissolve completely. The results showed that 50 % acetonitrile analytical system of standards and samples obviously improved the peak shapes especially for His, Arg, Orn, and Lys which always presented tailing peaks when using water as solvent (Guo et al. 2013; Wang et al. 2013; Zhou et al. 2013). According to our previous FAAs research (Zhou et al. 2013), the other strategy to improve peak shapes was to the application of different mobile phase additives. The results showed that ammonium acetate provided much improved resolution compared to ammonium formate when used as a salt additive of mobile phase. However, peak tailings were not eliminated for most of target peaks, especially Orn and Cyst. When acetic acid was added into the mobile phase, the peak shapes of Orn and Cyst were improved and the acetic acid concentration in the eluents was increased. The peak shapes of the rest target compounds, which were already satisfactory, were not affected by the administration of acetic acid in mobile phase. As a result, a mixed solution including A (water, 10 mmol/L ammonium acetate and 0.5 % acetic acid) and B (acetonitrile, 1 mmol/L ammonium acetate, and 0.1 % acetic acid) was chosen as the optimized mobile phase. In this study, the co-eluting analytes were separated by MS^2 using MRM mode enabling unambiguous detection and identification for Trp and Phe ("Analysis of tryptophan" in Supporting Information). For amino acids with similar MS^2 characteristics, like Leu (132.1 → 86.1) and

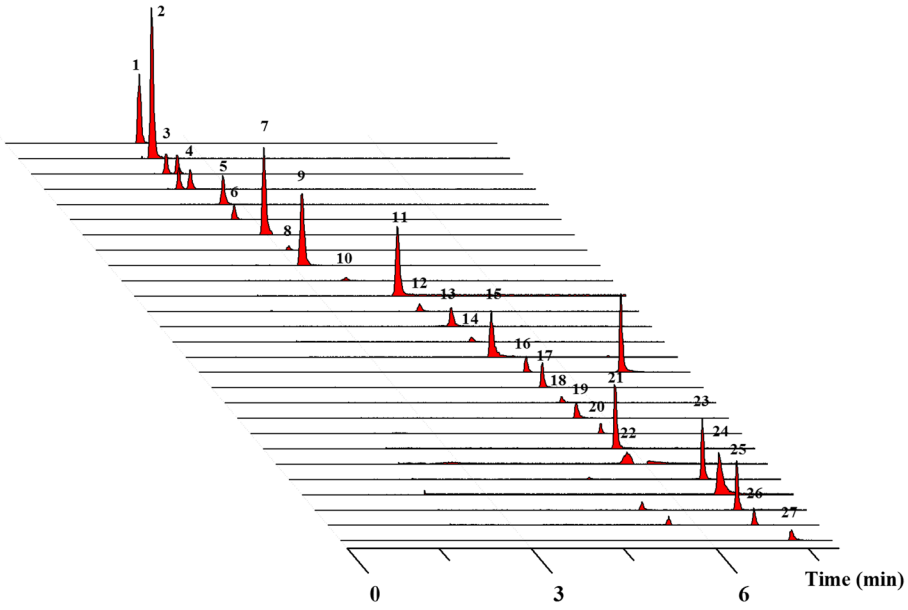

Fig. 3 Representative UHPLC-QTRAP®/MS^2 chromatogram of 27 target compounds in LHG sample (**S11**). *Peaks* **1** Trp, **2** Phe, **3** Leu, **4** Ile, **5** GABA, **6** Met, **7** Val, **8** Pro, **9** Tyr, **10** Cys, **11** Ala, **12** Hpro, **13** Thr, **14** Gly, **15** Glu, **16** Gln, **17** Ser, **18** GSH, **19** Asn, 20 Ala–Gln, **21** Cit, **22** Asp, **23** Arg, **24** His, **25** Lys, **26** Orn, and **27** Cyst

Ile (132.1 → 86.1), as well as Glu (148.1 → 83.9), Gln (147.1 → 84.0), and Lys (147.1 → 83.9), an evident baseline separation was still needed. To separate the above analytes, the elution procedure was optimized as the gradient elution: 0–3 min, 15–17.5 % A; 3–5.5 min, 17.5–30 % A; 5.5–7 min, 30–46 % A, and 7–8 min, 46 % A. The results showed that the optimized elution procedure could fully separate the above five problematic compounds. Additionally, in the preliminary experiments, to confirm the retention time of the above five problematic FAAs, each single standard solution (100 ng/mL) was, respectively, injected into the HILIC-UHPLC-QTRAP®/MS^2, and the retention time of Leu, Ile, Glu, Gln, and Lys were given at 2.19, 2.39, 5.00, 5.34, and 6.87 min using the above selected analytical conditions, respectively (Fig. S3). Finally, each compound was identified by respective special parameters of both MRM and retention time, and a mixture of 27 target compounds was used for the next research of method validation. The analysis time of the 27 target compounds was <7.5 min in this study, which was significantly shorter than that in all other reported LC separating FAAs (Wang et al. 2013; Zhou et al. 2013; Guo et al. 2013). Representative MRM chromatogram for LHG sample and 1.0 μg/mL mixture standard analytes are shown in Fig. 3 and Fig. S4, respectively.

