

Cite this: *RSC Adv.*, 2019, 9, 17754

# A new methodology for sensory quality assessment of garlic based on metabolomics and an artificial neural network†

Jian Liu,<sup>‡ab</sup> Lixia Liu,<sup>‡a</sup> Wei Guo,<sup>a</sup> Minglang Fu,<sup>a</sup> Minli Yang,<sup>a</sup> Shengxiong Huang,<sup>b</sup> Feng Zhang<sup>id</sup><sup>\*a</sup> and Yongsheng Liu<sup>\*b</sup>

This study has established a new method for the sensory quality determination of garlic and garlic products on the basis of metabolomics and an artificial neural network. A total of 89 quality indicators were obtained, mainly through the metabolomics analysis using gas chromatography/mass spectrometry (GC/MS) and high performance liquid chromatography coupled with tandem mass spectrometry (HPLC-MS/MS). The quality indicator data were standardized and fused at a low level, and then seven representative indicators including the  $a^*$  (redness) value, and the contents of *S*-methyl-L-cysteine, 3-vinyl-1,2-dithiacyclohex-5-ene, glutamic acid, L-tyrosine, D-fructose and propene were screened by partial least squares discriminant analysis (PLS-DA), analysis of variance (ANOVA) and correlation analysis (CA). Subsequently, the seven representative indicators were employed as the input data, while the sensory scores for the garlic obtained by a traditional sensory evaluation were regarded as the output data. A back propagation artificial neural network (BPANN) model was constructed for predicting the sensory quality of garlic from four different areas in China. The  $R^2$  value of the linear regression equation between the predicted scores and the traditional sensory scores for the garlic was 0.9866, with a mean square error of 0.0034, indicating that the fitting degree was high and that the BPANN model built in this study could predict the sensory quality of garlic accurately. In general, the method developed in this study for the sensory quality determination of garlic and garlic products is rapid, simple and efficient, and can be considered as a potential method for application in quality control in the food industry.

Received 14th March 2019

Accepted 6th May 2019

DOI: 10.1039/c9ra01978b

rsc.li/rsc-advances

## Introduction

Garlic (*Allium sativum* L.) is one of the most frequently consumed vegetables belonging to the Liliaceae family. Because of its special flavor and taste, garlic has been enjoyed as a food and condiment by people in almost all parts of the world for a long time.<sup>1,2</sup> Moreover, garlic is characterized by its enormous nutritional value and biological activities, including its effect in reducing blood lipid, resisting bacteria, eliminating inflammation, and preventing cardiovascular diseases.<sup>3–5</sup>

Recently, various garlic products have become available, such as garlic paste, garlic slices, salted garlic, picked garlic and fermented garlic, so as to suit the growing demands of consumers. As one of the important aspects of garlic quality, the

sensory quality of garlic and garlic products is easily influenced by the cultivars, geographical origin, and processing methods of the garlic.<sup>6,7</sup> In addition, researchers have proved that the sensory quality of garlic and garlic products is not only associated with its content of organosulfur compounds, but also closely related to the contents of sugars, amino acids, and phenolics. For example, Pardo *et al.*<sup>7</sup> analyzed relationships between the physicochemical and sensory parameters of 14 garlic varieties, and showed that the total sugar content was positively connected with the color parameters  $a^*$  and  $h^*$  (indicating red-green component and hue angle, respectively). Zhang *et al.*<sup>8</sup> found that the temperature had a significant influence on the basic indicators of sensory quality of black garlic, including the browning intensity, reducing sugar, total acids and alliin. Li *et al.*<sup>9</sup> identified several phenolic compounds in garlic which could inhibit the decomposition of alliin and alliinase. Alliin can be produced when alliin reacts with alliinase, and alliin is easily decomposed into a series of S-containing volatiles, which are the main source of the flavor of garlic. However, the literature on garlic sensory quality reported so far does not fully reflect the sensory quality of garlic and garlic products by considering both physical and chemical traits. Moreover, the relationship between physical and

<sup>a</sup>Institute of Food Safety, Chinese Academy of Inspection & Quarantine, Beijing 100176, China. E-mail: fengzhang@126.com; fengzhangchem@yahoo.com; Fax: +86-10-53898008; Tel: +86 13651290763

<sup>b</sup>School of Food Science and Engineering, Hefei University of Technology, Hefei 230009, China

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c9ra01978b

‡ These authors contributed equally to this work.



chemical traits and sensory quality needs a further study, which may promote the sensory evaluation of garlic and the development of garlic breeding for special flavor.

Currently, the most commonly used method for evaluating the sensory quality of garlic and garlic products depends on a descriptive sensory analysis and uses a trained panel,<sup>10–12</sup> but this has the disadvantages of being time-consuming and easily affected by subjective factors. Although some new instruments have already been used in sensory analysis of food, such as an electronic tongue and nose, they are still too expensive to be routinely used for quality control.

Since metabolomics offers a powerful means of monitoring all component concentrations that may be related to the sensory quality of a food sample,<sup>13</sup> it has been widely applied in the food industry and it makes the sensory quality assessment of garlic more accurate and comprehensive. Besides, a combination of multiple instrumental techniques can be more effective in discovering broader types of metabolites, and has been widely applied in the field of food.<sup>14,15</sup> In particular, multivariate statistical analysis is used to screen the most important quality indicators of a food, and includes analysis of variance (ANOVA), principal component analysis (PCA) and correlation analysis (CA).<sup>16,17</sup> These screened sensory quality indicators offer more accurate information about food quality and simplify the progression of the food quality assessment. However, the sensory evaluation of garlic based on metabolomics has not been reported.

The back propagation artificial neural network (BPANN) model is an artificial intelligence information processing system and is recognized as one of the mostly extensively used artificial neural network models. Since the BPANN model mainly imitates the human brain in processing complicated issues, a BPANN model generally includes three layers: an input layer, a hidden layer, and an output layer.<sup>18</sup> Central to the construction of a BPANN model is the continual training of the input data and the target output data until the accuracy of the predicted results is satisfactory, which is very convenient and does not require any mathematical formula, or weighting of input data.<sup>19</sup> Therefore, a BPANN model not only has the capability to solve linear problems, but can also be applied to the handling of nonlinear problems. Nevertheless, the BPANN model does have several drawbacks: for instance, the model-related parameters (which are the sensitivity factors for applying the BPANN model) are difficult for learners to confirm, and non-ideal reproducibility is often encountered. However, a number of investigations have reported satisfactory results for food quality assessment using a BPANN model. For example, Wang *et al.*<sup>20</sup> established a BPANN model based on the data of hyperspectral imaging and successfully discriminated rice variety and quality. Lu *et al.*<sup>21</sup> developed an efficient approach based on a grade classification model and a BPANN model, and the accuracy of this new approach for predicting the eating quality of rice reached 90%. In order to accurately assess the storage quality of fresh-cut green peppers, Meng *et al.*<sup>22</sup> built a BPANN model using oxygen, carbon dioxide, temperature and humidity as the input data, and *b* value, water loss, firmness

and vitamin C content as the output data, and the established BPANN model gave good predicted results.

The present study aims to establish a new methodology for sensory quality assessment of garlic and garlic products based on metabolomics and an artificial neural network. Firstly, the whole chemical constituents of garlic were acquired by a non-targeted metabolomics method, mainly using gas chromatography/mass spectrometry (GC/MS) and high performance liquid chromatography coupled with tandem mass spectrometry (HPLC-MS/MS) for detecting different classes of compounds. Secondly, the obtained data on the garlic constituents was combined through low-level fusion, and filtered by partial least squares discriminant analysis (PLS-DA), ANOVA and CA, in order to identify representative sensory quality indicators for garlic. Thirdly, a BPANN model was established for predicting the sensory quality of garlic; the representative sensory quality indicators of garlic were used as the input data of the BPANN model, and the sensory scores of garlic obtained by traditional sensory evaluation were regarded as the output data. A correlation coefficient between the predicted scores (from the developed BPANN model) and the sensory scores of garlic (obtained by traditional sensory evaluation) using a linear regression model was adopted to validate the BPANN model. This method is valuable not only for evaluating the quality of garlic and garlic products comprehensively and objectively, so determining the quality characteristics of different garlic cultivars, but also for developing garlic products and promoting the development of the garlic industry.

## Material and methods

### Garlic samples

A total of twenty-eight garlic samples (*Allium sativum* L.) were collected from the four main garlic-producing areas in China: Shandong (7 samples, SD1–SD7), Henan (7 samples, HN1–HN7), Jiangsu (7 samples, JS1–JS7) and Yunnan (7 samples, YN1–YN7). The detailed information for the samples is listed in Table S1.† The varieties were authenticated by Prof. Minli Yang at the Institute of Food Safety, Chinese Academy of Inspection & Quarantine (China). All of the garlic samples were harvested in June 2018, collected directly from the producers, and then stored at 4–8 °C until analysis.

