Journal of Integrative Agriculture 2021, 20(5): 1193–1203

Available online at www.sciencedirect.com

ScienceDirect

RESEARCH ARTICLE

Establishment and application of an accurate identification method for fragrant soybeans

ZHANG Yong-fang[1, 2, 3], ZHANG Chun-yan[3], ZHANG Bo[4], YIN Man[3], HONG Hui-long[3], YU Li-li[3], GAO Hua-wei[3], GU Yong-zhe[3], LIU Zhang-xiong[3], LI Fu-heng[1], QIU Li-juan[3]

[1] *College of Life Sciences, Northeast Agricultural University, Harbin 150030, P.R.China*

[2] *College of Life Sciences, Shanxi Datong University, Datong 037009, P.R.China*

[3] *National Key Facility for Crop Gene Resources and Genetic Improvement (NFCRI)/Key Laboratory of Germplamsm Utilization, Ministry of Agriculture and Rural Affairs/Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, P.R.China*

[4] *School of Plant and Environmental Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA 24060, USA*

Abstract

In order to screen the aroma characteristics of soybean, a new method was established which can quickly quantify the content of 2-acetyl-1pyrroline (2-AP), an important compound related to soybean aroma, using gas chromatography-mass spectrometry (GC-MS). Based on peak profile, total peak area and retention time as test indexes, an accurate identification method for fragrant soybeans was established. The optimum parameters of the protocol consisted of column temperature 70°C, sample injector temperature 180°C, optimum extraction alcohol content 1 mL, NaCl content 0.1 g, ultrasonication time 10 min, and extraction time 1 h, which were established by using the orthogonal test of single factors and three factors with four levels ($L_9(3)^4$). 2-AP content of leaves had significant correlations with seeds, which were easier to measure. The protocol was simple and easy to carry out, consumed only small amounts of reagents, and provided accurate and reliable results with good reproducibility. A total of 101 soybean genotypes from different geographical sources were analyzed using this protocol. The results showed that the average content of 2-AP was 0.29 mg L^{-1}, ranging from 0.094 to 1.816 mg L^{-1}, and the genetic diversity index was 0.54. Among all genotypes-tested, they were classified into three grades, including seven elite genotypes identified as “grade one fragrant soybeans”, which were Zhonglong 608, Heinong 88, Ha13-2958, Hongmiandou, Heinong 82, Huangmaodou, and Jiyu 21. These results provide both an identification technique and several elite aroma genotypes for gene discovery and good quality breeding in soybean.

Keywords: soybean, 2-acetyl-1pyrroline, GC-MS, quantification method, germplasm

Received 2 March, 2020 Accepted 10 July, 2020

ZHANG Yong-fang , E-mail: zyf_208@163.com; Correspondence QIU Li-juan, Tel: +86-10-82105843, Fax: +86-10-82105840, E-mail: qiulijuan@caas.cn; LI Fu-heng, Tel/Fax: +86-451-55190002, E-mail: lifuheng1963@126. com

doi: 10.1016/S2095-3119(20)63328-7

1. Introduction

Soybean is one of the major grain crops for most of the world’s population. Because of its rich nutritional value, pharmaceutical value, special flavor, and taste, it is widely loved by consumers (Wanchana *et al*. 2005). Different soybean varieties have different fragrance characteristics

due to different components and their contents. Along with the improvement of people's living quality, increasing attention has been paid to food sensory qualities. Therefore, it is of great significance to actively develop and utilize soybeans which provide better flavor.

Aroma is one of the most challenging plant characteristics in plant breeding because of its complex composition. For example, in early days, sensory evaluation methods such as boiling seeds in a small bottle and then chewing for evaluation or identification by simply smelling the leaves after KOH soaking were commonly used to analyze the fragrances of samples (Wanchana *et al*. 2005). Similar aroma evaluation methods were also reported in soybeans (Juwattanasomran *et al*. 2011). However, these methods are time-consuming and labor-intensive, and they are also affected by sensory differences of panelists and suffer from large differences in evaluation results. Therefore, further improvement of chemical analysis and instrumentation methods to efficiently evaluate the aroma quality of a large number of fragrant soybean samples is also important. Quantitative analysis of volatile components have been used, where samples are first heated to extract volatile components, such as by solvent extraction, direct distillation extraction, stem distillation-solvent extraction, or supercritical fluid extraction (Reineccius 2006), followed by analysis of the extracted compounds by gas chromatography-mass spectrometry (GC-MS). The method of GC-MS is very effective and has been widely used to analyze samples because of its robustness and easily-quantifiable results (Masuda *et al*. 1991; Fushimi and Masuda 2001; Plonjarean *et al*. 2007; Wu *et al*. 2009). However, the heat extraction method is time-consuming and labor-intensive, requiring many steps for sample preparation and extraction reagents, so it is not appropriate for analyzing large numbers of samples.

In recent years, many studies concerning the genetic and chemical aspects have shown that the fragrance of soybean is mainly determined by 2-acetyl-1-pyrroline (2-AP). 2-AP is a volatile compound imparting a flavor similar to popcorn, with molecular formula C_6H_9NO, a molecular weight of 111.14, and a boiling point of 182.9°C (760 mm Hg). 2-AP is formed naturally in various organisms, including plants (Widjaja *et al*. 1996; Yoshihashi 2002), animals (Brahmachary *et al*. 1990) and microorganisms (Snowdon *et al*. 2006; Adams and De 2007), and it can also be formed in food products (Buttery *et al*. 1983; Schieberle and Wener 1991). Fushimi and Masuda (2001) reported that vegetable soybean cultivars "dadachamame" and "chakaori" contained aromas similar to rice, such as jasmine and Indian fragrant rice. Further studies showed that the aroma of Dadachamam originated from 2-AP using a solvent extraction and GC-MS method (Plonjarean *et al*. 2007). 2-AP was also the same volatile compound reported as Buttery *et al*. (1983) and found in connection with fragrance in rice. Wu *et al*. (2009) also showed that 2-AP was the main aroma compound in soybeans. Arikit *et al*. (2010) and Juwattanasomran *et al*. (2010) demonstrated that aromatic soybeans lacked betaine dehydrogenase (GmBADH) activity. Sequencing analysis found that a 2-bp deletion in exon 10 of the *GmBADH2* gene resulted in early termination of *GmBADH2* expression and promoted the synthesis of 2-AP. A codominant PCR-based marker Gm2AP was designed based on the 2-bp deletion, and it can be used to identify the aroma trait in soybeans. This development led to the realization that 2-AP was naturally occurring and didn't need heating. However, 2-AP is easily oxidized, volatile and unstable, and the content of 2-AP in soybeans is extremely low which makes the quantification more difficult. Sriseadka *et al*. (2006) determined the aroma components of fragrant rice by headspace gas chromatography (HS-GC). Juwattanasomran *et al*. (2011) determined the concentration of soybean 2-AP by HS-GC with the internal standard substance 2,4-dimethylpyridine (DMP). The HS-GC method is fast and efficient, but the accuracy and repeatability of results are limited. The internal standard method can correct for errors caused by changes in the operating conditions, so the results are more accurate. However, it also has some disadvantages, such as difficulty in obtaining the internal standard, especially for internal standards labeled by isotopes; and the accuracy may be affected by environmental factors such as an operational error in adding the internal standard. The operation steps are especially complicated which makes it only suitable for the analysis of small numbers of samples. The external standard method has the advantages of simple pre-treatment steps and convenient operation, which make it suitable for the analysis of large numbers of samples. In addition, so far only five aromatic soybeans have been reported: Chamame (0.5795 mg L^{-1}, Japan), Kouri (0.5837 mg L^{-1}, Japan), Kaorihime (1.1600 mg L^{-1}, Japan), Yuagari musume (1.0085 mg L^{-1}, Japan), and Fukunari (0.6094 mg L^{-1}, Japan), while another five varieties (Okuhara Wase, Oishi Edamame, Shirono Mai, Chiang Mai 60 (CM60), and Jack) were non-aromatic soybeans. Whether other flavorful soybean varieties also contain 2-AP has not been determined.