Column temperature and flow rate were also important parameters that affect the retention of polar compounds in HILIC. Temperatures ranging from 15 to 40 °C and flow rates ranging from 0.1 to 0.5 mL/min were also investigated in this study (Fig. S5). It was suggested that separation achieved optimum when the flow rate was 0.4 mL/min and the column temperature was at 35 °C. Additionally, deviations of the retention time of 27 target compounds were also investigated in Acquity UPLC BEH Amide column, and the results indicated the robustness of retention times with the development HILIC method (Supporting Information "Retention time deviations in HILIC column" and Table S4).

Method validation

Validation of the method strictly complied with the International Conference on Harmonization (ICH) regulations for confirmation analysis procedure. Several performance parameters were studied, including matrix effects, linearity, LOD, LOQ, precision, repeatability, stability, and recovery (Rambla-Alegre et al. 2012).

Matrix effects were studied to ensure the bias-free analyzed results. LHG samples were spiked (spiking trials repeated 3 times), before extraction, with the FAAs and small peptides at low (0.2 mg/kg) and high (2.0 mg/kg) concentrations, respectively. The slopes of the calibration plots were compared with results obtained when the whole process was applied to standard solutions of the FAAs and small peptides. The slope ratio of matrix curve to neat solution curve was calculated; the ratio value of 1 indicated no matrix effect. When LHG samples were tested, most of the ratio values were close to 1, implying that sample preparation method and HILIC-UHPLC-QTRAP®/MS^2 conditions were suitable for the determination of 27 target compounds in such complex functional food matrices (Zhou et al. 2013). The results of matrix effect are listed in Table S5 and processing details were described in Supporting Information "matrix effects."

The calibration functions obtained by plotting the peak area *versus* the concentration of the compound were linear,

and the determination coefficient higher than 0.992 for all compounds. LODs for the 27 compounds were estimated to be 0.15–63.66 ng/mL, and the LOQs were 0.45–190.85 ng/mL. Such levels of sensitivity for FAA detection were noticeably higher than that in other reported studies using MRM methods (Wang et al. 2013; Zhou et al. 2013; Guo et al. 2013), probably because of the sensitive MRM detection by 5500 QTRAP®/MS^2 and the sharp peak shapes observed using 50 % acetonitrile as the solvent. Compared with the LC-HRMS methods, the LODs and LOQs values of many compounds from the report (Gökmen et al. 2012) were obvious higher 2 times than this study. Additionally, the sensitivities of Phe, Met, Pro, and Tyr in this study were also significantly higher than the one report using LC-HRMS method (Troise et al. 2015), and the sensitivities of the rest 16 compounds from the report (Troise et al. 2015) were close to our study. Although the reported LC-HRMS had considerable sensitivity (Troise et al. 2015), the main drawback of LC-HRMS was the poor separation of isomers (Ile and Leu). The intra- and inter-day precision variations, expressed as RSD, were from 0.82 to 3.69 % and from 1.89 to 6.28 %, respectively. The repeatability and stability presented as RSD were in the range from 1.92 to 6.81 % and from 1.13 to 6.69 %. Trueness was estimated through recovery studies. Before extraction, different aliquots of LHG samples (**S11**, $n = 3$) were spiked at two levels, 0.1 and 1.0 mg/g with the target compounds and were extracted with the developed method. The results indicated that recoveries ranged from 91.6 % (Lys) to 107.5 % (Ile) for the selected compounds at 0.1 mg/g, and from 90.5 % (Lys) to 106.9 % (Gln) at 1.0 mg/g. These results suggested the established method could provide sufficient accuracy for the quantification of LHG samples. Detailed data are summarized in Table S5.

Method application

The developed HILIC-UHPLC-QTRAP®/MS^2 method was applied to simultaneous quantification of 24 FAAs and 3 small peptides in three parts of tissue culture (**S4**) and cultivate (**S14**) of LHG (cv. Qingpiguo from Yongfu, Guangxi) which was a major merchandise resource. The total content of all target compounds ranged from 1581.9 to 12352.2 μg/g and the contents varied greatly among different parts. Distribution of total target compounds in three parts of LHG ranked as follows, mesocarp and endocarp, seed, epicarp (from highest to lowest). It was observed that the major FAAs in the mesocarp and endocarp were GABA, Pro, Gln and Glu, and Gln, Glu, Cyst (dipeptide), and GSH (tripeptide) in the seed. All target compounds showed the lowest level in the epicarp. The thiol small peptides of Cyst and GSH, which were important factors regulating cellular homeostasis and antioxidant defense networks, were principally distributed in the seed of LHG. In general, the mesocarp and endocarp part was distinguished from the other two parts. For example, the total contents of essential amino acids and 4 non-protein amino acids (GABA, Hpro, Cit and Orn) were 1768.25 and 1125.31 μg/g in the mesocarp and endocarp, respectively, and accounted for approximately 35.6 and 11.3 % of total amino acids; whereas they accounted for approximately 11.8–13.3 and 25.3–32.4 % in the other two parts (Fig. S6). There were many patents today about LHG products such as teas, glazed fruit, beverages, and tipple. From nutritional point of view, this study indicated that the mesocarp and endocarp of LHG might be the most proper raw material for manufacturing the LHG products.