### Chemicals

Methanol (MeOH), chloroform (CHCl<sub>3</sub>), *n*-hexane, and acetonitrile were obtained from Merck (Darmstadt, Germany). Methoxyamine hydrochloride, *N*-methyl-*N*-(trimethylsilyl) trifluoroacetamide (MSTFA), pyridine, normal alkane standards (C7–C30), isopentanol, ribitol, and standards of phenolic compounds (including apigenin, isorhamnetin, naringenin, hyperoside, luteolin, quercetin, quercitrin, myricetin, caffeic acid, ferulic acid, chlorogenic acid, *p*-coumaric acid, protocatechuic acid, resveratrol, and phthalic acid) were supplied by Sigma-Aldrich (St. Louis, Missouri, USA).



### Analysis of volatile metabolites

The method adopted for determination of volatile metabolites was modified on the basis of previous studies.<sup>23,24</sup> Garlic (2 g) was weighed and chopped, and then placed into a 20 mL glass vial. Isopentanol (5  $\mu$ L) was added into the above glass vial as the internal standard, and then sample preparation was carried out immediately on a CTC CombiPal autosampler (Zwingen, Switzerland). A solid-phase microextraction (SPME) coated fiber (50/30  $\mu$ m, DVB/CAR/PDMS) was selected based on the previous literature.<sup>25</sup> The sample was incubated (40  $^{\circ}$ C, 10 min) and the extraction of volatiles was then performed with continuous stirring (40  $^{\circ}$ C, 40 min). The volatiles were determined utilizing an Agilent 7890A gas chromatograph connected to an Agilent MS detector (5975C inert XL MSD with Triple-Axis Detector). After the extraction, the fiber was inserted into the GC inlet for analysis of volatiles (250  $^{\circ}$ C, 5 min, split ratio of 5 : 1). The volatiles were separated using a HP-Innowax column (60 m  $\times$  0.25 mm  $\times$  0.25  $\mu$ m, Agilent Technologies, USA), using helium as the carrier gas (1.0 mL min<sup>-1</sup>). The column temperature program was as follows: 40  $^{\circ}$ C (0–3 min), 40–240  $^{\circ}$ C (3–43 min), and 240  $^{\circ}$ C (43–48 min). The MS quadrupole and the transfer line were both operated at the same temperature of 150  $^{\circ}$ C. The operating temperature for MS and the mass range studied were 230  $^{\circ}$ C and 40–500  $m/z$ , respectively.<sup>26</sup> The measurements of volatile metabolites were repeated six times for each sample. The identification of volatile metabolites was mainly accomplished according to the NIST library and assisted by the retention indices, which were evaluated using C7–C30 alkanes.<sup>27</sup> Moreover, the content of volatile metabolites was determined relative to the content of isopentanol.

### Analysis of primary metabolites

Based on the method of Lisec *et al.*,<sup>28</sup> the procedure for testing primary metabolites was as follows: 100 mg lyophilized sample was weighed into a 2 mL Eppendorf tube, then mixed with 1.4 mL of 60% aqueous methanol and 60  $\mu$ L ribitol in distilled water (0.2 mg mL<sup>-1</sup>). A sonicator (30 Hz, 10 min) and a thermomixer (950 rpm, 70  $^{\circ}$ C, and 10 min) were used for thorough mixing. Then, the sample was centrifuged at 11 000g for 15 min at 4  $^{\circ}$ C, after which the supernatant was extracted with 0.75 mL chloroform and 1.5 mL distilled water, and vortexed for 10 s. Subsequently, the centrifugation step was repeated, and a polytetrafluoroethylene (PTFE) filter (0.2  $\mu$ m) was used for further purification of the supernatant, which was then dried for oximation. The oximation of the dried sample was performed with 40  $\mu$ L of methoxyamine hydrochloride in pyridine (20 mg mL<sup>-1</sup>), with a reaction temperature of 37  $^{\circ}$ C and a reaction time of 90 min. Finally, MSTFA (70  $\mu$ L) was added into the above solution and the temperature was held at 37  $^{\circ}$ C for another 30 min to finish the silylation. Another Agilent 7890A GC system equipped with a MS detector (5975C inert XL MSD with Triple-Axis Detector, Agilent) was applied for the analysis of primary metabolites. The sample (1  $\mu$ L) was inserted into the GC inlet (250  $^{\circ}$ C, 5 min, split ratio of 5 : 1). A DB-5MS capillary column (30 m  $\times$  0.25 mm  $\times$  0.25, Agilent Technologies, USA) was employed using helium as the carrier gas (1.5 mL min<sup>-1</sup>). The

column temperature program was as follows: 70  $^{\circ}$ C (0–2 min), 70–300  $^{\circ}$ C (2–23 min), and 300  $^{\circ}$ C (23–29 min). The MS quadrupole and the transfer line were both operated at the same temperature (150  $^{\circ}$ C). The operating temperature for MS and the mass range studied were 230  $^{\circ}$ C and 40–600  $m/z$ , respectively. The measurements of primary metabolites were repeated six times for each sample. Identification of primary metabolites was performed in the same way as that for volatile metabolites, by reference to the NIST library and the retention indices. The concentration of primary metabolites was determined relative to the content of ribitol.

### Analysis of secondary metabolites

A lyophilized sample (100 mg) was weighed into a 2 mL Eppendorf tube, and then extracted using 1.5 mL of 60% aqueous methanol. The mixture was put into a sonicator (30 Hz, 10 min) for sufficient homogenization. The next step was to centrifuge the sample at 11 000g for 15 min at 4  $^{\circ}$ C, after which a PTFE filter (0.2  $\mu$ m) was used for further purification of the supernatant. A Symbiosis Pharma HPLC linked with a triple quadrupole mass analyzer (API 6500QTRAP, Applied Biosystems/MDS Sciex) was used for the secondary metabolite analysis. The sample (10  $\mu$ L) was injected into a ZORBAX SB-C18 column (150 mm  $\times$  4.6 mm, 5  $\mu$ m particle size, Agilent, Santa Clara, USA) for separation, and a gradient system (0.8 mL min<sup>-1</sup>) was adopted. Formic acid aqueous solution (0.1%) (A) and acetonitrile (B) constituted the mobile phase. The gradient program was set as follows: 0–2 min, 90% A; 2–7 min, 90% to 40% A; 7–15 min, 40% to 20% A; 15–16 min, 20% to 90% A; and 16–17 min, 90% A. The spray voltage in negative mode was set at 4500 V. The other parameters were set as follows: nebulizer gas, 55 psi; curtain gas 40 psi; and auxiliary heating gas, 55 psi. Tandem MS analysis of the secondary metabolites was performed in multiple reaction monitoring mode. The evaluation of secondary metabolites was repeated six times for each sample. Identification of secondary metabolites was based on the MS information provided by authentic standards, and external standard curves of the authentic standards were used to gain quantitative information about the target metabolites.

### Color measurements

The measurement of garlic color was based on the method of Dan *et al.*<sup>29</sup> Each sample (50 g) was weighed and extracted using 50 mL aqueous methanol. After adequate homogenization, the mixture was centrifuged (10 000g, 4  $^{\circ}$ C, 10 min) to obtain a clear supernatant that was then tested by a colorimeter (NH300, Beijing Hongruida Technology Co., Ltd, Beijing, China). The color results of the garlic sample were represented with values of L\*, a\* and b\*, which represent lightness, redness and yellowness, respectively. Every sample was assessed and recorded three times.

### Determination of hardness

The hardness of the peeled garlic was determined using an electronic hardness tester (GY-4, Beijing Hongruida Technology Co., Ltd, Beijing, China). The peeled garlic sample was placed



on the plate of the tester, and a cylinder with a diameter of 3.5 mm was slowly inserted into the garlic to a depth of 10 mm. The electronic data reading of the tester was recorded and the mean value of three measurements was used to represent the hardness of each garlic sample.

### Other quality indicators

Moisture content was measured according to the Chinese national standard (GB5009.3-2016).<sup>30</sup> Reducing sugar content was determined by reference to the method of Zhang *et al.*<sup>31</sup> The content determination of allicin was performed according to the method of Ramirez *et al.*<sup>32</sup> Four sulfur-amino acids, namely (+)-*S*-allyl-L-cysteine (SAC), (+)-*S*-methyl-L-cysteine (SMC), (+)-*S*-allyl-L-cysteine sulfoxide (alliin) and (+)-*S*-methyl-L-cysteine sulfoxide (methiin) were evaluated on the basis of Kim *et al.*<sup>33</sup> In addition, the ascorbic acid content was assessed according to the method of Montañó *et al.*<sup>34</sup> All experiments were carried out in triplicate.

### Traditional sensory evaluation

A traditional sensory evaluation was performed using a method adapted from Horita *et al.*<sup>35</sup> and Constenla and Lozano.<sup>36</sup> Twelve educated panelists carried out the sensory evaluation and the overall scores were evaluated according to four characteristics (texture, taste, appearance and flavor) of the garlic samples, as presented in Table S2.† A 9-point hedonic scale was used to assess the overall sensory score of each garlic sample<sup>37</sup> and the sensory evaluation of every garlic sample was repeated three times.