In this study, a GC-MS external standard method for the quantification of flavor was established to improve on the above methods, increase the accuracy, repeatability and identification efficiency of flavor evaluation results, and provide a technical basis for transforming the results of soybean flavor evaluation from "subjective language description" to "objective numerical quantification". A total of 101 germplasms from different ecological regions were evaluated. The germplasms rich in 2-AP obtained by this

method lay a material and technical foundation for the genetic improvement of flavor in soybean varieties.

2. Materials and methods

2.1. Soybean materials

In this experiment, a total of 101 representative genotypes from China, Japan, Thailand, the USA and Poland were selected (Table 1). They included 93 Chinese soybean accessions (36 from the North region, 28 from the Huang-Huai region and 29 from the South region, and eight introduced accessions (one from Japan, one from Thailand, five from the USA, and one from Poland). Each accession was planted in a single row with a length of 2 m, between-row spacing 45 cm, in-row spacing 20 cm, and 10 plants per row. The field management followed local field production recommendations. Four protective rows on the side of the field were included to prevent edge effects. Top leaves in the flowering period or harvested soybean seeds were collected for 2-AP evaluation by GC-MS. These accessions were planted and sampled at the Shunyi Experimental Station of the Institute of Crop Sciences, Chinese Academy of Agricultural Sciences.

2.2. 2-AP identification

Preparation of standard stock solution A total of 10 mg of 2-AP standard substance was dissolved in 100 mL of anhydrous ethanol to obtain a 1 000-mg L^{-1} stock solution, which was placed in a 4°C refrigerator until needed.

Preparation of standard solution The standard stock solution was diluted to 4, 2, 1, 0.5, and 0.25 mg L^{-1} standard solutions sequentially by the method of stepwise dilution, and the 1 mL standard solutions were filtered with a pore diameter of 0.22 μm filter membrane and placed into a 32 mm ×11 mm glass vial. The vial was sealed immediately with a Polypropylene screw cap with a tetrafluoroethylene (TFE) septum (10-15-1256-TF, Shimadzu-GL Sciences (Shanghai) Laboratory China).

Preparation of soybean samples Soybean seeds were finely ground and the size of powder should be less than 60 mesh, or individual soybean leaves in the flowering period were sheared into pieces with diameters of 0.2–0.5 mm. Portions (0.4 g) of these samples were extracted with 1.5 mL ethanol (Yoshihashi *et al*. 2002) and Na_2SO_4 (0.2 g) in a 10-mL centrifuge tube and then were shaken and mixed for 1 min on the vortex mixer and subjected to ultrasonic extraction for 20 min in an ultrasonic cleaner. Then 0.2 g NaCl was added to the mixture and were extracted at room temperature for 2 h. After centrifugation (12 000 r min^{-1}, 10 min, 4°C), the 1 mL supernatant was filtered with a filter membrane (pore diameter 0.22 μm) and placed into vials. Polypropylene screw caps with TFE septa were used as vial covers.

GC-MS 2010 chromatography-mass spectrometry conditions Agilent DB-WAX Capillary (30 m×0.25 mm i.d.×0.25 μm film thickness), column programming-temperature: maintained at 60°C for 2 min, increased to 100°C at 10°C min^{-1}, increased to 230°C at 30°C min^{-1} and maintained for 5 min; constant linear velocity: 41.2 cm s^{-1}; injection port temperature: 180°C; carrier gas: high purity helium gas (purity>99.99%); unsplit stream sampling, sampling amount: 1 μL; and mass spectrometry conditions: electron bombardment ion source, ion source temperature: 200°C; ionization energy: 70 eV; interface temperature:

Table 1 The source and number of soybean germplasms tested in this study

Origin	Name of genotype	Number of genotype
North region, China	Zhonglong 608, Heinong 88, Ha13-2958, Heinong 82, Jiyu 21, Tonghuapingdingxiang, Jiyu 701, Zhengguang 1, Liushiribaidou, Dongnong 48, Longken 316, Dushidaqingdou, Suinong 79, Wangzhuangheidou, Heijiazi, Hefeng 30, JiNF58, Baipidou, Changchunmancangjin, Heidou, Wuxing 2, Tianlongyi, Beidou 36, Dongliaokeshuang, Zaoshu 1, Niuyajing, Xiaolihuang, Dahuangdou, Suinong 1, Heinong 2, Youhuangdou, Baogongdou, Beifeng 16, Hefeng 57, Hefeng 51, and Silihuang	36
Huanghui region, China	Minquanniumaohuang, Qiyuezha, Xudou 23, Zhongkemaodou 2, Zhengzhoudaziqingdou, Shangdou 6, Qihuang 39, Wandou 33, Qihuang 42, Qinyangshuidou, Runanpingdingshi, Qihuang 35, Ruidou 1, Shangdou 14, Shangdou 151, Suininghuangxudadou, Daqingpihuangdou, Xixiaxiaozihuang, Lingbaobaijingdou, Zhonghuang 14, Jidou 15, Jiuyuehan, Zhonghuang 20, Xixianpingdingshidadou, Qihuang 34, Ninglingtian'edan, Hongmiandou, and Huangdou	28
South region, China	Maodou 3, Xihuangdou 9, Fushendou, Zhexiandou 2, Liuyuebai, Pohuang, Yingshandalihuang, Xiangchundou 24, Houzimao, Huangpibayuezha, Guixiadou 2, Yantianqingpidou, Gongdou 10, Rugaociyutou'erbing, Qingpiqingren, Jiuyueba, Nannongdahongdou, Chahuangdaidou 1, Tianmendazihuang, Shanbaidou, Huangpishanzibai, Fengxiansuidaohuang, Edou 2, Zhexiandou 3, Liuchengshiyuehuang, Baoguhuang 8, Gandou 2, Jurongxiaozihuang, and Jiuyuebaimao	29
Plant introductions from Japan, Thailand, USA, and Poland	Kaorihime, CM60, PI509100, PI398682, PI416762, Essex, PI561395, and Dunajka	8

