

RAPID EVALUATION OF GARLIC CULTIVARS BY FRONT-FACE FLUORESCENCE AND INDEPENDENT COMPONENT ANALYSIS

X.Q. WU^{a,b}, H. XU^{c*} and J.Y. LIU^b

^aCollege of Pharmacy, Jinan University, Guangzhou 510632. China

^bDepartment of Food Science and Engineering, Jinan University. 601 Huangpu Avenue West, Guangzhou 510632. China

^cCollege of Life Science, Jiaying University, Meizhou 514015. China

(Received: 20 October 2018; accepted: 18 February 2019)

Garlic is widely used as food flavouring, and China is the world's largest garlic producer and exporter. To develop a convenient technique for evaluation of garlic cultivars would be worthwhile, and it would have wide application in such a huge market. In this research, 3D front-face fluorescence data of 8 garlic cultivars were recorded, and independent component analysis was used to decompose the overall fluorescence spectra into six independent components. The first, second, and fourth independent components showed a big difference among the cultivars, and the chemical fluorophores behind these three components were specified as protein, vitamin B6, and ATP, respectively, as fluorescent markers for evaluation. The result showed that all 8 cultivars cluster separately. The cultivar "YNQJ" have the highest quality in terms of protein and vitamin B6, "NXYC" and "SDHB" have the highest content of ATP, while "SXXA" is the poorest in terms of protein and ATP, and "HNJZ" has the lowest content of vitamin B6. Therefore, rapid evaluation of garlic cultivars can be accomplished successfully by using only the proportion values of three properly selected fluorescent markers.

Keywords: garlic cultivar, 3D front-face fluorescence spectroscopy, independent component analysis, fluorophores

Garlic (*Allium sativum*) has been used widely both as food flavouring and as traditional medicine due to its multiple beneficial properties. Interests in the analysis of garlic components have grown in recent years. Garlic has been found to contain approximately 65% water, 28% carbohydrate, 2.3% organosulphur compounds, 2% proteins, 1.2% free amino acids and 1.5% fibre (SANTHOSHA et al., 2013), and their amounts vary greatly depending on genetic, agronomic, and environmental factors (WATERER & SCHMITZ, 1994), which makes evaluation of garlic cultivars significant.

3D front-face fluorescence (3D-FFF), also referred to as the excitation-emission matrices (EEM), is a set of fluorescence spectra acquired at consecutive excitation wavelengths, creating a three-dimensional diagram (JIANG et al., 2010). It has an advantage over conventional fluorescence spectra, because it includes emission spectra excited at many different excitation wavelengths, making it possible to measure complex samples that contain many fluorophores. Relevant information on fluorophores derived from 3D-FFF data can be extracted by multivariate analysis (CHRISTENSEN et al., 2006). Independent component analysis (ICA) is one multivariate analysis method originally developed for separating a mixture of signals to their pure different sources, like dealing with the cocktail-party problem (COMON, 1994; WESTAD & KERMIT, 2009). ICA has been shown to be useful in solving the

* To whom correspondence should be addressed.
E-mail: xhx@jyu.edu.cn

blind source separation problem, and it also has widespread application in similar situations, like extracting the pure spectra of chemical compounds from mixed spectroscopy (HAHN & YOON, 2006). In this paper, ICA was applied to the 3D-FFF spectra of garlic samples to facilitate their interpretation.

1. Materials and methods

1.1. Sample preparation

Different garlic cultivars were collected from eight different locations across the north–south of China. These cultivars were named “HNJZ” (6 samples), “NXYC” (8 samples), “SDHB” (7 samples), “SDHZ” (7 samples), “SDLC” (6 samples), “SXXA” (6 samples), “YNQJ” (6 samples), and “ZJHZ” (6 samples) based on their locations, 52 garlic samples in total. The outer skins of all garlic cloves were removed before use.