The proposed method was also applied to analyze the different sources of 40 batches samples (using full fruit) from three cultivated varieties of LHG (Qingpiguo, Hongmaoguo and Dongguaguo) with different cultivated forms and regions (listed in Table 1). As shown in Table 2, the total content of 27 target compounds varied from 151.2 (**S17**) to 8689.6 μg/g (**S37**) among LHG samples from different sources. Our study revealed that FAAs exhibited a different concentration level in LHG samples with different cultivated varieties, forms, and regions. Obviously, the contents of total target compounds in Hongmaoguo (**S34**–**S36**), and Dongguaguo (**S40** and **S37**) samples were approximately 5 and 8 times higher than those in the other Qingpiguo samples, respectively. Furthermore, for a specified cultivated variety (Qingpiguo) and form (Cultivate), the total contents of FAAs and small peptides in major production region (Yongfu, Guangxi) were 3-5 times higher than other regions samples. Finally, for the specified cultivated variety of Qingpiguo and region of Yongfu, Guangxi, the total contents of target compounds in three cultivated forms of LHG were ranked in the order of cultivate > cuttage > tissue culture, from the highest to the lowest. As for the individual compounds, the results showed that all of these LHG samples were rich in the essential amino acids such as Trp, Phe, Leu, Ile, Met, Val, Thr, and Lys. Except for GABA and Hpro, the contents of the other two non-protein amino acids (Cit and Orn) in LHG were only detected at a trace level. In particular, the non-protein amino acid GABA was considered as an important neurotransmitter signal that had blood pressure lowering activity in functional foods (Syu et al. 2008). From a nutritional and functional point of view, these results indicated that LHG was healthy and nutritional as it was rich in FAAs, especially the essential FAAs and GABA. However, the contents of the three active small peptides (GSH, Ala–Gln, and Cyst) were all very low in LHG samples.

In a study published earlier, 18 amino acids from the hydrolysis products of tissue culture LHG protein were determined using ion-exchange chromatography in an

Table 2 Contents of 27 target compounds in the tested samples of LHG (μg/g)

	Trp	Phe	Leu	Ile	GABA	Met	Val	Pro	Tyr	Cys	Ala	Hpro	Thr	Gly
S1	3.0	190.3	219.4	189.6	50.0	2.1	342.1	78.9	257.8	5.6	322.9	24.5	+	+
S2	1.1	114.2	110.7	133.4	5.1	1.7	211.7	86.4	111.3	4.5	454.4	8.8	0.1	1.2
S3	0.8	26.2	13.1	20.6	8.5	2.0	28.7	11.3	19.8	7.3	70.0	3.9	+	+
S4	1.4	53.8	51.6	43.7	15.4	1.8	65.0	82.1	50.9	3.5	214.6	7.3	+	+
S5	2.7	136.7	154.4	129.4	30.3	2.2	184.2	115.2	147.3	9.6	252.8	12.8	+	8.2
S6	1.3	25.6	17.2	19.6	7.0	2.1	33.3	12.8	21.2	6.2	73.2	5.2	ND	+
S7	1.5	78.3	84.2	70.0	8.5	2.0	105.9	106.3	77.0	4.2	275.2	8.6	+	+
S8	1.3	115.1	130.7	137.8	8.9	1.6	234.1	90.2	128.2	4.6	417.1	11.5	+	+
S9	0.7	86.2	105.9	98.9	4.9	1.5	175.8	77.6	105.2	2.5	350.8	6.5	+	+
S10	0.9	86.1	88.8	81.5	6.8	2.3	156.5	75.4	98.0	3.7	331.6	5.3	+	+
S11	4.8	365.1	439.6	568.0	34.9	1.5	721.8	137.5	349.6	3.5	872.3	30.8	35.2	8.6
S12	6.1	83.6	87.3	74.9	25.8	5.4	96.5	101.4	82.0	12.9	311.4	10.8	+	+
S13	4.7	434.2	573.6	407.7	96.3	1.5	644.6	417.4	572.2	1.1	1105.6	19.9	84.8	10.6
S14	2.2	270.3	327.1	338.0	26.7	1.6	508.7	187.3	292.9	5.2	708.1	29.8	5.6	2.9
S15	1.1	27.4	17.0	21.1	6.1	2.0	34.0	8.7	20.6	5.8	72.4	4.5	+	+
S16	4.4	60.3	53.3	49.4	14.3	5.5	73.7	41.4	64.0	19.6	97.0	12.2	+	+
S17	0.8	22.0	13.8	15.9	1.5	1.5	18.9	5.4	13.8	+	31.1	2.2	ND	+
S18	2.5	46.3	44.8	48.1	13.4	2.5	60.6	38.0	41.4	6.8	134.6	7.7	+	+
S19	2.7	97.3	102.9	144.1	8.2	1.8	222.4	42.5	111.6	4.1	402.4	4.6	+	+
S20	1.8	89.2	102.5	105.3	15.2	1.7	164.1	84.7	106.7	5.2	274.0	5.8	+	+
S21	1.9	265.0	304.7	271.1	13.0	1.5	460.2	254.9	341.1	3.0	1084.6	20.1	10.3	+
S22	3.1	297.1	363.0	303.5	70.5	1.5	491.0	193.7	394.7	7.9	571.6	32.9	13.8	1.8
S23	1.3	171.3	203.7	176.3	14.2	1.3	303.9	125.5	203.5	2.0	479.4	13.0	+	+
S24	1.4	53.8	49.8	43.4	8.3	2.2	63.6	97.1	46.5	8.5	219.7	5.1	+	+
S25	1.2	42.0	37.8	31.9	5.0	1.9	53.1	29.5	38.6	4.9	121.6	5.3	+	+
S26	1.9	89.2	100.7	84.2	9.2	2.3	119.2	145.9	87.3	5.8	347.1	10.6	+	+
S27	1.7	71.5	73.2	60.7	11.2	2.0	95.9	47.8	79.5	6.5	117.3	10.4	+	+
S28	2.7	210.9	320.5	162.0	97.6	2.5	227.2	728.6	230.7	4.1	1826.9	20.5	10.3	38.3
S29	1.2	190.0	217.3	212.2	32.8	1.7	345.1	107.9	223.6	4.3	387.9	19.1	+	+
S30	0.7	173.8	216.2	224.4	8.1	1.4	364.0	263.4	215.9	2.1	835.9	14.0	+	+
S31	1.3	70.9	72.3	62.4	11.3	2.3	95.1	37.3	81.8	5.4	118.2	9.8	+	+
S32	1.0	39.9	33.5	33.7	5.5	2.1	52.6	19.8	38.5	6.7	69.0	3.7	+	+
S33	1.8	79.6	84.7	64.3	16.9	1.9	99.6	59.5	96.0	11.2	173.9	7.3	+	+
S34	33.3	511.9	391.5	587.6	48.5	1.4	613.6	551.3	743.8	4.7	880.4	55.3	197.3	22.7
S35	28.0	465.5	423.4	381.1	69.2	7.9	444.2	181.5	625.1	3.9	400.1	18.7	124.2	22.5
S36	21.6	365.5	290.2	312.2	66.2	9.1	317.6	412.3	627.6	11.5	266.1	14.1	128.7	23.0
S37	9.8	648.4	883.1	693.3	146.1	1.4	967.0	747.9	850.9	6.0	1353.3	56.4	93.2	22.4
S38	2.5	100.6	117.6	104.9	26.3	3.1	161.3	60.6	127.8	7.8	210.8	14.0	+	+
S39	2.2	302.6	472.2	390.8	118.1	1.6	679.5	219.5	437.5	4.1	950.6	33.2	26.7	10.6
S40	9.6	628.5	746.5	752.4	153.9	1.7	98.7	734.1	729.4	4.1	1183.7	44.9	80.1	18.3