220°C; full scanning mode, scanning range: 35–500 m/z.

Qualitative analysis 2-AP qualitative analysis information was retrieved from the NIST library. Taking the 1 mg L^{-1} standard as the test material, scanning and monitoring 2-AP according to GC-MS conditions described above were used to obtain a total ion flow diagram, then integrating the obtained peaks and searching for similarity allowed the determination of the peak time of 2-AP.

Quantitative analysis Quantitative analysis was carried out by determining the peak time of 2-AP and drawing the standard curve. Similarly, the selective ion monitoring (SIM) method was established according to the conditions of GC-MS, with ion fragment 83 as target ion, fragments 69 and 111 as reference ions, and 2-AP peak output time as the standard analysis standard. The standard solutions were arranged from low to high in the sample tray, and each concentration was injected with three needles. The average of the obtained peak areas was taken as the ordinate, and the corresponding standard solution concentration was taken as the abscissa to fit the standard curve. After the peak area of the sample to be measured was obtained, the 2-AP concentration was calculated from the obtained standard curve. After every 10 sample bottles, a sample of Zhengguang 1 was included for quality supervision.

Extraction optimization test scheme In the gas quality analysis, the column temperature of GC and the injector temperature of MS would affect the peak time and test effort of the samples. Therefore, the column temperature (gradient setting: 60, 70 and 80°C) and the injector temperature of the MS (gradient setting: 170, 180 and 190°C) were optimized by the single factor experimentation with sample Zhengguang 1 as the test material. In addition, alcohol content, NaCl content, ultrasonication time, and extraction time would all affect the quantification of 2-AP (such as peak area, retention time, peak shape, etc.). Therefore, these four factors of the NaCl quantity (A), the volume of alcohol (B), extraction time (C), and ultrasonication time (D) were taken as the examination factors simultaneously, and each factor has three levels to be optimized (Table 2). An orthogonal experimental design $L_9(3^4)$ (Table 3) in the extraction mode was used for optimizing the extraction conditions of 2-AP volatile substances.

The precision of the instrument and the method precision Zhengguang 1 was used to verify the precision of the instrument. Sample injection was repeated six times according to the selected chromatographic conditions to obtain the peak area, peak shape and peak time of 2-AP and to calculate its relative standard deviation. Four samples were prepared in parallel, and the contents of 2-AP were measured under the above selected conditions to evaluate the precision of the method.

2.3. Statistics

The average value, standard deviation, maximum value, minimum value, variation range, and coefficient of variation of 2-AP contents of each accession were calculated by Microsoft Excel 2013, and the accessions were further classified. The grading standards were as follows: firstly, the overall average value (x_i) and standard deviation (σ) of a certain trait of the tested varieties were calculated; and secondly, we divided these materials into three grades including first grade [$x_i \geq +\sigma$], second grade [$x_i \geq +0.1\sigma$] and third grade [$x_i < x$]. The ratio of the number of observed individuals to the total number in each level was used to calculate the diversity index. The calculation formula was: $H' = -\sum P_i \ln P_i$, where n is the number of phenotype levels of a certain trait, P_i is the ratio of the number of individuals

Table 2 Three levels for each factor in the orthogonal experiment

Factor level	A (NaCl) (g)	B (alcohol) (mL)	C (extraction time) (min)	D (ultrasonication time) (min)
1	0.1	1.0	1.0	10.0
2	0.2	1.5	1.5	20.0
3	0.3	2.0	2.0	30.0

Table 3 Factors and levels for optimizing test design

Horizontal combination	A (NaCl) (g)	B (alcohol) (mL)	C (extraction time) (min)	D (ultrasonication time) (min)
1	0.1	1.0	1.0	10.0
2	0.1	1.5	1.5	20.0
3	0.1	2.0	2.0	30.0
4	0.2	1.0	1.5	30.0
5	0.2	1.5	2.0	10.0
6	0.2	2.0	1.0	20.0
7	0.3	1.0	2.0	20.0
8	0.3	1.5	1.0	30.0
9	0.3	2.0	1.5	10.0

in level *i* of a certain trait to the total number, and ln is the natural logarithm. SPSS statistics 20 was used to analyze the correlation of the 2-AP contents between soybean seeds and leaves.

2.4. Genotyping soybean accessions for Gm2AP marker in the *BAHD2* gene

The polymerase chain reactions (PCR) were carried out by using two oligonucleotide primers of 5′-GGTCAG ATATGCAGTGCAAC-3′ and 5′-TTGACCCATTTCACAA TCCTAT-3′) that flanked the 2-bp deletion in exon 10 of the *BADH2* gene. The amplification products were identified with PAGE according to Arikit *et al*. (2011).

3. Results

3.1. Establishment of quantification protocol for 2-AP in soybean leaves

Control settings Because 2-AP is volatile, the use of odorous products such as plastic should be avoided throughout the entire experiment. The cleaning methods of sample bottles and other glassware were rigorous: clean with alcohol and a brush, treat with ultrasonication for 20 min, rinse with ultrapure water once, treat with ultrasonication for 20 min again, rinse with alcohol twice, ultrasound treatment twice again, and then dry samples at 40°C until the alcohol is completely volatilized. In each test, empty bottles and 1 mg L^{-1} standard were used as controls to calibrate the instrument.