1.2. Front-face fluorescence spectroscopy measurement

Front-face fluorescence spectra were obtained with an F-4600 fluorescence spectrophotometer (Hitachi, Tokyo, Japan) and a front-face solid sample cell holder, which was set single position at a 30° incidence angle. The slit widths for excitation and emission were both 5 nm. Measurements were carried out at room temperature directly on the surface of garlic cloves, which were located in solid sample holder to hold the flat surface of clove to the excitation light through a quartz glass. The 3D front-face spectra were obtained using excitation wavelengths from 230 to 400 nm and emission wavelengths from 300 to 550 nm with intervals of 5 nm and 2 nm, respectively, with a scan speed 1200 nm s⁻¹, giving a total of 35 excitation and 126 emission wavelengths. The data corresponding to each sample was arranged in a (126×35) matrix. All elementary matrices of all 52 garlic samples were gathered together in a (52×126×35) 3D cubic array.

1.3. Spectra preprocessing

The 3D-FFF cubic data (EEM landscapes) were needed to be preprocessed to reduce the scattering effects first. A method for handling scattering using interpolation in the areas affected by Rayleigh scatter and avoiding the so-called “inner-filter effects” was applied (BAHRAM et al., 2006). The two bands corresponding to the first- and the second-order Rayleigh scattering were detected and eliminated, and also the missing data were interpolated by shape-preserving piecewise cubic interpolation (YANG & HUIYAN, 1996).

1.4. Data unfolding for multivariate analysis

In order to apply ICA, the cube of data needs to be unfolded to a 2-D matrix. In this study, 3D-FFF data [Samples (52) × Emissions (126) × Excitations (35)] cube is unfolded to generate a 2D [Samples (52) × Emissions × Excitations (4410)] data matrix that keeps the sample dimension unchanged, facilitating interpretation of the results (Fig. 1).

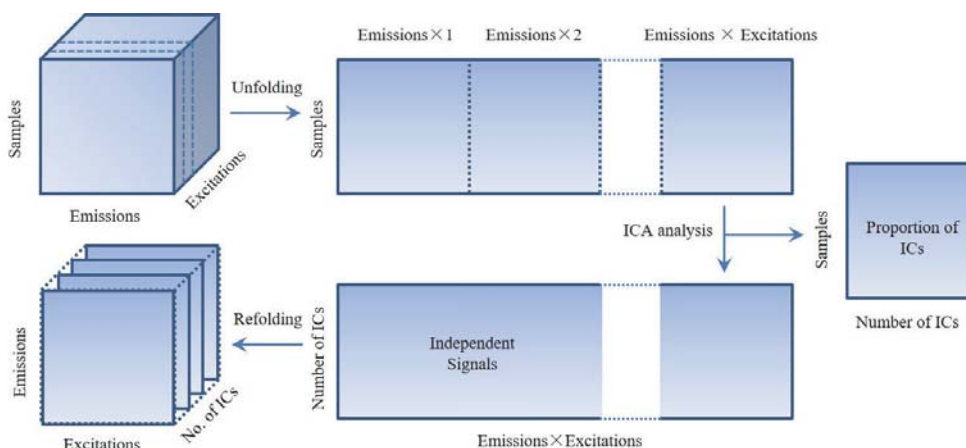


Fig. 1. Graphical representation of the independent component analysis (ICA) procedure in the case of 3D front-face fluorescence data

1.5. Independent component analysis (ICA)

ICA is a blind source separation technique developed to extract the pure underlying signals from a set of signals, where they are mixed in unknown proportions. The ICA algorithm JADE (Joint Approximate Diagonalization of Eigen matrices) (CARDOSO & SOULOUMIAC, 1993; RUTLEDGE & JOUAN-RIMBAUD BOUVERESSE, 2013) was used to analyse the unfolded cube of 3D-FFF matrices. An essential step in the application of ICA is the determination of the optimal number of independent components (IC) to be extracted. This was done using the method called “ICA-by-Blocks” (JOUAN-RIMBAUD BOUVERESSE et al., 2012). After ICA, the matrix [Proportion of ICs] gives the proportion value of each decomposed IC accounting for the mixed fluorescence data for each sample, and the matrix [Independent Signals] stands for the unknown “pure” source spectra, which need to be refolded to obtain spectrochemically interpretable 3D spectra data (Fig. 1).