### Statistical analysis

The data obtained after normalization of all the indicators was entered into the software SIMCA-P+ (Version 13.0) for PLS-DA, and the variable influence on projection (VIP) of the PLS-DA was used to perform a preliminary screen of quality indicators. Subsequently, ANOVA and CA were executed by SPSS software (version 20, IBM Inc. USA), to further select representative sensory quality indicators of the garlic samples. Finally, the BPANN model for predicting the sensory quality of garlic was established by Matlab (version R2016a, The Mathworks Inc., MA, USA).

## Result and discussion

### Traditional sensory evaluation

Sensory evaluation is one of the essential techniques used in the food area, and provides a reasonable assessment of the target food by evaluation experts based on their olfactory senses, taste senses and visual senses. According to the result of the traditional sensory evaluation presented in Table 1, the characteristics (texture, taste, appearance, flavor and overall impression) of 28 garlic samples from different provinces showed remarkable geographical differences. Specifically, garlic samples from Yunnan gained lower texture scores due to their hard texture, possibly caused by the drought climate of Yunnan which leads to the low content of water in the garlic. Besides, garlic samples

from Yunnan showed the strongest spicy taste and garlic flavor. Compared with garlic samples from Yunnan, Jiangsu and Henan, the taste and flavor of garlic samples from Shandong had the lowest scores, suggesting that garlic samples from Shandong have the weakest spicy taste and garlic flavor.

### Data fusion and preliminary screening of quality indicators

This research firstly investigated the physical and chemical properties of garlics sourced from four different areas, mainly based on metabolomics, to provide a data basis for the screening of sensory quality indicators. After data pre-processing, 76 metabolites were tentatively identified, including 25 volatile metabolites, 39 primary metabolites and 12 secondary metabolites. In the meantime, we also measured four sulfur-amino acids, allicin, ascorbic acid, reducing sugar, hardness, L\* value, a\* value, b\* value and moisture content of the garlic samples.

The next step was to combine the multiple data sets obtained from the different instrumental sources, to realize a complementarity between instruments and produce a more comprehensive and accurate sensory quality assessment of the garlic. Before the data fusion, the initial data were normalized to eliminate dimensional disturbances. Thus, zero-mean normalization of initial data was adapted in this study.<sup>38</sup> In order to distinguish the garlic samples and investigate the differences in the sensory quality indicators, the normalized data was entered into the software SIMCA-P+ to perform PLS-DA. A PLS-DA model with 3 components was established (Fig. 1a) and the model quality was evaluated, with  $R^2(X) = 0.583$ ,  $R^2(Y) = 0.969$ , and  $Q^2 = 0.948$ . A permutation test ( $n = 200$ ) was performed to validate the PLS-DA model and the results are shown in Fig. 1b, indicating that the PLS-DA model based on 89 sensory quality indicators was not overfitted and was statistically acceptable. Each dot in Fig. 1a represents a garlic sample and the corresponding garlic name is marked nearby. The garlic samples that are placed together have similar levels of quality indicators and the garlic samples that are well separated have extremely different levels of quality indicators. From the PLS-DA score, it can obviously be seen that the garlic samples collected from the same province are closely placed, which implies that they have similar levels of sensory quality indicators. In particular, the garlic samples produced in Shandong and Yunnan are separated significantly from those from Jiangsu and Henan, suggesting that the sensory quality of garlics in Shandong and Yunnan is quite different from that in garlics from the other two provinces. In addition, the garlic samples collected from Henan and Jiangsu are mixed together in the score plot, indicating that there is little difference in the sensory quality of garlic samples from these two places.

As one of the most frequently used criteria in variable selection methods, VIP clarifies the importance of variables in the projection and determines the variables that contribute the most to a PLS model.<sup>39</sup> Therefore, the variables with  $VIP > 1$  in the present study were thought to be important sensory quality indicators of garlics from the four





Table 1 Sensory evaluation of 28 garlic samples<sup>a</sup>

Sample	Texture	Taste	Appearance	Flavor	Overall impression
SD1	6.3 ± 0.5 <sup>abcde</sup>	6.3 ± 1.0 <sup>abcde</sup>	5.4 ± 0.9 <sup>abc</sup>	5.1 ± 0.7 <sup>abcd</sup>	5.8 ± 0.4 <sup>abcd</sup>
SD2	6.7 ± 0.6 <sup>abcde</sup>	5.6 ± 0.7 <sup>abc</sup>	5.7 ± 1.0 <sup>abc</sup>	5.0 ± 0.5 <sup>abc</sup>	5.6 ± 0.2 <sup>a</sup>
SD3	7.0 ± 1.0 <sup>abcde</sup>	5.6 ± 0.9 <sup>abc</sup>	5.7 ± 1.0 <sup>abc</sup>	4.9 ± 0.4 <sup>ab</sup>	5.6 ± 0.2 <sup>a</sup>
SD4	6.8 ± 0.7 <sup>bcde</sup>	5.5 ± 1.0 <sup>ab</sup>	5.7 ± 2.0 <sup>abc</sup>	5.3 ± 0.7 <sup>abcdef</sup>	5.7 ± 0.3 <sup>ab</sup>
SD5	7.0 ± 1.0 <sup>bcde</sup>	5.8 ± 0.8 <sup>abcd</sup>	5.1 ± 2.0 <sup>ab</sup>	5.3 ± 0.7 <sup>abcdef</sup>	5.8 ± 0.3 <sup>abc</sup>
SD6	6.8 ± 0.4 <sup>bcde</sup>	5.8 ± 0.9 <sup>abcd</sup>	6.5 ± 0.9 <sup>abcd</sup>	4.7 ± 0.7 <sup>a</sup>	5.7 ± 0.2 <sup>a</sup>
SD7	6.8 ± 0.2 <sup>bcde</sup>	6.2 ± 0.9 <sup>abcde</sup>	5.7 ± 1.0 <sup>abc</sup>	5.2 ± 0.6 <sup>abcde</sup>	5.9 ± 0.3 <sup>abcde</sup>
JS1	6.8 ± 0.7 <sup>bcde</sup>	5.4 ± 0.6 <sup>ab</sup>	5.7 ± 1.0 <sup>abc</sup>	5.4 ± 0.7 <sup>abcdefg</sup>	5.7 ± 0.2 <sup>ab</sup>
JS2	6.7 ± 0.5 <sup>abcde</sup>	5.8 ± 1.0 <sup>abcd</sup>	4.6 ± 1.0 <sup>a</sup>	5.6 ± 0.7 <sup>abcdefgh</sup>	5.8 ± 0.4 <sup>abc</sup>
JS3	7.1 ± 0.9 <sup>de</sup>	6.2 ± 1.0 <sup>abcde</sup>	7.0 ± 1.0 <sup>abcd</sup>	5.6 ± 0.4 <sup>abcdefgh</sup>	6.3 ± 0.2 <sup>efg</sup>
JS4	7.0 ± 0.8 <sup>cde</sup>	6.3 ± 0.8 <sup>abcde</sup>	6.8 ± 2.0 <sup>abcd</sup>	5.4 ± 0.5 <sup>abcdefg</sup>	6.2 ± 0.1 <sup>cdefg</sup>
JS5	6.6 ± 0.7 <sup>abcde</sup>	6.5 ± 2.0 <sup>abcde</sup>	4.9 ± 1.0 <sup>a</sup>	5.9 ± 0.8 <sup>bcdefghi</sup>	6.1 ± 0.4 <sup>bcdef</sup>
JS6	6.7 ± 0.6 <sup>abcde</sup>	5.2 ± 1.0 <sup>a</sup>	6.3 ± 2.0 <sup>abcd</sup>	5.0 ± 0.2 <sup>abcd</sup>	5.6 ± 0.2 <sup>a</sup>
JS7	6.8 ± 0.9 <sup>bcde</sup>	6.6 ± 0.8 <sup>abcde</sup>	7.1 ± 1.0 <sup>abcd</sup>	5.1 ± 0.8 <sup>abcd</sup>	6.2 ± 0.4 <sup>bcdef</sup>
HN1	6.3 ± 0.7 <sup>abcde</sup>	6.7 ± 2.0 <sup>abcde</sup>	6.5 ± 2.0 <sup>abcd</sup>	5.7 ± 0.7 <sup>abcdefgh</sup>	6.2 ± 0.1 <sup>defg</sup>
HN2	7.1 ± 0.8 <sup>e</sup>	7.5 ± 0.4 <sup>de</sup>	6.2 ± 0.5 <sup>abcd</sup>	6.2 ± 0.6 <sup>defghi</sup>	6.8 ± 0.2 <sup>ijk</sup>
HN3	6.9 ± 0.6 <sup>bcde</sup>	6.2 ± 1.0 <sup>abcde</sup>	6.2 ± 1.0 <sup>abcd</sup>	6.2 ± 1.0 <sup>cdefghi</sup>	6.3 ± 0.4 <sup>efgh</sup>
HN4	6.5 ± 0.8 <sup>abcde</sup>	7.3 ± 0.6 <sup>cde</sup>	7.6 ± 0.9 <sup>bed</sup>	6.1 ± 0.7 <sup>bcdefghi</sup>	6.7 ± 0.3 <sup>hijk</sup>
HN5	7.2 ± 0.4 <sup>e</sup>	6.2 ± 1.0 <sup>abcde</sup>	6.2 ± 0.5 <sup>abcd</sup>	6.2 ± 1.0 <sup>cdefghi</sup>	6.4 ± 0.3 <sup>fghi</sup>
HN6	7.0 ± 0.9 <sup>bcde</sup>	6.8 ± 2.0 <sup>abcde</sup>	5.9 ± 2.0 <sup>abcd</sup>	6.0 ± 0.3 <sup>bcdefghi</sup>	6.5 ± 0.1 <sup>fghi</sup>
HN7	6.5 ± 1.0 <sup>abcde</sup>	6.5 ± 0.4 <sup>abcde</sup>	6.5 ± 2.0 <sup>abcd</sup>	6.2 ± 1.0 <sup>cdefghi</sup>	6.4 ± 0.4 <sup>fgh</sup>
YN1	5.6 ± 0.3 <sup>a</sup>	7.8 ± 0.5 <sup>e</sup>	7.0 ± 1.0 <sup>abcd</sup>	5.9 ± 1.0 <sup>bcdefghi</sup>	6.6 ± 0.4 <sup>ghijk</sup>
YN2	6.1 ± 0.4 <sup>abcde</sup>	7.5 ± 1.0 <sup>de</sup>	5.9 ± 2.0 <sup>abcd</sup>	6.7 ± 0.9 <sup>hi</sup>	6.8 ± 0.4 <sup>hijk</sup>
YN3	5.6 ± 0.1 <sup>a</sup>	7.9 ± 0.4 <sup>e</sup>	6.2 ± 2.0 <sup>abcd</sup>	6.5 ± 0.9 <sup>fghi</sup>	6.8 ± 0.4 <sup>hijk</sup>
YN4	5.9 ± 0.2 <sup>abc</sup>	7.8 ± 0.9 <sup>e</sup>	6.8 ± 2.0 <sup>abcd</sup>	7.1 ± 0.6 <sup>i</sup>	7.1 ± 0.2 <sup>k</sup>
YN5	6.1 ± 0.4 <sup>abcde</sup>	7.7 ± 0.5 <sup>e</sup>	7.8 ± 1.0 <sup>cd</sup>	6.5 ± 0.5 <sup>ghi</sup>	7.0 ± 0.2 <sup>jk</sup>
YN6	5.9 ± 0.1 <sup>abcd</sup>	7.1 ± 0.5 <sup>bcde</sup>	6.5 ± 0.9 <sup>abcd</sup>	6.4 ± 0.7 <sup>efghi</sup>	6.6 ± 0.1 <sup>fghij</sup>
YN7	5.8 ± 0.2 <sup>ab</sup>	7.4 ± 0.6 <sup>de</sup>	8.1 ± 1.0 <sup>d</sup>	6.8 ± 0.8 <sup>hi</sup>	6.9 ± 0.2 <sup>jk</sup>