Determination of retention time of 2-AP and plotting of the standard curve We used 1 mg L^{-1} 2-AP as the standard material. A full scan was performed according to the GC-MS conditions described above to obtain a GC-MS total ion chromatogram (Fig. 1). The obtained peak was integrated, and the 2-AP standard mass spectrum was searched through the NIST library (Fig. 2). The peak time of 2-AP was determined as 7.014 min, and the similarity to the 2-AP of the library was 91%. Similarly, a SIM method was established in which a 2-AP fragment ion chromatogram with different concentration gradients was obtained by taking the average value of the peak areas obtained as the vertical ordinate and the corresponding standard solution concentrations as the abscissa for fitting a standard curve (Fig. 3). The regression equation was y=115 373x+1 483.1 and the correlation coefficient (R^2) was 0.9998, which showed that the regression equation had a good fitting effect and a good linear correlation degree, so it can be used for quantification.

Single factor comparison (1) Effect of column temperature on total peak area and peak shape. The fixed injector temperature was set at 180°C. Because the column temperature was usually about 13°C lower than the temperature of the extraction reagent, the reagent used in this test was alcohol with a boiling point of 78°C, and the column temperature was sequentially set at 50°C, 60 and 70°C to analyze its influences on the peak time, peak area and peak shape of volatiles (Fig. 4). When the column temperature was 50°C, the peak time of the sample was 7.067 min and the peak area was 9 551 (Fig. 4). We observed that the peak shape was wide, not smooth, and the peak spacing was small and could not be separated easily. When the column temperature was 60°C, the peak time of the sample was 6.997 and the peak area was 20 354, which was 2.1 times the peak area at 50°C. In addition, the peak shape was smooth and the peak bottom was relatively flat. When the column temperature was 70°C, the peak time of

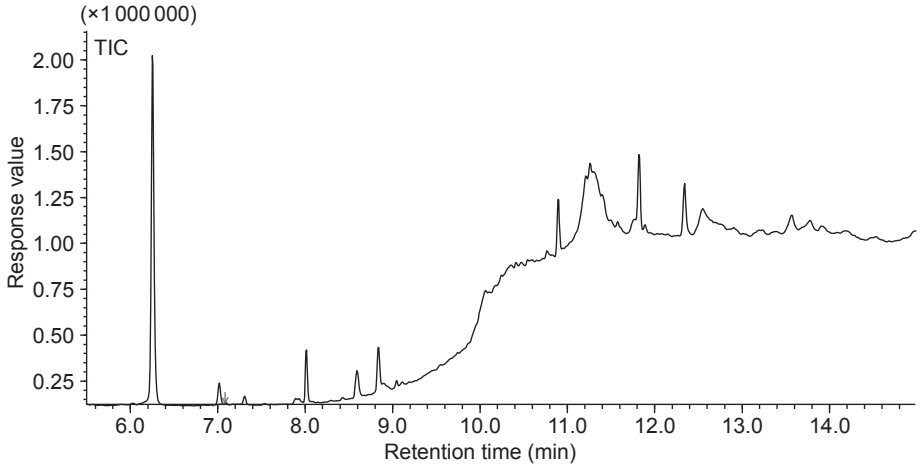

Fig. 1 Total ion flow diagram of 1 mg L^{-1} standard. TIC, total ions chromatograph.

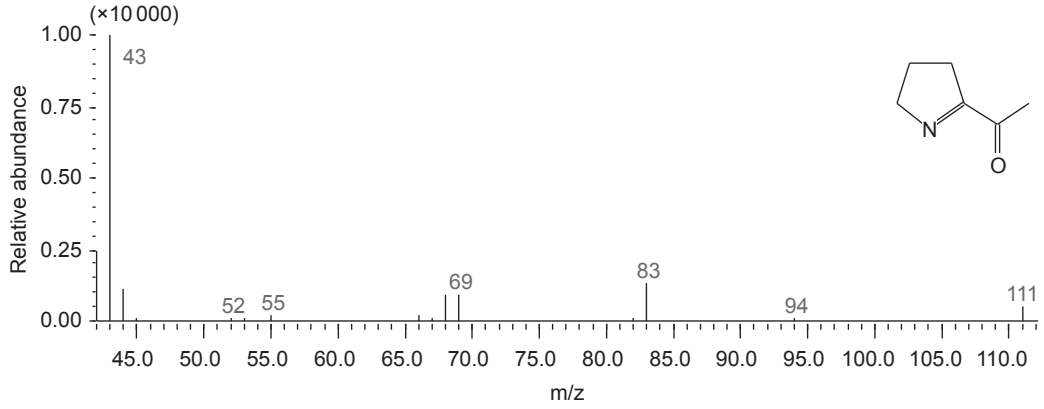

Fig. 2 National Insititute of Standards and Technology library ion fragment table.

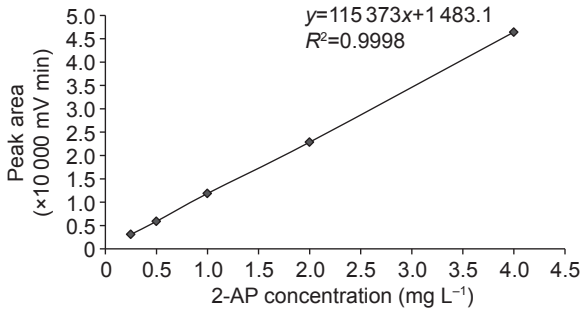

Fig. 3 2-Acetyl-1pyrroline (2-AP) standard curve established using peak areas of different 2-AP concentrations.

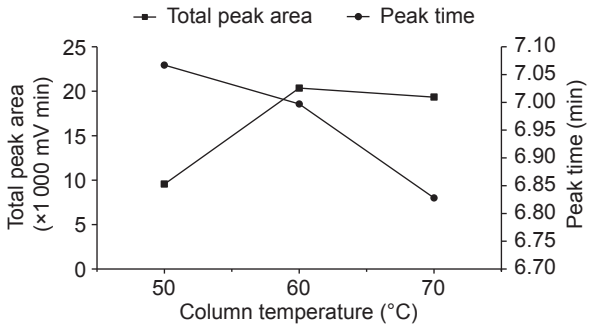

Fig. 4 Effect of column temperature on peak time and peak area.

the sample was 6.828 and the peak area was 19 342, which was about 0.95 times the peak area at 60°C, and the peak shape was narrower than that at 60°C. Therefore, increasing the column temperature will enhance the movement and accumulation speed of the samples, reduce the dissolution time of the target substance in the extraction reagent, shorten the peak times of volatile substances, and decrease the analysis time. However, a higher temperature may also affect the extraction balance of the substance to be detected, resulting in a narrower peak. Therefore, 60°C was used as the best column temperature in this study.