1.6. Chemical fluorophore behind each IC

In order to explore the actual chemical fluorophores behind each decomposed independent signal, the content of probable fluorophores in garlic, such as protein, vitamins (A, B2, B6, E), ATP, and coenzymes (NADH) were determined, and then, by Pearson product-moment correlation analysis, the high correlation between certain fluorophore concentration and proportion value of each decomposed IC was used as an indicator to unveil the actual fluorophore behind each IC.

Each garlic clove, which was used to get 3D front-face fluorescence data before, was precisely weighed and homogenized using liquid nitrogen. Then the prepared garlic powder was stored at $-20\text{ }^{\circ}\text{C}$ and used for various additional determinations. The further processing were carried out according to the following assay methods.

The total protein of each clove was determined by using a BCA Protein Assay Kit (Beyotime Biotechnology, China). The contents of vitamin A (retinol), E (alpha-tocopherol), B2 (riboflavin), and B6 (pyridoxine) were determined using Shimadzu Prominence UPLC LC20A coupled with a C18 column (250 mm \times 4.6 mm, 5 μm , Dikma Technologies Inc.,

Beijing, China) by diode array detector according to the previous reports (JIN et al., 2012; WU et al., 2016). Retinol, alpha-tocopherol, riboflavin, and pyridoxine reference standards were purchased from Sigma-Aldrich (St. Louis, USA). Meanwhile, the content of ATP and NADH was determined by ATP assay kit (Beyotime Biotechnology, China) and NAD/NADH Quantitation Kit (Solarbio biotechnology Ltd., China), respectively. The assays were performed according to the manufacturer's instruction. All chemicals used were of analytical grade and solvents for chromatography of HPLC-grade.

2. Results and discussion

2.1. Spectra pretreatment

In order to remove the first and second-order Rayleigh scattering, the widths of the two scattering bands were assessed for the dataset to be ± 25 and ± 20 nm, respectively. Then, the removed areas were replaced with shape-preserving piecewise cubic interpolation. Finally, the 3D-FFF data $[52 \times 126 \times 35]$ was unfolded to be a 2D $[52 \times 4410]$ data matrix for ICA.

2.2. Independent component analysis

The method "ICA-by-Blocks" with from 1 to 20 ICs and 2 block numbers was applied to the unfolded matrix (Fig. 1) to determine the number of ICs. The lowest correlations between source signals extracted from two blocks were calculated for every possible IC number, and the largest IC number with the relatively high correlation value would be the optimal IC number. Since the results may be influenced by the distribution of the measured data in the two blocks, the ICA-by-Blocks calculation was repeated 100 times with random distribution of the samples between the two blocks, and the chance of every IC number becoming the optimal one out of 100 repetitions was calculated. According to these calculations, 6 was the most probable IC number for this dataset (Fig. 2).

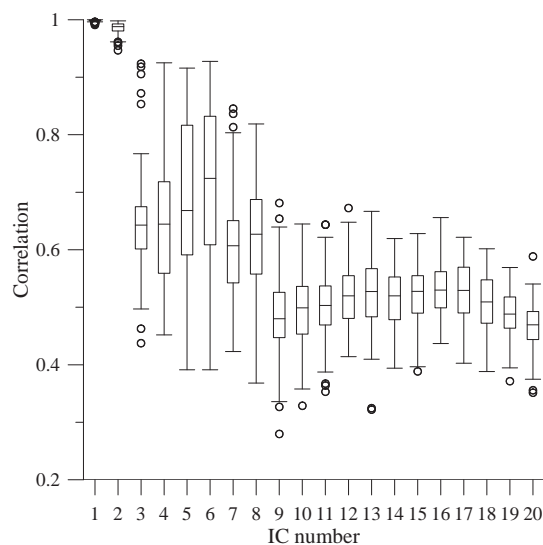


Fig. 2. The boxplot of the correlation values calculated by ICA-by-Blocks after 100 repetitions with the samples randomly distributed between two blocks for every possible IC number. Small circles are outliers