<sup>a</sup> The results are indicated as mean ± std. error; different lowercase letters in the same column indicate significant differences ( $p < 0.05$ ).

different origins. According to the VIP values in the PLS-DA model (Table 2), 43 sensory quality indicators were identified initially from the original 89 indicators. Simultaneously, these 43 sensory quality indicators showed significant differences by Duncan's multiple-range test ( $p < 0.05$ ), probably reflecting the regional differences in the garlic samples.

#### Data distribution of differential quality indicators.

The data distribution of the 43 differential quality indicators is shown in Table S3† and presents various degrees of difference. For example, the L\* value indicates the brightness of the garlic color and a larger L\* value means a brighter color. From Table S3,† the variation coefficient of the L\* value was the smallest, at only 1.83%, indicating that the L\* value has little contribution to the sensory quality of garlics from different areas. Except for sucrose, the variation coefficients of the other quality indicators all exceeded 10%, among which the variation coefficient of gluconic acid was the largest (93.06%), followed by those of ornithine, L-homoserine and arabinose (85.19%, 74.63% and 70.95%, respectively). Overall, the differential sensory quality indicators based on data distribution mainly included volatile compounds, amino acids and sugars. These chemical constituents provide the foundations for the sensory quality of garlic and garlic products according to the previous literature.<sup>7–9</sup> Besides, these results prove the necessity to further define the important sensory quality of garlics.

#### Correlation analysis of differential quality indicators

Our study aimed to predict garlic sensory quality using several representative indicators, so it was necessary to do further research on the screening of the 43 sensory quality indicators. CA was adapted in this study to evaluate the relationships among different sensory quality indicators and provide evidence for simplifying related indicators. Moreover, the relationships between the different sensory quality indicators were demonstrated by correlation coefficients ( $r$  values) that varied from  $-1$  to  $+1$ . A high correlation was determined between two variables when the  $r$  value was close to  $+1$  or  $-1$ , while a very weak correlation was suggested when the  $r$  value was close to  $0$ .

Different degrees of correlation of the 43 differential quality indicators are shown in Fig. 2 and Table S4.† For instance, high correlations were found between methyl propenyl disulfide, dimethyl trisulfide, diallyl disulfide and diallyl tetrasulfide ( $p < 0.05$ ), implying that one of these four compounds can be chosen to represent the others. As a consequence, we could conclude that there is a certain degree of correlation between the sensory quality indicators that causes overlapping among these indicators. Thus, the 43 sensory quality indicators could reasonably be reduced to several independent indicators according to the CA.

Colors represent the Pearson's correlation coefficient values: the red color indicates a positive ( $0 < r < 1$ ) correlation and the blue color represents a negative ( $-1 < r < 0$ ) correlation. X1–X43:



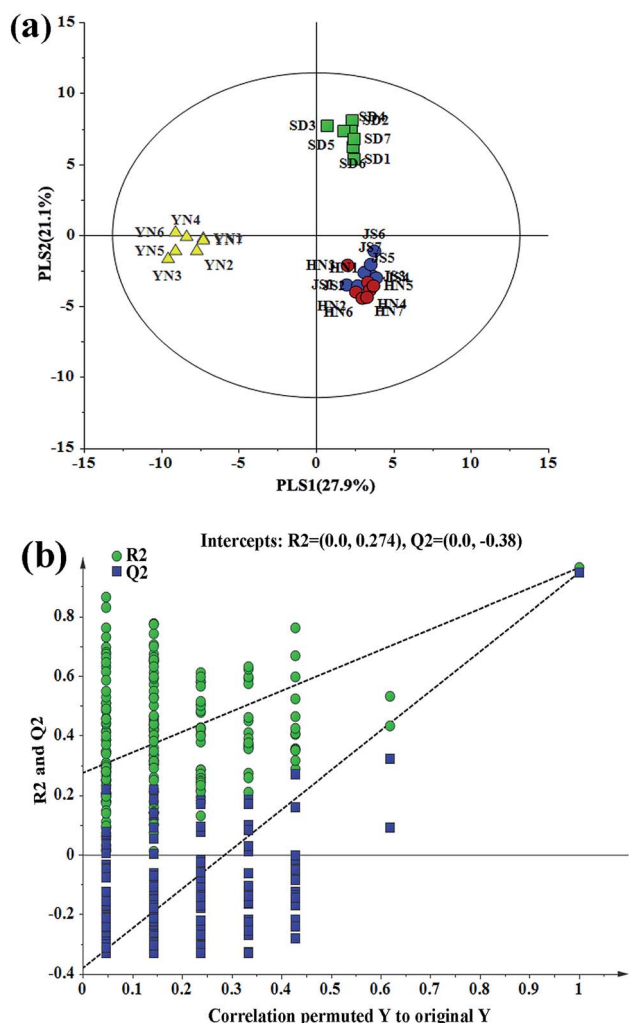


Fig. 1 (a) PLS-DA score plot and (b) permutation test plots (200 permutation tests) for garlic samples.

L\* value, a\* value, b\* value, ascorbic acid, *S*-methyl-L-cysteine, quercitrin, propene, propene sulfide, 2-propen-1-ol, di(1-propenyl) sulfide, 2-ethylthiophene, methyl propenyl disulfide, dimethyl trisulfide, diallyl disulfide, diallyl tetrasulfide, (*Z*)-benzaldehyde, 3-vinyl-1,2-dithiacyclohex-4-ene, diallyl trisulfide, 3-vinyl-1,2-dithiacyclohex-5-ene, propionic acid, 1,3-propanediol, L-alanine, ethanimidic acid, serine, L-threonine, L-proline, *n*-octanoic acid, xylulose, ornithine, glutamic acid, arabinose, L-asparagine, lanthionine, L-tyrosine, isocitric acid, D-fructose, D-glucose, L-lysine, gluconic acid, L-homoserine, myo-inositol, succinylacetone, sucrose.