(2) Effect of injector temperature on total peak area and peak shape. With leaf of Zhengguang 1 as the sample, the effect of injector temperature was tested using the fixed column temperature of 60°C. Because 2-AP's boiling point is 182.9°C and the maximum operating temperature of DB-WAX column is 250°C, the temperature at the injection port was gradually set to 150, 180 and 210°C for analyzing the effect of injection temperature on the peak time, peak area and peak shape. When the injector temperature was 150°C, the peak time of the sample was 7.004, the peak area was 8 880 and the peak shape was wider (Fig. 5). When the column temperature was 180°C, the peak time of the sample was 7.006, and the peak area was 21 421, which was 2.4 times the area at injection temperature of 150°C. The peak shape was smooth and the peak bottom was flat. When the column temperature was 210°C, the peak time of the sample was 7.008, the peak area was 19 870, which was less than that at column temperature of 180°C, but the peak shape was narrower. Therefore, 180°C was used as the injector temperature in this study.

Results and analysis of the orthogonal experiment On the basis of a single factor test, an orthogonal test $L_9(3)^4$ with four factors and three levels was carried out (Table 4). The analysis showed that when the column temperature was 60°C and the injector temperature was 180°C, the peak times of different combinations ranged from 6.989 to 6.998 min, all within the allowable error range (retention time±0.1 min) (Table 4). The extreme difference analysis was simple and intuitive, and showed that the order of influence of the various factors on the extraction effect based on the effect level was: alcohol amount (B)>NaCl amount (A)>ultrasonic time (D)>extraction time (C) when the peak area was used as the evaluation index. For each factor,

the influence on peak area was in the order of: A1 (0.1 g) >A2 (0.2 g)>A3 (0.3 g), B1 (1 mL)>B2 (1.5 mL)>B3 (2 mL), C1 (1 mL)>C3 (2 mL)>C2 (1.5 mL), and D1 (10 min)>D3 (30 min)>D2 (20 min). Therefore, the optimum extraction conditions for GC-MS injection using soybean leaves were as follows: 0.1 g NaCl, 1 mL alcohol, 1 h extraction time, and 10 min ultrasonication time.

Alcohol content (B), NaCl content (A) and ultrasonication time (D) had significant effects on the extraction of 2-AP from the variance analysis (Table 5), while extraction time (C) had little effect on it, which was consistent with the factor influence results as they also showed extraction time as having the least effect. The major and minor relationships of the four factors were the same as the results of the range analysis.

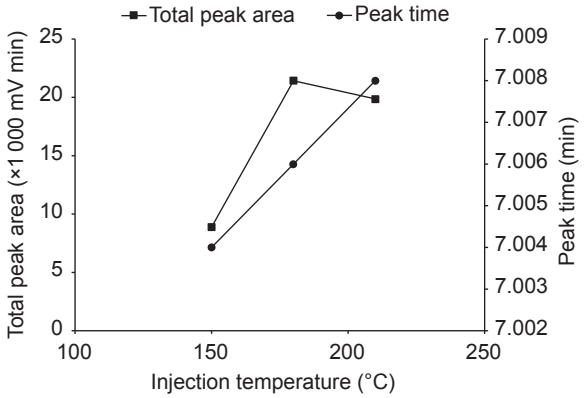

Fig. 5 Effect of injector temperature on peak time and peak area.

Precision evaluation (1) Instrument precision: The samples of 2-AP from Zhengguang 1 were run six times according to the selected optimal chromatographic conditions) and the sample treatment scheme). The results of 2-AP contents in the six runs were 0.398, 0.387, 0.395, 0.388, 0.396, and 0.389 mg L^{-1}, with an average value of 0.392 mg L^{-1} and a relative standard deviation RSD of 1.20%, indicating that the repeatability of the instrumentation met requirements of RSD less than 10%.

(2) Method precision: Four parallel measurements were conducted to quantify 2-AP in Zhengguang 1 samples, the 2-AP contents were 0.404, 0.410, 0.402, and 0.400 mg L^{-1} with an average value of 0.404 mg L^{-1}, and the relative standard deviation was 1.23%, indicating that the method had high precision.

Stability test Chromatographic analysis was carried out at 0, 1, 2, 4, 8, and 12 h according to the selected optimal chromatographic conditions and sample treatment scheme. The 2-AP content was calculated using the established standard curve. The results showed that the relative

Table 4 Combinations of three factors for the 2-acetyl-1pyrroline (2-AP) orthogonal test

Factor classification[1)]	A (NaCl) (g)	B (alcohol) (mL)	C (extraction time) (min)	D (ultrasonic time) (min)	Retention time (min)	Peak area (mV min)
1	0.1	1	1	10	6.997	21421.0
2	0.1	1.5	1.5	20	6.990	9818.5
3	0.1	2	2	30	6.990	7633.5
4	0.2	1	1.5	30	6.998	17183.5
5	0.2	1.5	2	10	6.991	13492.0
6	0.2	2	1	20	6.990	8009.5
7	0.3	1	2	20	6.992	14521.5
8	0.3	1.5	1	30	6.993	8891.0
9	0.3	2	1.5	10	6.989	5560.0
K_1	16825.17	17708.67	12773.83	13491.00		
K_2	12895.00	10733.83	10854.00	10783.17		
K_3	9657.50	7067.67	11882.33	11236.00		
R	7167.67	10641.00	1919.83	2707.83		

[1)] K_1, K_2 and K_3 represent the sums of the corresponding test results when the horizontal number in any column is one, two and three, and R represents the extreme difference.

Table 5 Variance analysis of the orthogonal experiment for peak area

Factor	SS	*df*	MS	*F*-value	Significance
A (NaCl)	18230830	2	9115415	9.006	0.002
B (alcohol)	74791390	2	37395695	36.947	0.000
C (extraction time)	1542098	2	771049	0.762	0.481
D (ultrasonication time)	14670660	2	7335330	7.247	0.005
Error	18218808	18	1012156		

standard deviation of 2-AP content was 0.12%, indicating that the tested samples had good stability within 12 h without decomposition.