ICA with 6 ICs was applied to the unfolded matrix (52×4410), and the 2D matrices of decomposed independent signals were refolded to get EEM landscape data (Fig. 1). Based on the data, the average proportion value of each garlic cultivar on 6 ICs is calculated as seen in Figure 3. The result indicated that the crucial differences between the eight cultivars mainly appears in the first, second, and fourth IC (IC1, IC2, and IC4). The proportion value of the eight garlic cultivars showed the biggest difference on IC1, and the garlic cultivar YNQJ had the highest score, while the SXXA the lowest. Due to the similarity of scores on all 6 ICs, “NXYC” and “SDHB”, “SDHZ” and “SDLC” both seemed to have a close relationship.

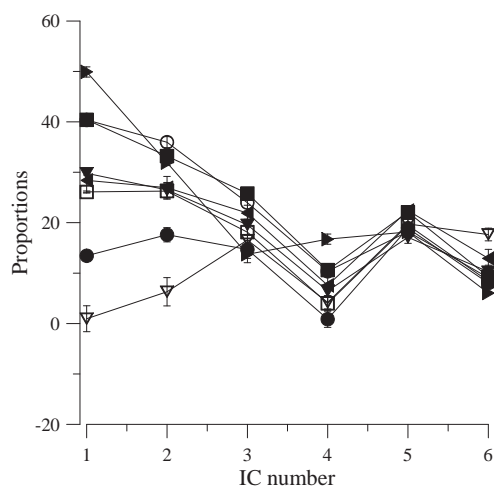


Fig. 3. The average proportions of six ICs of eight garlic cultivars (▽: SDLC; ■: SDHB; ●: HNJZ; ◻: SDHZ; ○: NXJC; ▽: SXXA; ▶: YNQJ; ◀: ZJHZ)

2.3. Chemical fluorophores behind decomposed ICs

The contents of possible fluorescent substances, proteins (total protein), vitamins (A, B2, B6, E), ATP, and coenzymes (NADH) were determined, and the results are listed in Table 1. Vitamin A and E were not detected in any garlic cultivar. “YNQJ” had the highest contents of protein and vitamin B6, while cultivars “NXJC”, “SDHB”, and “SXXA” showed good results at ATP, NADH, and vitamin B2, respectively.

Based on the calculated Pearson correlation coefficients listed in Table 2, it can be stated that the concentration of protein is in strong relation with the proportion values of IC1, and it indicates that protein, more specifically tryptophan, is most likely the fluorophore behind this IC1 signal, which is in agreement with some fluorophores data reported before (GHISAIDOOBE & CHUNG, 2014). Meanwhile, the concentration of VB6 has a strong linear relation with proportions of IC4, vitamin B2 with IC6, and ATP with IC2, indicating that the fluorophores behind IC2, IC4, and IC6 are likely to be ATP, vitamin B6, and vitamin B2, respectively. However, the concentration of protein also has a strong linear relation with IC4, VB6 with IC1, and ATP with IC1, indicating that the concentration of protein, ATP, and vitamin B6 show a similar trend among different garlic cultivars. That means that garlic with higher protein content also has more ATP and vitamin B6, and vice versa.