Regarding the color parameters, the a\* value was highly positively associated with the b\* value and negatively related with the L\* value ( $p < 0.05$ ), suggesting that the a\* value can be chosen to represent the L\* value and b\* value. The organosulfur compounds among the 43 differential sensory quality indicators included *S*-methyl-L-cysteine, propene sulfide, di(1-propenyl) sulfide, 2-ethylthiophene, methyl propenyl disulfide, dimethyl trisulfide, diallyl disulfide, diallyl tetrasulfide, 3-vinyl-1,2-dithiacyclohex-4-ene, diallyl trisulfide, and 3-vinyl-1,2-

dithiacyclohex-5-ene. These organosulfur compounds showed different degrees of correlation with each other. In particular, *S*-methyl-L-cysteine was positively connected with di(1-propenyl) sulfide, methyl propenyl disulfide, dimethyl trisulfide, diallyl disulfide, and diallyl tetrasulfide ( $p < 0.05$ ), and negatively correlated with propene sulfide, 2-ethylthiophene and 3-vinyl-1,2-dithiacyclohex-4-ene ( $p < 0.05$ ). This might be because *S*-methyl-L-cysteine is one of the key garlic flavor precursors, and a number of volatile organosulfur compounds could be released by the action of alliinase on *S*-alk(en)yl-L-cysteine sulfoxides when garlic is crushed and disrupted.<sup>33,40</sup> Therefore, *S*-methyl-L-cysteine can be used to stand for di(1-propenyl) sulfide, methyl propenyl disulfide, dimethyl trisulfide, diallyl disulfide, diallyl tetrasulfide, propene sulfide, 2-ethylthiophene and 3-vinyl-1,2-dithiacyclohex-4-ene in our study. Moreover, a notable positive correlation was discovered between diallyl trisulfide and 3-vinyl-1,2-dithiacyclohex-5-ene (CV (Coefficient of Variance) values were 21.65% and 25.98%, respectively); herein, 3-vinyl-1,2-dithiacyclohex-5-ene was chosen to represent diallyl trisulfide.

Organic acids and amino acids are important constituents that contribute to the taste, flavor and color of garlic and garlic products. Based on Fig. 2, glutamic acid was positively associated with L-alanine, serine, L-threonine, L-asparagine and L-lysine ( $p < 0.05$ ), whereas negatively connected with lanthionine, L-homoserine, ascorbic acid, propionic acid, *n*-octanoic acid and isocitric acid ( $p < 0.05$ ), indicating that glutamic acid can be selected to represent the variations of L-alanine, serine, L-threonine, L-asparagine, L-lysine, lanthionine, L-homoserine, ascorbic acid, propionic acid, *n*-octanoic acid and isocitric acid. L-Tyrosine was positively linked to ornithine and negatively correlated with L-proline ( $p < 0.05$ ) and thus, L-tyrosine can indicate the changes in ornithine and L-proline.

Considering the sugars and other compounds in the garlic, D-fructose was positively correlated with D-glucose and sucrose, and negatively associated with myo-inositol ( $p < 0.05$ ), suggesting that D-fructose has the ability to predict the variations in D-glucose, sucrose and myo-inositol. In addition, as an odorless compound in garlic, propene displayed significant positive associations with 2-propen-1-ol and (*Z*)-benzaldehyde ( $p < 0.05$ ), and thus, propene was adopted to represent 2-propen-1-ol and (*Z*)-benzaldehyde.

In summary, CA of differential quality indicators in our study proved that the 43 sensory quality indicators can be reduced to several representative indicators. Furthermore, the association between differential sensory quality indicators and sensory attributes should be a major consideration when screening the representative sensory indicators of garlic.

### Correlation analysis between differential quality indicators and sensory attributes

It is worth noting that the sensory quality of garlic and garlic products is characterized by complex forming processes and influenced by multiple chemical compounds. The color of a garlic sample is the first sensory attribute that the consumer observes, and is usually indicated by the color parameters of the L\* value, a\* value and b\* value. It has been previously



Table 2 Selected indicators and VIP values

Indicator	VIP value	Indicator	VIP value
L* value	1.04	Ethanimidic acid	1.12
a* value	1.12	Serine	1.34
b* value	1.11	L-Threonine	1.2
Ascorbic acid	1.06	L-Proline	1.71
S-Methyl-L-cysteine	1.08	n-Octanoic acid	1.39
Quercitrin	1.61	Xylulose	1.51
Propene	1.16	Ornithine	1.09
Propene sulfide	1.15	Glutamic acid	1.24
2-Propen-1-ol	1.01	Arabinose	1.28
Di(1-propenyl) sulfide	1.12	L-Asparagine	1.24
2-Ethylthiophene	1.23	Lanthionine	1.19
Methyl propenyl disulfide	1.06	L-Tyrosine	1.22
Dimethyl trisulfide	1.15	Isocitric acid	1.51
Diallyl disulfide	1.19	D-Fructose	1.26
Diallyl tetrasulfide	1.15	D-Glucose	1.13
(Z)-Benzaldoxime	1.12	L-Lysine	1.38
3-Vinyl-1,2-dithiacyclohex-4-ene	1.3	Gluconic acid	1.6
Diallyl trisulfide	1.18	L-Homoserine	1.06
3-Vinyl-1,2-dithiacyclohex-5-ene	1.08	Myo-inositol	1.09
Propionic acid	1.16	Succinylacetone	1.04
1,3 Propanediol	1.18	Sucrose	1.06
L-Alanine	1.25		

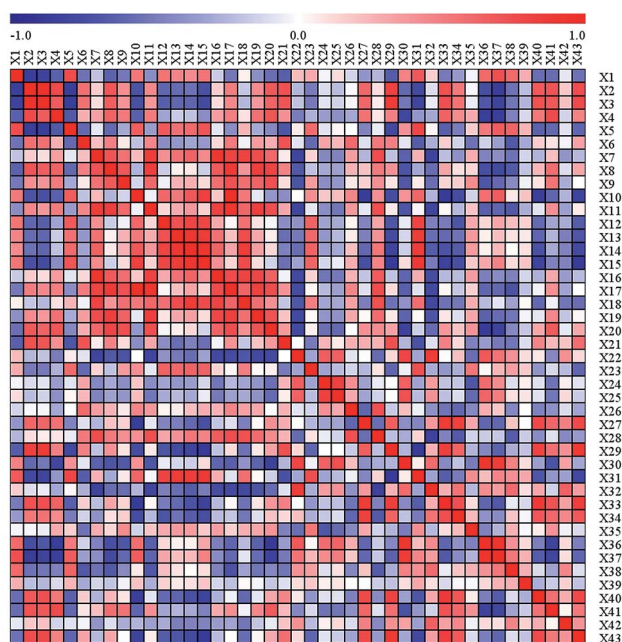


Fig. 2 Correlations between 43 quality indicators.

demonstrated that the L\* value, a\* value and b\* value can be used to calculate the change in total color of garlic puree,<sup>41</sup> while Pardo *et al.*<sup>7</sup> found that the L\* value of the external color of the garlic samples was the most useful color parameter for discriminating different garlic varieties. In our study, the L\* value, a\* value and b\* value were all different sensory quality indicators among the four garlic producing areas (VIP > 1 and  $p < 0.05$ ), and also showed strong associations with garlic texture, taste, appearance, flavor and overall impression ( $p < 0.05$ ).

Considering the high correlations between the a\* value, b\* value and L\* value ( $p < 0.05$ ), it was reasonable to select the a\* value to represent the L\* value and b\* value. Moreover, a higher a\* value indicated a relatively soft garlic texture, a more spicy taste, a yellower appearance, a stronger garlic flavor and a higher overall impression.

S-Methyl-L-cysteine is one of the garlic flavor precursors, and is closely related to the release of organosulphur compounds in garlic, which are the main source of garlic aroma.<sup>2</sup> According to our results (Table 3), a strong positive correlation was discovered between S-methyl-L-cysteine and garlic texture ( $p < 0.05$ ), suggesting that a high content of S-methyl-L-cysteine may be associated with a softer garlic texture, while significant negative associations were shown between S-methyl-L-cysteine, and garlic taste, flavor and overall impression ( $p < 0.05$ ). This may be because a higher retained S-methyl-L-cysteine content in a garlic extract means that lower levels of S-methyl-L-cysteine are converted to organosulfur compounds, leading to a weak spicy taste, a lighter garlic flavor and a lower overall impression. Volatile organosulfur compounds are recognized as the major contributors to the characteristic taste and flavor of fresh garlic. Furthermore, the different types and concentrations of volatile organosulfur compounds are the main factors in the sensory differences among garlic varieties.<sup>42,43</sup> On the basis of our results (Tables 2 and 3), a number of volatile organosulfur compounds including propene sulfide, di(1-propenyl)sulfide, methyl propenyl disulfide, dimethyl trisulfide, diallyl disulfide, diallyl tetrasulfide, 3-vinyl-1,2-dithiacyclohex-4-ene, diallyl trisulfide, and 3-vinyl-1,2-dithiacyclohex-5-ene were actually found to be differential sensory quality indicators, and exhibited various degrees of correlations with sensory attributes of garlic. In particular, propene sulfide presented notable correlations with garlic texture, taste, appearance, flavor and overall