Standard addition recovery test A total of 0.4 g sample was weighed accurately and repeated for nine times. Three 2-AP solutions with three gradients of 80, 100 and 120% of 2-AP contents in the samples were added. We prepared the samples and performed chromatographic analysis according to the selected optimal chromatographic conditions and sample treatment scheme, and then calculated the 2-AP contents and recovery rates using the established standard interval. The calculated average recoveries (Rate of recovery=Actual measured 2-AP content after adding the standard substance–2-AP quality in the sample)/2-AP content in standard) of 2-AP were 96.16, 98.55 and 98.24% of the standard solutions, respectively, and the relative standard deviations (RSD) were less than 8.07%, indicating that the detection method was accurate (Table 6).

Correlation analysis of soybean seed and leaf fragrance Correlation analysis of the 2-AP contents of 30 representative soybean seeds and leaves were shown in Table 7. Kaorihime (aromatic soybean) and CM60 (non-aromatic soybean) were the control materials. The results showed a significant correlation (r=0.760) between the 2-AP contents of soybean seeds and leaves (P<0.01). Since it was more difficult to grind soybean seeds into powder than leaves when measuring a large number of samples, it is better to use leaves as samples. Therefore, we selected soybean leaves as the test materials when selecting fragrant soybean materials.

3.2. Identification of soybean germplasm with high 2-AP content flavor

Variation of 2-AP content among soybean accessions The 2-AP content of 101 soybean genotypes from different geographical sources had extensive genetic variation. The variation range was 0.094–1.816 mg L^{-1}, with an average content of 0.29 mg L^{-1}, a coefficient of variation of 0.95 and a genetic diversity index of 0.54. The results of variance analysis showed an extremely significant difference in 2-AP content among the genotypes (Table 8).

According to the grading standard, the 101 soybean samples were divided into three grades. There were seven first-grade materials. The top three varieties of 2-AP content from high to low were Zhonglong 608

Table 6 Data of 2-acetyl-1pyrroline (2-AP) plus standard recovery test

Leaf quality (g)	2-AP quality in the sample (μg)	2-AP content in standard (μg)	Actual measured 2-AP content (μg)	Rate of recovery (%)	Relative standard deviation (%)	Average recovery (%)
0.401	0.265	0.200	0.450	92.500	8.070	96.160
0.400	0.263	0.200	0.452	94.500		
0.402	0.271	0.200	0.474	101.500		
0.402	0.278	0.300	0.576	99.330	1.210	98.550
0.403	0.276	0.300	0.564	96.000		
0.402	0.266	0.300	0.576	103.330		
0.406	0.269	0.360	0.615	96.110	4.120	98.240
0.404	0.302	0.360	0.667	101.39		
0.405	0.281	0.360	0.634	97.22		

Table 7 2-Acetyl-1pyrroline (2-AP) contents of leaves and seeds for 16 representative soybean genotypes

Genotype	Leaf (mg L^{-1})	Seed (mg L^{-1})	Genotype	Leaf (mg L^{-1})	Seed (mg L^{-1})
Kaorihime	0.360	0.620	CM 60	0.094	0.051
Qihuang 34	0.424	0.648	Qihuang 42	0.175	0.061
Tonghuapingdingxiang	0.441	0.503	Dahuangdou	0.172	0.053
Zhonglong 608	1.816	1.043	PI561395	0.261	0.062
Ha13-2958	1.465	0.737	Dunajka	0.184	0.053
Hongmiandou	1.213	0.535	Ruidou 1	0.182	0.050
Zhengguang 1	0.403	0.765	Zhexiandou 3	0.181	0.055
Heinong 88	0.364	0.649	Silihuang	0.150	0.063
Huangmaodou	0.610	0.533	Minquanniumaohuang	0.150	0.055
Xihuangdou 9	0.396	0.160	Zhongkemaodou 2	0.154	0.064
PI561395	0.261	0.062	Baogongdou	0.158	0.058
Dunajka	0.184	0.053	Youhuangdou	0.159	0.054
Zexiandou 3	0.181	0.055	Qihuang 39	0.170	0.070
Heijiazi	0.234	0.543	Baoguhuang-8	0.172	0.055
Liushiribaidou	0.361	0.150	Jiuyuebaimao	0.149	0.052

(1.816 mg L^{-1}), Heinong 88 (1.510 mg L^{-1}) and Ha13-2958 (1.466 mg L^{-1}); followed by Hongmiandou (1.213 mg L^{-1}) from Shandong Province, Heinong 82 (0.913 mg L^{-1}), Huangmaodou (0.610 mg L^{-1}) from Guizhou Province and Jiyu 21 (0.571 mg L^{-1}) from Jilin Province. There were 12 second-grade materials. The top three varieties by 2-AP content were Ninglingtian'edan (0.516 mg L^{-1}) from Henan Province, Qihuang 34 (0.424 mg L^{-1}) from Shandong Province, Tonghuapingdingxiang (0.441 mg L^{-1}) from Jilin Province; followed by Jiyu 701 (0.427 mg L^{-1}) from Jilin Province, Xixianpingdingdadou (0.424 mg L^{-1}) from Henan Province, Zhonghuang 20 (0.423 mg L^{-1}) from Beijing and Zhengguang 1 (0.403 mg L^{-1}) from Jilin Province, Xihuangdou 9 (0.396 mg L^{-1}) from Guizhou, Fushengdou (0.362 mg L^{-1}) from Zhejiang, Liushiribaidou (0.361 mg L^{-1}) from Shanxi Province, Kaorihime (0.360 mg L^{-1}) from Japan, and Jiuyuehan (0.335 mg L^{-1}) from Shaanxi Province. The remaining 82 accessions were third-grade materials, and the lowest content was Silihuang (0.150 mg L^{-1}) from Heilongjiang Province.

Genotyping of soybean accessions with the marker Gm2AP in the *BADH2* gene The genotypes of the 101 soybean accessions were analyzed with marker Gm2AP that reported by Arikit *et al*. (2011) (Fig. 6). The results showed only one accession Dahuangdou having the same genotypy as the positive control of Kaorihime. This was also confirmed by sequencing the targeted fragment of *BADH2* gene which showed that there was the 2-bp deletion on exon 10 and there's no other SNP or InDel. However, the 2-AP content was only 0.172 mg L^{-1} which was second grade, indicating that the accessions with high content of 2-AP were not controlled by targeted allele of 2-bp deletion. There must have novel gene or allele controlled the high content of 2-AP in this study.