	Glu	Gln	Ser	GSH	Asn	Ala Gln	Cit	Asp	Arg	His	Lys	Orn	Cyst	Total
S1	+	8.5	5.0	+	120.0	ND	7.2	301.6	65.6	9.6	+	4.8	3.8	2212.2
S2	0.8	5.5	0.8	8.7	63.5	ND	7.6	243.6	+	4.5	+	+	4.2	1581.4
S3	+	4.2	+	0.2	+	ND	9.2	28.6	+	4.9	+	+	5.0	264.2
S4	+	4.6	2.3	4.2	+	+	9.5	89.4	+	5.1	+	+	5.3	711.2
S5	+	6.6	3.2	+	26.4	+	11.3	125.3	68.3	9.9	+	2.2	5.3	1444.4
S6	+	6.0	+	2.0	+	+	9.8	61.0	+	4.1	+	+	5.0	312.4
S7	+	6.9	+	+	11.1	+	10.9	153.3	+	7.4	+	+	4.6	1015.6

Table 2 continued

	Glu	Gln	Ser	GSH	Asn	Ala Gln	Cit	Asp	Arg	His	Lys	Orn	Cyst	Total
S8	+	5.8	5.3	8.6	134.6	+	7.4	328.4	+	5.4	+	+	3.8	1780.2
S9	+	4.9	1.7	+	34.2	+	6.7	209.5	+	3.5	ND	+	3.3	1280.1
S10	+	5.4	0.7	12.2	11.8	+	7.2	171.5	+	4.4	+	16.3	3.7	1170.0
S11	63.9	23.7	66.0	10.9	615.6	+	9.5	806.1	323.8	19.3	+	5.7	3.3	5520.9
S12	+	8.9	26.2	9.7	9.6	2.3	18.6	122.3	19.9	11.1	+	+	5.4	1132.0
S13	68.8	10.2	55.9	41.0	371.6	0.2	7.8	537.8	209.7	21.2	+	16.4	3.7	5718.5
S14	+	9.9	24.2	14.6	378.0	+	8.2	595.1	59.5	10.3	+	+	3.5	3809.4
S15	+	5.8	+	0.0	+	ND	9.2	25.7	+	4.3	+	+	4.7	270.4
S16	+	9.5	+	9.1	+	0.1	15.6	62.9	+	10.6	+	+	6.1	608.9
S17	+	4.2	+	+	ND	+	8.6	5.0	+	2.6	+	+	3.9	151.2
S18	+	6.5	+	+	+	+	11.0	88.4	+	6.6	+	+	5.2	564.3
S19	+	7.2	9.2	ND	72.6	+	6.6	293.1	9.1	8.4	+	+	3.3	1553.9
S20	+	4.3	7.9	6.3	13.3	+	8.1	212.3	38.7	8.1	ND	+	4.4	1259.4
S21	+	7.1	18.1	+	426.0	+	8.0	658.0	110.4	9.7	ND	+	3.7	4272.5
S22	+	8.0	24.0	+	216.2	ND	8.0	425.5	178.4	19.9	+	+	4.2	3630.0
S23	+	6.6	4.5	2.6	166.8	+	7.0	418.4	5.5	6.3	+	+	3.6	2316.7
S24	+	4.5	+	0.1	+	+	10.6	91.5	+	5.7	+	+	6.0	717.7
S25	+	5.3	+	5.6	+	+	10.1	45.1	+	4.5	+	+	4.7	448.0
S26	+	5.4	3.1	2.3	39.7	+	10.0	139.1	+	6.5	+	11.7	4.6	1225.7
S27	+	5.4	+	8.1	+	ND	11.9	56.6	5.4	6.5	+	+	5.3	676.7
S28	398.8	8.8	132.6	13.5	606.8	ND	10.2	631.4	125.6	18.5	+	+	4.1	5833.0
S29	+	5.2	+	+	87.5	+	7.2	273.1	1.4	5.4	+	+	4.0	2127.1
S30	+	4.9	8.2	+	225.4	ND	6.7	519.0	+	4.8	+	+	3.1	3091.9
S31	+	6.0	ND	1.7	+	+	10.3	65.5	4.7	7.5	+	+	5.4	669.2
S32	+	4.8	+	2.3	+	0.1	9.6	55.7	+	3.2	+	+	4.3	385.9
S33	+	5.7	+	14.4	+	ND	11.2	93.9	0.5	6.2	+	+	4.2	832.8
S34	326.8	19.4	132.6	+	677.0	1.9	43.6	996.1	495.4	120.1	68.7	63.7	3.6	7592.0
S35	222.6	63.2	117.6	0.5	367.3	0.1	16.8	622.1	368.6	71.4	77.3	22.9	3.8	5149.5
S36	215.1	31.4	204.3	14.4	209.1	ND	92.8	598.6	343.4	77.4	103.6	46.9	5.8	4808.2
S37	127.9	10.0	78.5	2.3	722.7	1.0	19.8	580.7	588.8	47.9	9.4	17.9	3.9	8689.6
S38	+	7.7	+	+	+	ND	11.4	134.1	27.1	7.5	+	+	5.2	1130.2
S39	6.2	6.8	41.3	+	357.8	+	10.0	658.6	203.3	25.6	ND	6.8	3.5	4968.8
S40	82.8	8.7	55.3	17.0	664.4	+	16.6	481.7	621.1	67.5	18.5	13.5	3.9	7236.7