Table 1. The concentration of possible fluorophores in 8 garlic cultivars

| Cultivar | Protein (mg ml ⁻¹) | Vitamin B2 (µg ml ⁻¹) | Vitamin B6 (µg ml ⁻¹) | ATP (µg ml ⁻¹) | NADH (µg ml ⁻¹) |
|----------|-----------------------------------|--------------------------------------|--------------------------------------|-------------------------------|--------------------------------|
| HNJZ | 2.93±0.68 | 0.87±0.10 | 0.33±0.11 | 17.60±3.83 | 14.68±6.57 |
| NXYC | 9.8±0.96 | 0.91±0.04 | 1.45±0.15 | 35.96±2.77 | 23.97±3.6 |
| SDHB | 8.81±0.96 | 0.73±0.05 | 1.48±0.15 | 33.22±2.86 | 25.79±2.72 |
| SDHZ | 4.82±1.14 | 0.73±0.09 | 0.74±0.18 | 19.64±2.14 | 18.12±6.3 |
| SDLC | 6.49±1.01 | 1.03±0.09 | 1.07±0.16 | 26.49±3.62 | 19.58±5.62 |
| SXXA | 0.21±1.54 | 1.77±0.25 | 0 | 6.29±4.46 | 16.53±5.1 |
| YNQJ | 10.88±0.32 | 0.80±0.06 | 1.59±0.05 | 26.89±1.11 | 13.72±5.05 |
| ZJHZ | 6.19±1.59 | 1.29±0.12 | 1.17±0.25 | 26.94±3.58 | 21.91±7.06 |

The concentrations of possible fluorophores in garlic were determined by using all 52 garlic samples of 8 cultivars. The values were listed in the table as the mean of samples from each cultivar ± standard deviation.

Table 2. Pearson product-moment correlation coefficients between the content of fluorophores and the proportion value of each decomposed IC

| Fluorophore | IC1 | IC2 | IC3 | IC4 | IC5 | IC6 |
|-------------|----------|----------|----------|----------|----------|----------|
| Protein | 0.9867** | 0.7793** | 0.0914 | 0.9277** | 0.4314** | 0.5351** |
| Vitamin B2 | 0.6890** | 0.5351** | 0.3717* | 0.5889** | 0.6095** | 0.9345** |
| Vitamin B6 | 0.9611** | 0.8563** | 0.0544 | 0.9784** | 0.3040* | 0.4497** |
| ATP | 0.8751** | 0.9367** | 0.1568 | 0.9137** | 0.2655 | 0.4481** |
| NADH | 0.1979 | 0.4832** | 0.8322** | 0.3398* | 0.5163** | 0.3153* |

*: P<0.05; **: P<0.01

The concentration of NADH did not show strong relation with any of the decomposed fluorescence signals, probably due to the bound form of NADH in the mitochondria, which gives about 80% fluorescence, and the much lower concentration in the solution leads to uncertainty of content determination (BLINOVA et al., 2005). However, from the results in Table 2, NADH may possibly be related to the signal of IC3.

2.4. Evaluation of garlic cultivars

Due to the big divergence of proportion value among all 8 garlic cultivars appearing at IC1, IC2, and IC4 (Fig. 3), the fluorescence signals behind these three components (protein, ATP, vitamin B6) were chosen as fluorescence markers for evaluating garlic cultivars. As shown in Figure 4, the proportion values behind protein, ATP, and vitamin B6 was used to construct a relation plot. The results show that all 8 garlic cultivars clustered separately and could be divided into 5 groups. The first group contains one cultivar “YNQJ”, of which the concentration of protein and vitamin B6 are the highest compared with the other cultivars. This result is a little surprising, because this cultivar comes from the most southern province of China, from a typical subtropical monsoon climate. Under such circumstances, the life cycle of a plant is rather shorter than of others, and also the quality normally is a little poorer. However, these results show that the good breed of the garlic “YNQJ” has the predominant role in high