4. Discussion

4.1. The 2-AP determination method is rapid, cheap, convenient, and less toxic for screening fragrant soybeans

Based on the peak area and peak shape, the optimum analysis conditions for the determination of 2-AP were determined by single cross and orthogonal experiments. This method revealed several characteristics: (1) The peak time of 2-AP in this study was about 7 min, which was 5.5 min shorter than that in the methods used by Ariki *et al*. (2011), so it is conducive to the mass screening of 2-AP. (2) The reagents used in this method were only chromatographic grade alcohol, while other solid phase microextraction methods use at least two reagents, such as alcohol and dichloromethane (Ying *et al*. 2010), or toxic reagents, such as ethylether (Zhang *et al*. 2009), so this method involved fewer extraction reagents and less toxicity. (3) Simple and easy to operate: Huang *et al*. (2012) used distillation and extraction devices to obtain 2-AP and other researchers used methods of adding internal standards to control for the stability of the instrument (Ariki *et al*. 2011;

Table 8 Analysis of variance for 2-acetyl-1pyrroline (2-AP) of soybean accessions

Source of variation	*df*	ANOVO SS	MS	*F*-value	*Pr>F*
Genotype	98	6.107	3.054	196.081	<0.00010
Error	2	1.526	0.016		
Total	100	7.633			

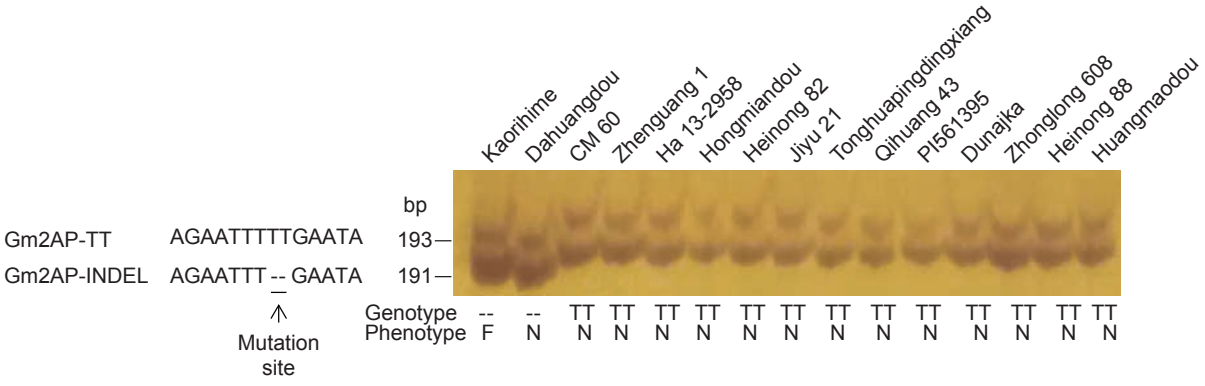

Fig. 6 Genotyping soybean accessions with the marker Gm2AP in the *BADH2* gene by PAGE analysis using Kaorihime as a positive control. The sequence variation, a 2-bp deletion, on chromosome 10 and expected sizes of the PCR products are provided. F, fragrant; N, non-fragrant.

Juwattanasomran *et al*. 2011). In this study, low 2-AP concentrations can be monitored without distillation and extraction devices, and the stability of the instrument can be controlled by only 1 mg L^{-1} standard and empty bottles or by adding a repeated control every 10 sample bottles, without internal standard, thus greatly reducing the test cost. (4) The results were accurate. In this study, the GC-MS instrument is used to evaluate the aroma grade by measuring 2-AP content, which is more accurate and objective than the subjective evaluation of traditional odor and KOH soaking methods (Wanchana *et al*. 2005). Therefore, the method described in this study is not only rapid, but also the reagents used are less toxic, it is simple and easy to operate, and the results are accurate. Therefore, it is an ideal technique for screening large samples of elite genotypes to find those with higher 2-AP content. However, it should be noted that the 2-AP content of soybean leaves does not fully represent the 2-AP content of seeds. So, while we can carry out non-destructive screening by measuring leaves, we still need to test seeds to confirm the 2-AP levels.

4.2. 2-AP content of tested genotypes has a wide range of variation

In recent years, the research on soybean quality has not been limited to the contents of protein and fat. The appearance, flavor, texture, sweetness, and other quality indexes of soybeans have been put on the agenda and have become the research hotspots. Soybeans, especially vegetable soybeans, have a good taste, large seeds, high soluble content, unique flavor characteristics, and other important qualities. As an important index of flavor components, the 2-AP content of the tested soybeans was accurately identified by GC-MS. In our study, the highest 2-AP content was found in Zhonglong 608 (1.816 mg L^{-1}), which was nearly three times higher than that in Kaorihime (0.620 mg L^{-1}). In addition, five other soybean genotypes with 2-AP contents higher than 0.6 ppm were identified: Heinong 88, Ha13-2958, Heinong 82, Huangmaodou, and Jiyu 21. 2-AP content variations may be due to different experiment materials, genotypes, environmental factors, the different treatment methods of the materials etc., or because the fragrance character has been selected artificially through artificial breeding. This study provides materials and methods to support soybean flavor breeding in China. There was a wide range of genetic variation in the 2-AP content among different soybean varieties, which showed that some soybean germplasms may have extremely high 2-AP content and can be used for high 2-AP breeding. There were significant differences in 2-AP content among different ecological types of soybeans. The content of 2-AP in the northern region was significantly higher than that in the southern region, indicating that climatic conditions also had an impact on the accumulation of 2-AP in soybeans. Therefore, it is very important to obtain accurate and reliable 2-AP content levels under various environmental conditions.

4.3. New alleles contribute to elite accessions with high 2-AP content

It was reported that the increase of 2-AP content in aromatic soybean varieties was related to a 2-bp deletion of 928 nucleotide (TT) in exon 10 for *BADH2* gene (Arikit *et al*. 2010). This mutation leads to the mutation and early termination of BADH2 which resulted shortening length of amino acids from 448 to 309. Despite one accession of Dahuangdou had this allele, it had pretty lower 2-AP content. All accessions with high 2-AP content provide useful materials for discovering novel genes or alleles.

5. Conclusion

In this study, a rapid and convenient method for the accurate identification of 2-AP content in soybean was established. Through a single factor test, the best parameters of the instrument operation are column temperature at 70°C, and injection port temperature at 180°C. Through the analysis of the range and variance of the orthogonal test design, the optimal extraction conditions are 1 mL of alcohol, 0.1 g of NaCl, 10 min of ultrasound time, and 1 h of extraction time. The influence of the factors on the determination of 2-AP content is in the order of: alcohol content>NaCl content>ultrasound time>extraction time. The extraction time had no significant effect on 2-AP content, indicating that 2-AP can be extracted quickly. Under the optimal conditions, the method is rapid, cheap and convenient, and has good reproducibility, indicating that this method can be used for the accurate and rapid determination of 2-AP content and the mass screening of flavor in soybeans.