automatic amino acid analyzer (He et al. 2012). However, so far there was no updated study reporting the detailed information of free amino acids in LHG. In general, the differences between the previously reported and this study were considered as the reflection of variances in the technique, instrument, conditions of experiments, and the sources of samples.

Chemometric analysis

To evaluate whether the FAAs and small peptides profiles could effectively distinguish different sources of 40 batches LHG samples, PCA was chosen because it was a commonly used technique for reducing the dimensionality of the data, which could best differentiate the differences between samples and allow the visualization of cluster and outliers (Kumar et al. 2014). PCA worked by replacing a large number of variables with a new, smaller number of variables, namely *PC*s, which represented most of the original data (Bro and Smilde 2014; Vanloot et al. 2014). PC1 explained 57.2 % of the total variance in the data set while PC2 explained 16.9 %. The first two *PC*s simplified the multidimensional dataset to a two-dimensional dataset. The remaining *PC*s with less effect on the model were discarded. According to the loading plot, PC1 showed a positive correlation with Tyr, Phe, and Arg, while PC2 showed a positive correlation with Met, Cyst, Cys, Cit, and Lys. Interestingly, Cys and Cyst (Cys–Cys) exhibited good correlation with PC2; thus, Cys and Cyst might be important markers for distinguish LHG samples. The component

loading matrix (Fig. 4a) suggested that all 27 compounds may contribute to the classification of the samples. As shown in the score plot (Fig. 4b) of PC1 (*x* axis) and PC2 (*y* axis), the samples can be classified into three groups, and the main contributors to the three groups were representing as follows: **S34–36** were in group **1**, **S37** and **S40** were in group **2**, and the rest samples were in group **3**. Obviously, it could be observed that samples with similar chemical profiles were commonly divided into same group. As for **S37** and **S40** (Dongguaguo), the content of each target compound was relatively higher than that in the other samples. However, for all the cultivated variety of Qingpiguo (a major merchandise resource) samples, the contents of the target compounds were relatively lower than those in the other cultivated varieties samples (Dongguaguo and Hongmaoguo). Thus, it could be concluded that Dongguaguo had a high content of these nutritional and functional substances than other cultivated varieties. Dongguaguo therefore might be the preferable raw material for making the LHG products, from the nutritional and functional point of view. These results indicated the factor of cultivated variety of LHG might play an important role in samples classification, and chemical profiles of FAAs and small peptides of LHG were dominantly affected by cultivated varieties.

Results from the PCA analysis could discern the differences between samples, but could not provide a clear class separation for the groups of samples under investigation (Hakimzadeh et al. 2014). In order to make a satisfactory classification and prediction of the different LHG samples (**S1–40**), a feasible chemometric method should be employed. CP-ANN, one of the most popular ANNs, recently has been regarded as a powerful supervised classification tool. CP-ANN mimicked the action of a biological network of neurons, where each neuron accepted different signals from neighboring neurons and processed them for a defined number of epochs (Ballabio and Vasighi 2012). The data matrix of 40 samples with 27 variables was used as input for training network. The elements of the output vector were assigned to be A, B, or C according to the PCA clusters of samples (**1–3**). The parameters of network size and epochs (i.e., number of learning iterations) were optimized by grid search method. The network size of 10 × 10 (100 neurons) and the epochs of 100 were obtained as the optimum parameters for the studied samples. After training the network, the samples were separated correctly on the map ("Repetitive operation of CP-ANN" in Supporting Information). Figure 5 showed distribution of the samples on the CP-ANN map where the semblable samples were separated in a more distinct pattern, especially for cluster C (group **3** involving 35 samples from PCA) based on the 27 target compounds. Comparing with the PCA scores plot in Figs. 4b, 5 showed good coincidence between the obtained results. For example, **S37** and **S40** in the score plot were located close to **S28** and **S13**. Therefore, the data structure observed in the CP-ANN map was in concordance with the score plot of PCA although the basics of the two methods in samples classification were different.