quality, rather than growing environment. The second group includes two cultivars, “NXYC” and “SDHB”. These two cultivars have almost the same content of protein and vitamin B6, and the highest content of ATP compared with others, but a slight difference in ATP concentration. The third group contains three cultivars, “SDHZ”, “SDLC”, and “ZJHZ”. The concentrations of all three fluorophores in these three cultivars are similar and tangled together. The last two groups both contained one cultivar, “HNJZ” and “SXXA”, respectively. The content of vitamin B6 is the lowest in cultivar “HNJZ”, while cultivar “SXXA” has the poorest quality in terms of protein and ATP, and also exhibits big individual variance, probably because of bad storage conditions. At last, an interesting phenomenon can be seen in Fig. 4A, that the cultivar, which has a higher content of protein, may also have a relatively higher content of ATP and vitamin B6. This feature not only makes these three fluorophores good markers for garlic quality evaluation, but triggers some relevant thought about the possible metabolic connection of these three fluorescent substances.

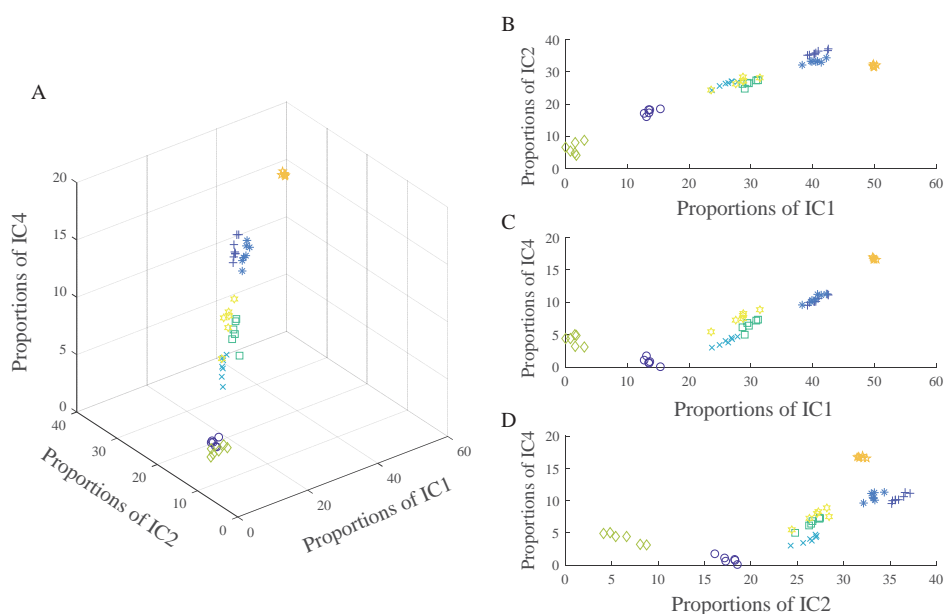


Fig. 4. Relation of proportion values between three responsible fluorophores (protein behind IC1, ATP behind IC2, and vitamin B6 behind IC4) of eight garlic cultivars (+: NXYC; *: SDHB; □: SDLC; ×: SDHZ; ○: HNJZ; ◇: SXXA; ★: YNQJ; ♠: ZJHZ). (A) 3D relation plot of IC1, IC2, and IC4; (B) 2D relation plot of IC1 and IC2; (C) 2D relation plot of IC1 and IC4; (D) 2D relation plot of IC2 and IC4

3. Conclusions

For evaluating a total of 8 garlic cultivars coming from different geographic locations in China, the application of 3D front-face fluorescence spectroscopy combined with independent component analysis was presented in this study. The 3D front-face fluorescence technique can get a set of intact fluorophore data of garlic, and independent component analysis is a powerful tool for the extraction of independent signals from mixed source data. In this study, six independent components were extracted, and the chemical fluorophore behind each

component also has been unveiled. Among them, the fluorophores, protein behind the first independent component, ATP behind the second, and vitamin B6 behind the fourth, were selected as fluorescence markers for evaluating garlic cultivars. Finally, rapid evaluation of garlic cultivars was accomplished by comparing proportion values between the three fluorophores behind the first, second, and fourth independent components.