A total of 101 soybean genotypes from different geographical sources were tested for 2-AP content and divided into three grades of flavor type, among which seven genotypes with grade-one flavor type could be used as elite resources for soybean aroma breeding and gene discovery.

This study provides a standard identification method of aroma, useful not only for soybean aroma breeding, but also for other crops.

Acknowledgements

This work was supported by the National Key R&D Program for Crop Breeding (2016YFD0100201) and Scientific Innovation Program of of the Chinese Academy of Agricultural Sciences (CAAS). We thank the Core

Facility Platform, Institute of Crop Sciences, CAAS for their assistance with GC-MS analyses.

Declaration of competing interest

The authors declare that they have no conflict of interest.

References

Adams A, De K N. 2007. Formation of pyrazines and 2-acetyl-1-pyrroline by *Bacillus cereus*. *Food Chemistry*, **101**, 1230–1238.

Arikit S, Yoshihashi T, Wanchana S, Tanya P, Juwattanasomran R, Srinives P, Vanavichit A. 2011. A PCR-based marker for a locus conferring aroma in vegetable soybean (G*lycine max* L.). *Theoretical and Applied Genetics*, **122**, 311–316.

Arikit S, Yoshihashi T, Wanchana S, Uyen T T. 2010. Deficiency in the amino aldehyde dehydrogenase encoded by *GmAMADH2*, the homologue of rice *Os2-AP*, enhances 2-acetyl-1-pyrroline biosynthesis in soybeans (*Glycine max* L.). *Plant Biotechnology Journal*, **122**, 311–316.

Brahmachary R L, Sarkar P, Dutta J. 1990. The aroma of rice and tiger. *Nature*, **344**, 26.

Buttery R G, Ling L C, Juliano O B, Turnbaugh J G. 1983. Cooked rice aroma and 2-acetyl-1-pyrroline. *Journal of Agricultural and Food Chemistry*, **31**, 823–826.

Fushimi T, Masuda R. 2001. 2-acetyl-1-pyrroline concentration of the vegetable soybean. In: Lumpkin T, Shanmugasundaram S, eds., *Proceeding of the 2nd International Vegetable Soybean Conference*. Washington State University, Pullman. p. 39.

Huang Z L, Tang X R, Wang Y L, Chen M J, Zhao Z K, Duan M Y, Pan S G. 2012. Effects of increasing aroma cultivation on aroma and grain yield of aromatic rice and their mechanism. *Scientia Agricultura Sinica*, **45**, 1054–1065. (in Chinese)

Juwattanasomran R, Somta P, Chankaew S. 2011. A SNP in *GmBADH2* gene associates with fragrance in vegetable soybean variety “Kaori”and SNAP marker development for the fragrance. *Theoretical and Applied Genetics*, **122**, 533–541.

Juwattanasomran R, Somta P, Kaga A. 2010. Identification of a new fragrance allele in soybean and development of its functional marker. *Molecular Breeding*, **29**, 13–21.

Masuda R. 1991. Quality requirement and improvement of vegetable soybean. In: Shanmugasundaram S, ed., *Vegetable Soybean*: *Research Needs for Production And Quality Improvement*. Asian Vegetable Research and Development Center, Taiwan of China. pp. 92–102.

Plonjarean S, Phutdhawong W, Siripin S, Suvannachai N, Sengpracha W. 2007. Flavour compounds of the Japanese vegetable soybean “Chakaori growing in Thailand. *Majeo International Journal of Science and Technology*, **1**, 1–9.

Reineccius G A. 2006. *Flavor Chemistry and Technology*. Taylor & Francis Group (CRC Publishers), BocaRaton, FL. pp. 391–418.

Schieberle P, Grosch W. 1991. Potent odorants of the wheat bread crumb differences to the crust and effect of a longer dough fermentation. *Zeitschrift Für Lebensmittel Untersuchung Und Forschung*, **192**, 130–135. (in German)

Snowdon E M, Bowyer M C, Grbin P R, Bowyer P K. 2006. Mousy off-flavor: A review. *Journal of Agricultural and Food Chemistry*, **54**, 6465–6474.

Sriseadka T, Wongpornchai S, Kitsawatpaiboon P. 2006. Rapid method for quantitative analysis of the aroma impact compound, 2-acetyl-1-pyrroline, in fragrant rice using automated headspace gas chromatography. *Journal of Agricultural and Food Chemistry*, **54**, 8183–8189.

Wanchana S, Kamolsukyunyoung W, Ruengphayak S, Tragoonrung S, Toojinda T, Vanavichit A. 2005. A rapid construction of a physical contig across a 4.5 cM region for rice grain aroma facilitates marker enrichment for positional cloning. *ScienceAsia*, **31**, 299–306.

Widjaja R, Craske J and Wootton M. 1996. Comparative studies on volatile components of non-fragrant and fragrant rices. *Journal of Science of Food and Agriculture*, **70**, 151–161.

Wu M L, Chou K L, Wu C R, Chen J K, Huang T C. 2009. Characterization and the possible formation mechanism of 2-acetyl-1pyrroline in aromatic vegetable soybean (*Glycine max* L.). *Journal of Food Science*, **74**, 192–197.

Ying X H, Xu X, Chen M X, OuYang Y N, Zhu Z W, Min J. 2010. Determination of 2-acetyl-1-pyrroline in aroma rice using gas chromatography-mass spectrometry. *Chinese Journal of Chromatography*, **28**, 782–785. (in Chinese)

Yoshihashi T. 2002. Quantitative analysis on 2-acetyl-1-pyrroline of an aromatic rice by stable isotope dilution method and model studies on its formation during cooking. *Journal of Food Science*, **67**, 619–622.

Yoshihashi T, Huong N T T, Inatomi H. 2002. Precursors of 2-acetyl1-pyrroline, a potent flavor compound of an aromatic rice variety. *Journal of Agricultural Food Chemistry*, **50**, 2001–2004.

Zhang Y B, Miao C Q, Cui J J. 2009. Comparative analysis of volatile oil from lotus leaves by purge and trap-thermal desorption gas chromatography-mass spectrometry. *Acta Chimica Sinica*, **67**, 2368–2374. (in Chinese)