Conclusion

Firstly, UAE was optimized by PB and BBD designs to obtain maximum extraction efficiency of 27 FAAs and

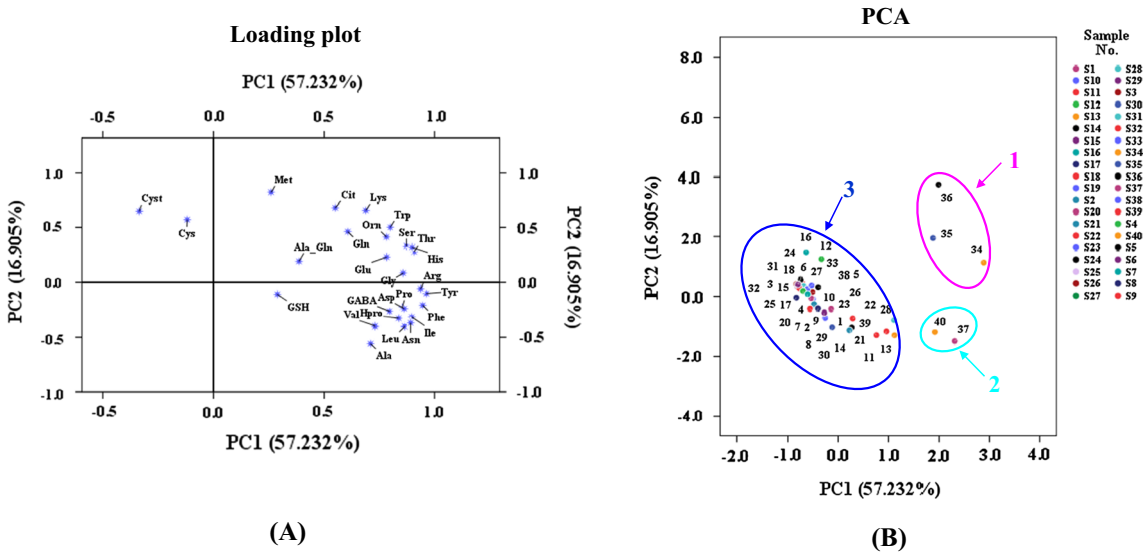

Fig. 4 Loading plot **a** obtained by PCA of the 27 target compounds and the scatter plot **b** obtained by PCA of the 40 batches of LHG samples. The 40 batches of LHG samples are the same as in Table 1, and the 27 amino acids are the same as in Table S1

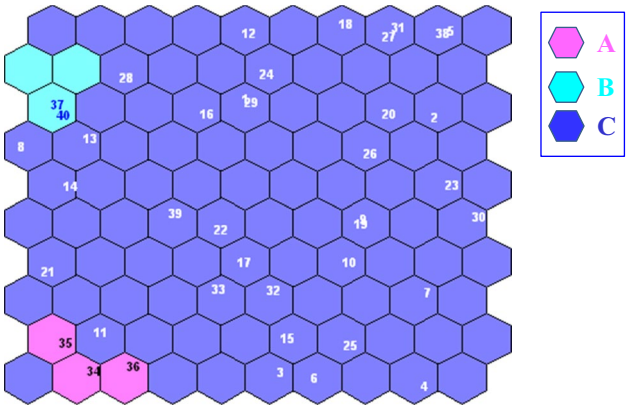

Fig. 5 Samples distribution on the CP-ANN map

small peptides. The optimized UAE technique provided higher extraction efficiency of target compounds in LHG on a simple and inexpensive instrument. Secondly, a sensitive, rapid, accurate, and repeatable HILIC-UHPLC-QTRAP®/MS^2 method was successfully developed and employed to determine 27 underivatized FAAs and small peptides in LHG. By combining the rapid HILIC-UHPLC and sensitive QTRAP®/MS^2 methods, the 27 polar compounds were separated appropriately within 7.5 min. This method could be used to analyze FAAs and small peptides simultaneously in medicinal or edible plant materials and could also be considered as an attractive alternative for laboratories to detect these compounds in bio-samples such as blood or urine. The analysis results of LHG showed that LHG was a kind of healthy food rich in essential FAAs. Finally, a chemometric strategy combining unsupervised PCA and supervised CP-ANN was developed for comprehensive analysis of samples diversity, according to the levels of the 27 target markers. Results showed that this chemometric method was not only accurate and unbiased, but also acquired additional information which made the data set more complete.

Acknowledgments We are pleased to thank Dr. Bin Hao (School of Pharmacy, Shanghai Jiao Tong University) for the help in HILIC-UHPLC-QTRAP®/MS^2 tests.

Conflict of interest The authors have no conflicts of interest.