Table 3 Correlation analysis between differential quality indicators and sensory attributes<sup>a</sup>

Indicator	Texture	Taste	Appearance	Flavor	Overall impression
L* value	0.66**	−0.66**	−0.41*	−0.51**	−0.58**
a* value	−0.83**	0.72**	0.47**	0.63**	0.65**
b* value	−0.78**	0.67**	0.38*	0.65**	0.62**
Ascorbic acid	−0.71**	0.76**	0.42*	0.73**	0.73**
S-Methyl-L-cysteine	0.73**	−0.52**	−0.26	−0.42*	−0.41*
Quercitrin	−0.36*	0.10	0.18	0.14	0.08
Propene	−0.05	0.32*	0.27	0.50**	0.46**
Propene sulfide	−0.54**	0.66**	0.46**	0.69**	0.69**
2-Propen-1-ol	−0.33*	0.65**	0.53**	0.59**	0.69**
Di(1-propenyl) sulfide	0.72**	−0.53**	−0.29	−0.48**	−0.45**
2-Ethylthiophene	−0.32*	0.46**	0.31	0.65**	0.57**
Methyl propenyl disulfide	0.53**	−0.22	0.09	−0.04	−0.04
Dimethyl trisulfide	0.51**	−0.13	0.03	0.08	0.06
Diallyl disulfide	0.55**	−0.16	0.01	0.03	0.03
Diallyl tetrasulfide	0.66**	−0.39*	−0.14	−0.25	−0.25
(Z)-Benzaldoxime	−0.02	0.36*	0.15	0.53**	0.48**
3-Vinyl-1,2-dithiacyclohex-4-ene	−0.37*	0.48**	0.40*	0.55**	0.54**
Diallyl trisulfide	0.27	0.10	0.18	0.25	0.26
3-Vinyl-1,2-dithiacyclohex-5-ene	−0.28*	0.61**	0.44**	0.62**	0.68**
Propionic acid	−0.52**	0.72**	0.60**	0.84**	0.83**
1,3-Propanediol	−0.58**	0.32	0.37*	0.29	0.28
L-Alanine	0.17	−0.52**	−0.44*	−0.72**	−0.69**
Ethanimidic acid	0.46**	−0.01	0.19	−0.04	0.10
Serine	0.07	−0.42*	−0.21	−0.62**	−0.56**
L-Threonine	0.28	−0.50**	−0.28	−0.60**	−0.57**
L-Proline	0.15	−0.23	0.00	−0.27	−0.24
n-Octanoic acid	−0.65**	0.58**	0.34*	0.53**	0.53**
Xylulose	0.07	0.06	0.21	0.21	0.18
Ornithine	−0.69**	0.46**	0.27	0.41*	0.38*
Glutamic acid	0.59**	−0.79**	−0.46**	−0.86**	−0.84**
Arabinose	0.62**	−0.31	−0.07	−0.19	−0.16
L-Asparagine	0.04	−0.30	−0.32*	−0.48**	−0.45**
Lanthionine	−0.75**	0.61**	0.39*	0.52**	0.53**
L-Tyrosine	−0.49**	0.44**	0.05	0.30	0.32*
Isocitric acid	0.02	0.42*	0.14	0.46**	0.49**
D-Fructose	0.70**	−0.88**	−0.54**	−0.91**	−0.91**
D-Glucose	0.82**	−0.85**	−0.50**	−0.86**	−0.84**
L-Lysine	0.45**	−0.39*	−0.17	−0.38*	−0.36*
Glucuronic acid	0.17	0.02	0.09	−0.17	−0.02
L-Homoserine	−0.66**	0.50**	0.23	0.45**	0.42*
Myo-inositol	−0.66**	0.76**	0.29	0.72**	0.72**
Succinylacetone	0.01	−0.02	0.07	−0.09	−0.04
Sucrose	−0.73**	0.48**	0.21	0.45**	0.39*

<sup>a</sup> \*Represents significance levels of  $p < 0.05$ . \*\*Represents significance levels of  $p < 0.01$ .

impression ( $p < 0.01$ ). Di(1-propenyl)sulfide and 2-ethylthiophene showed notable associations with garlic texture, taste, flavor and overall impression ( $p < 0.01$ ). Methyl propenyl disulfide, dimethyl trisulfide and diallyl disulfide were significantly correlated only with garlic texture ( $p < 0.01$ ), while diallyl tetrasulfide correlated with garlic texture and taste ( $p < 0.01$ ); the relationship between these four compounds and the overall impression of garlic was not significant ( $p < 0.05$ ), indicating that methyl propenyl disulfide, dimethyl trisulfide, diallyl disulfide, and diallyl tetrasulfide may not be key indicators contributing to the sensory quality of garlic. A previous literature study has proved that 3-vinyl-1,2-dithiacyclohex-4-ene and 3-vinyl-1,2-dithiacyclohex-5-ene are isomers, and originate from diallyl disulfide and allyl 1-propenyl disulfide, respectively.<sup>44</sup> In

our study, they were both positively associated with garlic taste, appearance, flavor and overall impression ( $p < 0.05$ ), implying that they have an important role in garlic sensory quality. According to our CA of these organosulfur compounds (Fig. 2), S-methyl-L-cysteine and 3-vinyl-1,2-dithiacyclohex-5-ene were finally chosen as the representative organosulfur compounds.

Previous literature reports have suggested that garlic is an excellent source of free amino acids, and that amino acid composition has an important influence on the taste attributes of garlic and garlic products.<sup>34,45</sup> L-Alanine, L-proline and L-tyrosine may be responsible for the sweet taste, while glutamic acid is likely related with the umami taste.<sup>46</sup> Free amino acids also play an important role in the color formation of garlic and garlic products. For example, the research of Zhang *et al.*<sup>8</sup>





indicated that racemization of amino acids is probably correlated with the color change of black garlic during its processing. Cho *et al.*<sup>47</sup> identified candidate amino acids that may be connected with the color formation in crushed garlic; the results showed that amino acids other than glycine participated in the generation of blue pigments. Among the 43 differential sensory quality indicators of the garlic samples collected from four areas, a number of amino acids and organic acids showed various associations with sensory attributes of garlic (Table 3). Specifically, the garlic texture was not significantly correlated with L-alanine, serine, L-threonine, L-proline, L-asparagine or isocitric acid ( $p > 0.05$ ), but it was tightly linked to ornithine, glutamic acid, lanthionine, L-tyrosine, L-lysine, L-homoserine, ascorbic acid, propionic acid, ethanimidic acid and *n*-octanoic acid ( $p < 0.05$ ). Moreover, lanthionine, ascorbic acid and propionic acid were positively correlated with the garlic appearance, and in contrast, L-alanine, glutamic acid and L-asparagine presented negative correlations with the garlic appearance. Furthermore, lanthionine, L-tyrosine, L-homoserine, ornithine, ascorbic acid, propionic acid, *n*-octanoic acid and isocitric acid were obviously positively associated with the garlic taste, flavor and overall impression ( $p < 0.05$ ), implying that they make notable contributions to the sensory quality differences between garlic samples. Negative relationships were observed between L-alanine, serine, L-threonine, glutamic acid, L-asparagine and L-lysine and the garlic taste, flavor, and overall impression ( $p < 0.05$ ). Combining the correlations among these amino acids and organic acids, glutamic acid and L-tyrosine were selected as the representative indicators.

The main constituents of raw garlic are carbohydrates, of which the majority are known as fructans, a kind of fructose polymer with water solubility that provides a significant source of soluble dietary fiber.<sup>48,49</sup> Furthermore, carbohydrates account for not only the major constituents of the cell wall but also the soluble sugars in the cytoplasm, and so carbohydrates (*i.e.* fructose, glucose and sucrose) are also closely related to the garlic texture.<sup>50</sup> As seen in Table 3, arabinose, D-fructose and D-glucose were positively connected with the garlic texture, which suggests that high contents of arabinose, D-fructose and D-glucose may cause a soft garlic texture, whereas, sucrose presented a negative correlation with the garlic texture, indicating that a high level of sucrose probably leads to a hard garlic texture. However, negative associations between D-fructose, D-glucose and the garlic taste, appearance, flavor and overall impression were observed. This could be explained according to previous research:<sup>51</sup> with an increasing content of fructans, the water-retentive performance of garlic cells may be reduced and the content of organosulfur compounds responsible for pungency may be relatively increased. Moreover, fructans can decompose into fructose, glucose and sucrose, which are responsible for the sweet taste of garlic. Thus, we can conclude that high concentrations of D-fructose and D-glucose are likely to lead to a decrease in garlic taste, appearance, flavor and overall impression, while sucrose displays the opposite results. After correlations among these sugars and other compounds were analyzed, D-fructose and propene were ultimately determined to be representative indicators.

Taken together, a total of seven representative indicators of garlic sensory quality were identified according to the multivariate statistical analysis. These indicators were the  $a^*$  value, and the contents of *S*-methyl-L-cysteine, 3-vinyl-1,2-dithiacyclohex-5-ene, glutamic acid, L-tyrosine, D-fructose and propene, all of which were tightly correlated with the garlic sensory attributes and had the ability to reflect the sensory quality of garlic samples based on both physical and chemical traits.