References

- BAHRAM, M., BRO, R., STEDMON, C. & AFKHAM, A. (2006): Handling of Rayleigh and Raman scatter for PARAFAC modeling of fluorescence data using interpolation. *J. Chemometr.*, *20*, 99–105.
- BLINOVA, K., CARROLL, S., BOSE, S., SMIRNOV, A.V., HARVEY, J.J., ... & BALABAN, R.S. (2005): Distribution of mitochondrial NADH fluorescence lifetimes: Steady-state kinetics of matrix NADH interactions. *Biochemistry*, *44*, 2585–2594.
- CARDOSO, J.F. & SOULOUMIAC, A. (1993): Blind beamforming for non-Gaussian signals. *IEE Proc. F.*, *140*, 362–370.
- CHRISTENSEN, J., NØRGAARD, L., BRO, R. & ENGELSEN, S.B. (2006): Multivariate autofluorescence of intact food systems. *Chem. Rev.*, *106*, 1979–1994.
- COMON, P. (1994): Independent component analysis, a new concept? *Signal Process.*, *36*, 287–314.
- GHISAIDOUBE, A.B.T. & CHUNG, S.J. (2014): Intrinsic tryptophan fluorescence in the detection and analysis of proteins: A focus on Förster resonance energy transfer techniques. *Int. J. Mol. Sci.*, *15*, 22518–22538.
- HAHN, S. & YOON, G. (2006): Identification of pure component spectra by independent component analysis in glucose prediction based on mid-infrared spectroscopy. *Appl. Optics*, *45*, 8374–8380.
- JIANG, J.K., WU, J. & LIU, X.H. (2010): Fluorescence properties of lake water. *Guang Pu Xue Yu Guang Pu Fen Xi*, *30*, 1525–1529. (In Chinese)
- JIN, P., XIA, L., LI, Z., CHE, N., ZOU, D. & HU, X. (2012): Rapid determination of thiamine, riboflavin, niacinamide, pantothenic acid, pyridoxine, folic acid and ascorbic acid in vitamins with minerals tablets by high-performance liquid chromatography with diode array detector. *J. Pharm. Biomed. Anal.*, *70*, 151–157.
- JOUAN-RIMBAUD BOUVERESSE, D., MOYA-GONZÁLEZ, A., AMMARI, F. & RUTLEDGE, D.N. (2012): Two novel methods for the determination of the number of components in independent components analysis models. *Chemometr. Intell. Lab.*, *112*, 24–32.
- RUTLEDGE, D.N. & JOUAN-RIMBAUD BOUVERESSE, D. (2013): Independent components analysis with the JADE algorithm. *TrAC, Trend. Anal. Chem.*, *50*, 22–32.
- SANTHOSHA, S.G., JAMUNA, P. & PRABHAVATHI, S.N. (2013): Bioactive components of garlic and their physiological role in health maintenance: A review. *Food Biosci.*, *3*, 59–74.
- WATERER, D. & SCHMITZ, D. (1994): Influence of variety and cultural practices on garlic yields in Saskatchewan. *Can. J. Plant Sci.*, *74*, 611–614.
- WESTAD, F. & KERMIT, M. (2009): Independent component analysis. -in: BROWN, S.D., TAULER, R. & WALCZAK, B. (Eds) *Comprehensive chemometrics*. Elsevier, Oxford, pp. 227–248.
- WU, K., LIU, J.J., LI, X. & ZOU, L. (2016): Determination of vitamin A, E in foods for infants and young children, milk and milk products. *Farm Products Processing*, *9*, 41–42. (in Chinese).
- YANG, L. & HUIYAN, Z. (1996): Shape preserving piecewise cubic interpolation. *Applied Mathematics*, *11*, 419–424.