References

Akamatsu S, Mitsuhashi T (2013) Development of a simple analytical method using capillary electrophoresis-tandem mass spectrometry for product identification and simultaneous determination of free amino acids in dietary supplements containing royal jelly. J Food Compos Anal 30:47–51

Bakirdere S, Bramanti E, D'ulivo A, Ataman OY, Mester Z (2010) Speciation and determination of thiols in biological samples using high performance liquid chromatography-inductively coupled plasma-mass spectrometry and high performance liquid chromatography-Orbitrap MS. Anal Chim Acta 680:41–47

Ballabio D, Vasighi M (2012) A MATLAB toolbox for self organizing maps and supervised neural network learning strategies. Chemometr Intell Lab Syst 118:24–32

Ballabio D, Vasighi M, Consonni V, Kompany-Zareh M (2011) Genetic algorithms for architecture optimisation of counter-propagation artificial neural networks. Chemometr Intell Lab Syst 105:56–64

Bernal J, Mendiola J, Ibáñez E, Cifuentes A (2011) Advanced analysis of nutraceuticals. J Pharm Biomed 55:758–774

Bro R, Smilde AK (2014) Principal component analysis. Anal Methods 6:2812–2831

Carrera C, Ruiz-Rodríguez A, Palma M, Barroso CG (2015) Ultrasound-assisted extraction of amino acids from grapes. Ultrason Sonochem 22:499–505

Castro-Perez J, Plumb R, Liang L, Yang E (2005) A high-throughput liquid chromatography/tandem mass spectrometry method for screening glutathione conjugates using exact mass neutral loss acquisition. Rapid Commun Mass Spectrum 19:798–804

Chen X, Zhuang J, Liu J, Lei M, Ma L, Chen J, Shen X, Hu L (2011) Potential AMPK activators of cucurbitane triterpenoids from *Siraitia grosvenorii* Swingle. Bioorgan Med Chem 19:5776–5781

Dago À, González I, Ariño C, Manuel Díaz-Cruz J, Esteban M (2014) Chemometrics applied to the analysis of induced phytochelatins in *Hordeum vulgare* plants stressed with various toxic non-essential metals and metalloids. Talanta 118:201–209

Deshmukh S, Kamde K, Jana A, Korde S, Bandyopadhyay R, Sankar R, Bhattacharyya N, Pandey R (2014) Calibration transfer between electronic nose systems for rapid in situ measurement of pulp and paper industry emissions. Anal Chim Acta 841:58–67

Duan L, Guo L, Liu K, Liu E, Li P (2014) Characterization and classification of seven *Citrus* herbs by liquid chromatography-quadrupole time-of-flight mass spectrometry and genetic algorithm optimized support vector machines. J Chromatogr A 1339:118–127

Fiechter G, Mayer HK (2011) UPLC analysis of free amino acids in wines: profiling of on-lees aged wines. J Chromatogr B 879:1361–1366

Friedman M (2004) Applications of the ninhydrin reaction for analysis of amino acids, peptides, and proteins to agricultural and biomedical sciences. J Agr Food Chem 52:385–406

Gillette MA, Carr SA (2013) Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry. Nat Methods 10:28–34

Gökmen V, Serpen A, Mogol BA (2012) Rapid determination of amino acids in foods by hydrophilic interaction liquid chromatography coupled to high-resolution mass spectrometry. Anal Bioanal Chem 403:2915–2922

Guna M, Londry FA (2011) Tandem ion trap design with enhanced mass analysis capabilities for large populations of ions. Anal Chem 83:6363–6367

Guo S, Duan J, Qian D, Tang Y, Qian Y, Wu D, Su S, Shang E (2013) Rapid determination of amino acids in fruits of *Ziziphus jujuba* by hydrophilic interaction ultra-high-performance liquid chromatography coupled with triple-quadrupole mass spectrometry. J Agr Food Chem 61:2709–2719

Hakimzadeh N, Parastar H, Fattahi M (2014) Combination of multivariate curve resolution and multivariate classification techniques for comprehensive high-performance liquid chromatography-diode array absorbance detection fingerprints analysis of *Salvia reuterana* extracts. J Chromatogr A 1326:63–72

He C, Zhu X, Liu L, He W (2012) Analysis and comparison the nutrient compositions of fresh and dry tissue culture seedling Luohanguo. Guihaia 32:706–709

Ilisz I, Berkecz R, Péter A (2008) Application of chiral derivatizing agents in the high-performance liquid chromatographic separation of amino acid enantiomers: a review. J Pharmaceut Biomed 47:1–15

Kıvrak İ, Kıvrak Ş, Harmandar M (2014) Free amino acid profiling in the giant puffball mushroom (*Calvatia gigantea*) using UPLC–MS/MS. Food Chem 158:88–92

Kumar N, Bansal A, Sarma G, Rawal RK (2014) Chemometrics tools used in analytical chemistry: an overview. Talanta 123:186–199

Kuśmierek K, Bald E (2008) Reduced and total glutathione and cysteine profiles of citrus fruit juices using liquid chromatography. Food Chem 106:340–344

Lahrichi SL, Affolter M, Zolezzi IS, Panchaud A (2013) Food peptidomics: large scale analysis of small bioactive peptides-a pilot study. J Proteomics 88:83–91

Li C, Lin L, Sui F, Wang Z, Huo H, Dai L, Jiang T (2014) Chemistry and pharmacology of *Siraitia grosvenorii*: a review. Chin J Nat Med 12:89–102

Megías C, Cortés-Giraldo I, Girón-Calle J, Vioque J, Alaiz M (2015) Determination of L-canavanine and other free amino acids in *Vicia disperma* (Fabaceae) seeds by precolumn derivatization using diethyl ethoxymethylenemalonate and reversed-phase high-performance liquid chromatography. Talanta 131:95–98