### Establishment and validation of the BPANN model

Among the seven representative indicators of garlic sensory quality, the  $a^*$  value was the key color parameter, *S*-methyl-L-cysteine and 3-vinyl-1,2-dithiacyclohex-5-ene were mainly responsible for the spicy taste and garlic flavor, glutamic acid and L-tyrosine had an important influence on the taste attributes of garlic, D-fructose contributed to the sweet taste of garlic and was positively correlated with the garlic texture, and propene was used to represent other compounds in the garlic.

In order to establish a BPANN model for the sensory quality assessment of garlic and garlic products, the seven representative indicators ( $a^*$  value, *S*-methyl-L-cysteine, 3-vinyl-1,2-dithiacyclohex-5-ene, glutamic acid, L-tyrosine, D-fructose and propene) were standardized (such that they ranged from 0 to 1) and set as the input data of the BPANN model, while the sensory scores obtained by the traditional sensory evaluation were treated as the output data of the BPANN model. Because the measurement of garlic quality indicators and the traditional sensory evaluation were both repeated three times, altogether 84 groups of data were generated. 60 groups of data were randomly selected and regarded as training samples for the BPANN model. After the training process using the 60 groups of data was completed, the remaining 24 groups of data were used for predictions to test the accuracy of the model.

The hidden layers are the most important layers in the architecture of the BPANN model; the number of hidden layers and the number of neurons in the hidden layers influence the prediction accuracy of the BPANN model.<sup>52</sup> A lower number of neurons in the hidden layers is not competent in obtaining information from a complicated data set, while an excessive number of neurons in the hidden layers may increase the time required for data training and cause "overfitting". Finally, one hidden layer with six neurons was determined in this study to offer the best prediction of the sensory quality of garlic. A continuous training process was completed until the mean square error (MSE) between the training data and test data was lower than 0.006, and the other parameters were set as follows: maximum number of cycles, 500 000; learning rate, 0.7; and momentum factor, 0.5. Fig. 3 shows the predicted results by the developed BPANN model. Test data (24 groups) was used to validate the prediction accuracy of the BPANN model, and a linear regression equation between the predicted score and the traditional sensory score for the garlic was  $y = 1.0538x - 0.3421$ . The  $R^2$  value and MSE were 0.9866 and 0.0034, respectively. This result indicates that the BPANN model matches well to the sensory evaluation and that the prediction precision is



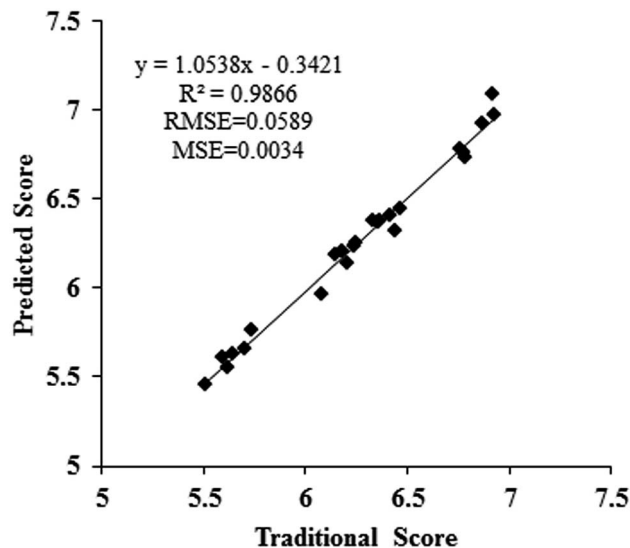


Fig. 3 Relationship between predicted score and traditional sensory score.

high. Thus, the sensory quality of garlic can be precisely predicted by the developed BPANN model. Furthermore, a BPANN model with a more reliable training set calibrated from a bigger group set and a larger amount of data is expected to be developed in the future. This could be used to predict the garlic sensory quality of garlic not only from China but also from elsewhere in the world.

## Conclusions

A rapid and simple method based on metabolomics and an artificial neural network has been developed for the sensory quality determination of garlic and garlic products. Quality indicators were obtained mainly using GC/MS and HPLC-MS/MS, and the quality indicator data were fused at a low level. Through PLS-DA, ANOVA and CA, seven representative indicators ( $a^*$  value, and contents of *S*-methyl-L-cysteine, 3-vinyl-1,2-dithiacyclohex-5-ene, glutamic acid, L-tyrosine, D-fructose, and propene) were identified and employed as the input data for a BPANN model. Finally, a BPANN model for predicting the sensory quality of garlic and garlic products was constructed. The fitting degree according to the linear regression equation between the predicted scores and the traditional sensory scores of garlic was high, which suggests that the BPANN model built in this study could predict the sensory quality of garlic accurately. In summary, this is the first method reported for sensory quality assessment of garlic and garlic products using metabolomics and an artificial neural network, while the BPANN model could simplify the quality assessment process and improve the precision of prediction.

## Conflicts of interest

The authors declare that they have no conflict of interest.

## Acknowledgements

This work has been carried out with support from the National Key Research and Development Programs of China (2016YFD0400201-8).

## References

- 1 H. Amagase, B. L. Petesch, H. Matsuura, S. Kasuga and Y. Itakura, Intake of garlic and its bioactive components, *J. Nutr.*, 2001, **131**, 955S–962S.
- 2 V. Lanzotti, The analysis of onion and garlic, *J. Chromatogr. A*, 2006, **1112**, 3–22.
- 3 J. Reiter, N. Levina, M. vander Linden, M. Gruhlke, C. Martin and A. J. Slusarenko, Diallylthiosulfinate (allicin), a volatile antimicrobial from garlic (*Allium sativum*), kills human lung pathogenic bacteria, including MDR strains, as a vapor, *Molecules*, 2017, **22**, 1711.
- 4 N. Morihara, A. Hino, S. Miki, M. Takashima and J. I. Suzuki, Aged garlic extract suppresses inflammation in apolipoprotein E-knockout mice, *Mol. Nutr. Food Res.*, 2017, **61**, 1700308.
- 5 M. Atkin, D. Laight and M. H. Cummings, The effects of garlic extract upon endothelial function, vascular inflammation, oxidative stress and insulin resistance in adults with type 2 diabetes at high cardiovascular risk. A pilot double blind randomized placebo controlled trial, *J. Diabetes Complicat.*, 2016, **30**, 723–727.
- 6 D. T. Constenla and J. E. Lozano, Effect of pretreatments and processing conditions on the chemical, physical, microbiological and sensory characteristics of garlic paste, *J. Food Process Eng.*, 2005, **28**, 313–329.
- 7 J. E. Pardo, J. Escribano, R. Gomez and A. Alvarruiz, Physical-chemical and sensory quality evaluation of garlic cultivars, *J. Food Qual.*, 2007, **30**, 609–622.
- 8 X. Zhang, N. Li, X. Lu, P. Liu and X. Qiao, Effects of temperature on the quality of black garlic, *J. Sci. Food Agric.*, 2016, **96**, 2366–2372.
- 9 W. Q. Li, H. Zhou, M. Y. Zhou, X. P. Hu, S. Y. Ou, R. A. Yan, X. J. Liao, X. S. Huang and L. Fu, Characterization of phenolic constituents inhibiting the formation of sulfur-containing volatiles produced during garlic processing, *J. Agric. Food Chem.*, 2015, **63**, 787–794.
- 10 N. Utama-ang, T. Cheewinworasak, N. Simawonthamgul and R. S. Samakradhamrongthai, Effect of drying condition of Thai garlic (*Allium sativum* L.) on physicochemical and sensory properties, *Int. Food Res. J.*, 2018, **25**, 1365–1372.
- 11 J. Wakamatsu, T. D. Stark and T. Hofmann, Taste-active maillard reaction products in roasted garlic (*Allium sativum*), *J. Agric. Food Chem.*, 2016, **64**, 5845–5854.
- 12 L. Eroman Unni, O. P. Chauhan and P. S. Raju, High pressure processing of garlic paste: effect on the quality attributes, *Int. J. Food Sci. Technol.*, 2014, **49**, 1579–1585.
- 13 D. S. Wishart, Metabolomics: applications to food science and nutrition research, *Trends Food Sci. Technol.*, 2008, **19**, 482–493.