Meisel H, FitzGerald RJ (2003) Biofunctional peptides from milk proteins: mineral binding and cytomodulatory effects. Curr Pharm Des 9:1289–1295

Mudiam MKR, Ratnasekhar C (2013) Ultra sound assisted one step rapid derivatization and dispersive liquid–liquid microextraction followed by gas chromatography-mass spectrometric determination of amino acids in complex matrices. J Chromatogr A 1291:10–18

Ni Y, Zhuang H, Kokot S (2014) A high performance liquid chromatography and electrospray ionization mass spectrometry method for the analysis of the natural medicine, *Forsythia Suspensa*. Anal Lett 47:102–116

Pirman DA, Reich RF, Kiss A, Heeren RMA, Yost RA (2013) Quantitative MALDI tandem mass spectrometric imaging of cocaine from brain tissue with a deuterated internal standard. Anal Chem 85:1081–1089

Rambla-Alegre M, Esteve-Romero J, Carda-Broch S (2012) Is it really necessary to validate an analytical method or not? That is the question. J Chromatogr A 1232:101–109

Rezazadeh M, Yamini Y, Seidi S, Aghaei A (2015) Pulsed electromembrane extraction for analysis of derivatized amino acids: a powerful technique for determination of animal source of gelatin samples. Talanta 136:190–197

Sarmadi BH, Ismail A (2010) Antioxidative peptides from food proteins: a review. Peptides 31:1949–1956

Schultz CL, Moini M (2003) Analysis of underivatized amino acids and their D/L-enantiomers by sheathless capillary electrophoresis/electrospray ionization-mass spectrometry. Anal Chem 75:1508–1513

Stokvis E, Rosing H, Beijnen JH (2005) Stable isotopically labeled internal standards in quantitative bioanalysis using liquid chromatography/mass spectrometry: necessity or not? Rapid Commun Mass Spectrum 19:401–407

Syu KY, Lin C, Huang H, Lin J (2008) Determination of theanine, GABA, and other amino acids in green, oolong, black, and Pu-erh teas with dabsylation and high-performance liquid chromatography. J Agric Food Chem 56:7637–7643

Takasaki M, Konoshima T, Murata Y, Sugiura M, Nishino H, Tokuda H, Matsumoto K, Kasai R, Yamasaki K (2003) Anticarcinogenic activity of natural sweeteners, cucurbitane glycosides, from *Momordica grosvenori*. Cancer Lett 198:37–42

Tian M, Zhang J, Mohamed AC, Han Y, Guo L, Yang L (2014) Efficient capillary electrophoresis separation and determination of free amino acids in beer samples. Electrophoresis 35:577–584

Troise AD, Fiore A, Roviello G, Monti SM, Fogliano V (2015) Simultaneous quantification of amino acids and Amadori products in foods through ion-pairing liquid chromatography-high-resolution mass spectrometry. Amino Acids 47:111–124

Tuberoso CIG, Congiu F, Serreli G, Mameli S (2015) Determination of dansylated amino acids and biogenic amines in Cannonau and Vermentino wines by HPLC-FLD. Food Chem 175:29–35

Vanloot P, Bertrand D, Pinatel C, Artaud J, Dupuy N (2014) Artificial vision and chemometrics analyses of olive stones for varietal identification of five French cultivars. Comput Electron Agric 102:98–105

Wang H, Duan J, Guo S, Qian D, Shang E (2013) Development and validation of a hydrophilic interaction ultra-high-performance liquid chromatography with triple quadrupole MS/MS for the absolute and relative quantification of amino acids in *Sophora alopecuroides* L. J Sep Sci 36:2244–2252

Wu G (2013) Functional amino acids in nutrition and health. Amino Acids 45:407–411

Yan J, Cai Y, Wang Y, Lin X, Li H (2014) Simultaneous determination of amino acids in tea leaves by micellar electrokinetic chromatography with laser-induced fluorescence detection. Food Chem 143:82–89

Zhang H, Yang H, Zhang M, Wang Y, Wang J, Yau L, Jiang Z, Hu P (2012) Identification of flavonol and triterpene glycosides in *Luo-Han-Guo* extract using ultra-high performance liquid chromatography/quadrupole time-of-flight mass spectrometry. J Food Compos Anal 25:142–148

Zhang K, Wong JW, Krynitsky AJ, Trucksess MW (2014) Determining mycotoxins in baby foods and animal feeds using stable isotope dilution and liquid chromatography tandem mass spectrometry. J Agr Food Chem 62:8935–8943

Zhou G, Pang H, Tang Y, Yao X, Mo X, Zhu S, Guo S, Qian D, Qian Y, Su S (2013) Hydrophilic interaction ultra-performance liquid chromatography coupled with triple-quadrupole tandem mass spectrometry for highly rapid and sensitive analysis of underivatized amino acids in functional foods. Amino Acids 44:1293–1305

Zhou G, Yao X, Tang Y, Qian D, Su S, Zhang L, Jin C, Qin Y, Duan J (2014) An optimized ultrasound-assisted extraction and simultaneous quantification of 26 characteristic components with four structure types in functional foods from ginkgo seeds. Food Chem 158:177–185

Zhou G, Fu L, Li X (2015) Optimisation of ultrasound-assisted extraction conditions for maximal recovery of active monacolins and removal of toxic citrinin from red yeast rice by a full factorial design coupled with response surface methodology. Food Chem 170:186–192