- 14 S. Y. Lee, S. Lee, S. Lee, J. Y. Oh, E. J. Jeon, H. S. Ryu and C. H. Lee, Primary and secondary metabolite profiling of doenjang, a fermented soybean paste during industrial processing, *Food Chem.*, 2014, **165**, 157–166.
- 15 W. L. Hung and Y. Wang, Metabolite Profiling of Candidatus *Liberibacter* Infection in Hamlin Sweet Oranges, *J. Agric. Food Chem.*, 2018, **66**, 3983–3991.
- 16 J. Bi, X. Wang, Q. Chen, X. Liu, X. Wu, Q. Wang, J. Lv and A. Yang, Evaluation indicators of explosion puffing Fuji apple chips quality from different Chinese origins, *LWT–Food Sci. Technol.*, 2015, **60**, 1129–1135.
- 17 J. Lyu, L. Y. Zhou, J. F. Bi, X. Liu and X. Y. Wu, Quality evaluation of yellow peach chips prepared by explosion puffing drying, *J. Food Sci. Technol.*, 2015, **52**, 8204–8211.
- 18 Y. Liu, X. Sun and A. Ouyang, Nondestructive measurement of soluble solid content of navel orange fruit by visible-NIR spectrometric technique with PLSR and PCA-BPNN, *LWT–Food Sci. Technol.*, 2010, **43**, 602–607.
- 19 J. Sun, Learning algorithm and hidden node selection scheme for local coupled feedforward neural network classifier, *Neurocomputing*, 2012, **79**, 158–163.
- 20 L. Wang, D. Liu, H. Pu, D. W. Sun, W. Gao and Z. Xiong, Use of hyperspectral imaging to discriminate the variety and quality of rice, *Food Anal. Methods*, 2015, **8**, 515–523.
- 21 L. Lu, C. Fang, Z. Hu, X. Hu and Z. Zhu, Grade classification model tandem BpNN method with multi-metal sensor for rice eating quality evaluation, *Sens. Actuators, B*, 2019, **281**, 22–27.
- 22 X. Meng, M. Zhang and B. Adhikari, Prediction of storage quality of fresh-cut green peppers using artificial neural network, *Int. J. Food Sci. Technol.*, 2012, **47**, 1586–1592.
- 23 A. C. Kimbaris, N. G. Siatis, D. J. Daferera, P. A. Tarantilis, C. S. Pappas and M. G. Polissiou, Comparison of distillation and ultrasound-assisted extraction methods for the isolation of sensitive aroma compounds from garlic (*Allium sativum*), *Ultrason. Sonochem.*, 2006, **13**, 54–60.
- 24 F. Priego-Capote and M. D. L. D. Castro, HS-GC/MS volatile profile of different varieties of garlic and their behavior under heating, *Anal. Bioanal. Chem.*, 2016, **408**, 3843–3852.
- 25 S. N. Lee, N. S. Kim and D. S. Lee, Comparative study of extraction techniques for determination of garlic flavor components by gas chromatography-mass spectrometry, *Anal. Bioanal. Chem.*, 2003, **377**, 749–756.
- 26 J. Liu, F. Ji, F. Chen, W. Guo, M. Yang, S. Huang, F. Zhang and Y. Liu, Determination of garlic phenolic compounds using supercritical fluid extraction coupled to supercritical fluid chromatography/tandem mass spectrometry, *J. Pharm. Biomed. Anal.*, 2018, **159**, 513–523.
- 27 E. Kovats, Gas chromatographic characterization of organic substances in the retention index system, *Adv. Chromatogr.*, 1965, **1**, 229–247.
- 28 J. Lisec, N. Schauer, J. Kopka, L. Willmitzer and A. R. Fernie, Gas chromatography mass spectrometry-based metabolite profiling in plants, *Nat. Protoc.*, 2006, **1**, 387–396.
- 29 T. D. Dan, Z. Bin, Z. L. Yao, H. X. Song, L. X. Jun and Z. Yan, Kinetics of “laba” garlic greening and its physiochemical properties treated by dense phase carbon dioxide, *LWT–Food Sci. Technol.*, 2015, **64**, 775–780.
- 30 GB5009.3-2016, *National Food Safety Standard, Determination of Moisture in Foods*, Standards Press of China, China, Beijing, 2016.
- 31 Y. Q. Zhang, Z. P. Wang, Y. M. Song, D. L. Luan and C. H. Xue, Comparative study on the determination of reducing sugar, *Sci. Technol. Food Ind.*, 2010, **6**, 321–326.
- 32 D. A. Ramirez, D. A. Locatelli, C. A. Torres-Palazzolo, J. C. Altamirano and A. B. Camargo, Development of garlic bioactive compounds analytical methodology based on liquid phase microextraction using response surface design. Implications for dual analysis: cooked and biological fluids samples, *Food Chem.*, 2017, **215**, 493–500.
- 33 S. Kim, S. L. Park, S. Lee, S. Y. Lee, S. Ko and M. Yoo, UPLC/ESI-MS/MS analysis of compositional changes for organosulfur compounds in garlic (*Allium sativum* L.) during fermentation, *Food Chem.*, 2016, **211**, 555–559.
- 34 A. Montañó, F. J. Casado, A. de Castro, A. H. Sánchez and L. Rejano, Vitamin content and amino acid composition of pickled garlic processed with and without fermentation, *J. Agric. Food Chem.*, 2004, **52**, 7324–7330.
- 35 C. N. Horita, A. M. Farias-Campomanes, T. S. Barbosa, E. A. Esmerino, A. Gomesda Cruz, H. M. A. Bolini, M. A. A. Meireles and M. A. R. Pollonio, The antimicrobial, antioxidant and sensory properties of garlic and its derivatives in Brazilian low-sodium frankfurters along shelf-life, *Food Res. Int.*, 2016, **84**, 1–8.
- 36 D. T. Constenla and J. E. Lozano, Effect of pretreatments and processing conditions on the chemical, physical, microbiological and sensory characteristics of garlic paste, *J. Food Process Eng.*, 2005, **28**, 313–329.
- 37 M. C. Meilgaard, B. T. Carr and G. V. Civille, *Sensory evaluation techniques*, CRC Press, 1999.
- 38 Y. K. Jain and S. K. Bhandare, Min max normalization based data perturbation method for privacy protection, *International Journal of Computer & Communication Technology*, 2011, **2**, 45–50.
- 39 M. Farrés, S. Platikanov, S. Tsakovski and R. Tauler, Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation, *J. Chemom.*, 2015, **29**, 528–536.
- 40 M. Keusgen, H. Schulz, J. Glodek, I. Krest, H. Krüger, N. Herchert and J. Keller, Characterization of some *Allium* hybrids by aroma precursors, aroma profiles, and alliinase activity, *J. Agric. Food Chem.*, 2002, **50**, 2884–2890.
- 41 J. Ahmed and U. S. Shivhare, Thermal kinetics of color change, rheology, and storage characteristics of garlic puree/paste, *J. Food Sci.*, 2001, **66**, 754–757.
- 42 R. E. Gonzalez, V. C. Soto, M. M. Sance, A. B. Camargo and C. R. Galmarini, Variability of solids, organosulfur compounds, pungency and health-enhancing traits in garlic (*Allium sativum* L.) cultivars belonging to different ecophysiological groups, *J. Agric. Food Chem.*, 2009, **57**, 10282–10288.
- 43 A. Montano, V. M. Beato, F. Mansilla and F. Orgaz, Effect of genetic characteristics and environmental factors on



- organosulfur compounds in garlic (*Allium sativum* L.) grown in Andalusia, Spain, *J. Agric. Food Chem.*, 2011, **59**, 1301–1307.
- 44 M. Molina-Calle, F. Priego-Capote and M. D. de Castro, HS–GC/MS volatile profile of different varieties of garlic and their behavior under heating, *Anal. Bioanal. Chem.*, 2016, **408**, 3843–3852.
  - 45 Y. Ueda, H. Kawajiri, N. Miyamura and R. Miyajima, Content of some sulfur-containing components and free amino acids in various strains of garlic, *J. Jpn. Soc. Food Sci.*, 2009, **38**, 429–434.
  - 46 H. Kato, M. R. Rhue and T. Nishimura, *Role of free amino acids and peptides in food taste*, 1989, pp. 158–174.
  - 47 J. Cho, E. J. Lee, K. S. Yoo, S. K. Lee and B. S. Patil, Identification of candidate amino acids involved in the formation of blue pigments in crushed garlic cloves (*Allium sativum* L.), *J. Food Sci.*, 2009, **74**, 11–16.
  - 48 P. M. Chandrashekhara and Y. P. Venkatesh, Immunostimulatory properties of fructans derived from raw garlic (*Allium sativum* L.), *Bioact. Carbohydr. Diet. Fibre*, 2016, **8**, 65–70.
  - 49 A. T. Mankarios, C. F. G. Jones, M. C. Jarvis, D. R. Threfall and J. Friend, Hydrolysis of plant polysaccharides and GLC analysis of their constituent neutral sugars, *Phytochemistry*, 1979, **18**, 419–422.
  - 50 C. Ohsumi and T. Hayashi, Carbohydrate analysis of an interspecific hybrid between onion and garlic, *Biosci., Biotechnol., Biochem.*, 1994, **58**, 959–960.
  - 51 M. J. Havey, C. R. Galmarini, A. F. Gökçe and C. Henson, QTL affecting soluble carbohydrate concentrations in stored onion bulbs and their association with flavor and health-enhancing attributes, *Genome*, 2004, **47**, 463–468.
  - 52 S. Karsoliya, Approximating number of hidden layer neurons in multiple hidden layer BPNN architecture, *International Journal of Engineering Trends and Technology*, 2012, **3**, 714–717.

