



OPEN

Analysis of volatile compounds by GCMS reveals their rice cultivars

Shengying Hu^{1,2,4}, Hongbo Ren³, Yong Song^{1,2}, Feng Liu³, Lili Qian⁵, Feng Zuo⁵ & Li Meng^{1,2}✉

Due to the similarity in the grain and difference in the market value among many rice varieties, deliberate mislabeling and adulteration has become a serious problem. To check the authenticity, we aimed to discriminate rice varieties based on their volatile organic compounds (VOCs) composition by headspace solid phase microextraction (HS-SPME) coupled with gas chromatography mass spectrometry (GC-MS). The VOC profiles of Wuyoudao 4 from nine sites in Wuchang were compared to 11 rice cultivar from other regions. Multivariate analysis and unsupervised clustering showed an unambiguous distinction between Wuchang rice and non-Wuchang rice. Partial least squares discriminant analysis (PLS-DA) demonstrated a goodness of fit of 0.90 and a goodness of prediction of 0.85. The discriminating ability of volatile compounds is also supported by Random forest analysis. Our data revealed eight biomarkers including 2-acetyl-1-pyrroline (2-AP) that can be used for variation identification. Taken together, the current method can readily distinguish Wuchang rice from other varieties which it holds great potential in checking the authenticity of rice.

Rice (*Oryza sativa* L.) is the staple food for approximately 3.5 billion people worldwide with an estimated production of 480 million tons annually¹. Rice provides up to 50% of the dietary caloric supply and a substantial part of the protein intake for about 520 million people living in poverty in Asia². The market value of rice is primarily determined by the grain quality, which in turn depends on the genetic background, agronomic management, geographic factors and postharvest factors^{3,4}. Although the grain quality and thus the price for rice of different geographic origins could differ vastly, their appearance may be very alike. Thus, deliberate mislabeling and adulteration represent a serious issue, which not only threatens the credit of the traders, but also infringes the right of consumers⁵. To prevent such actions, a high-throughput, sensitive and precise method is urgently needed to discriminate rice of different geographic origins. The key to determining the geographic origin was to identify specific biomarkers for a specific rice variety. Previous studies had used a wide range of properties such as stable isotope, mineral element content that are associated with geographic features to track the origin of rice^{6,7}. Notably, multiple parameters can be used together as they may provide complementary information. For instance, the isotope composition of $\delta^{13}\text{C}$, $\delta^{15}\text{N}$, $\delta^2\text{H}$, $\delta^{18}\text{O}$ had been employed in junction with the multi-elemental concentrations analysis to determine the origin of different rice varieties⁸. For a more comprehensive comparison of the components among various rice cultivars, spectroscopy approaches such as near-infrared reflectance spectroscopy (NIR), nuclear magnetic resonance spectroscopy (NMR), Raman spectroscopy (RS) had also been used^{9–11}. These methods requires large amount of input for data collection and has no direct relationship with the quality of rice.

Volatile organic compounds (VOCs) are the main source of aroma in rice. Although difficult to describe by words, aroma differ greatly among varieties due to a different composition in VOCs. Indeed, a previous study has reported the geographical discrimination of rice based on VOCs¹². Headspace solid phase microextraction (HS-SPME) was applied in volatile substance concentrate to improve the analytes placed on column¹³. Due to its simplicity, high sensitivity and reproducibility, SPME is particularly suitable for the analysis of volatiles in food^{14,15}. Coupled with GC-MS, the aroma compounds can be identified precisely.

Different rice varieties have different aromatic characteristics and volatiles of rice can be considered to be biomarker candidates of geographical varieties determination. Rice considered to be has high quality because of the attractive aroma and good taste¹⁶. So, to characterize the aroma profile of fragrant rice is meaningful not only on adulteration but also good for breeding. Wuchang rice, as its name suggests, grows in Wuchang county

¹Engineering Research Center of Agricultural Microbiology Technology, Ministry of Education, Heilongjiang University, Harbin 150500, China. ²Key Laboratory of Molecular Biology of Heilongjiang Province, College of Life Science, Heilongjiang University, Harbin 150080, China. ³Quality and Safety Institute of Agricultural Products, Heilongjiang Academy of Agricultural Sciences, Harbin 150086, China. ⁴Shandong Yanggu Huetai Chemical Co., Ltd., Shandong 252300, China. ⁵College of Food Science and Technology, Heilongjiang Bayi Agricultural University, Daqing 163319, China. ✉email: mengli198026@126.com

in Heilongjiang province of northeast China¹⁷. It is known for its intense aroma and enlisted as a National Geographical Indication Product of China¹⁸. Wuyoudao 4 is the dominant variety in this area since 1990s. Despite big differences in composition and the grain quality between WuChang rice and other varieties, it is still challenging to distinguish them due to a lack of robust and sensitive method^{19–21}.

In this study, we aimed to address this challenge by profiling Wuchang rice VOCs with HS-SPME extraction and GC–MS detection. First, an untargeted metabolomic approach was used to profile the VOCs in both WuChang and non-WuChang rice varieties. Subsequent multivariate analysis showed an unambiguous classification of the samples according to their geographic origins. In addition, our data led to the identification of eight VOCs that can be used as biomarkers in variety determination. The GCMS approach described is readily applicable to determine the geographic origin of other premium rice.

Materials and methods

Material. Twenty rice samples grown in 2018 were purchased from either Wuchang or other regions (see Table 1 for details). It was confirmed that all methods were performed in accordance with the relevant legislation. For each sample group, three biological replicates were obtained. All samples were stored at $-80\text{ }^{\circ}\text{C}$ until analysis. Headspace vials (20 ml in volume) were purchased from Agilent, each with a silver aluminum cap and a polytetrafluoroethylene (PTFE)/silicone rubber septum. Supplies for solid phase microextraction (SPME) including a manual injection handle and a $75\text{ }\mu\text{m}$ carboxen/polydimethylsiloxane (CAR/PDMS) extraction head were purchased from SUPELCO (St. Louis, USA)²².

Extraction of volatiles by HS-SPME. Four grams of rice grain were stored in a 20 mL headspace vial. The headspace bottle was heated at $80\text{ }^{\circ}\text{C}$ for 30 min in a water bath. Subsequently, a SPME fiber was inserted into the headspace portion of the vial and then incubated for 60 min to absorb VOCs from the rice grain²³.

GC–MS analysis. GC–MS analysis was performed using a TRACE GC ULTRA (Thermo) coupled with a SQ QUANTUM XLS (Thermo). The SPME fiber was inserted into the inlet of the GC–MS with the fiber head pushed out, and VOCs were desorbed at $250\text{ }^{\circ}\text{C}$ for 5 min. Helium was used as the carrier gas with a flow rate of 1 mL/min and a splitless injection was used. A DB-5 capillary column ($30\text{ m}\times 0.25\text{ mm}\times 0.25\text{ }\mu\text{m}$) was used for separation with the following temperature gradient: an initial temperature of $40\text{ }^{\circ}\text{C}$ for 3 min; a linear increase to $100\text{ }^{\circ}\text{C}$ at $5\text{ }^{\circ}\text{C}/\text{min}$ and hold for 3 min; and ramping to $250\text{ }^{\circ}\text{C}$ at $10\text{ }^{\circ}\text{C}/\text{min}$ and hold for 4 min. Eluted compounds are ionized by electron ionization with an electron energy of 70 eV and an ion source temperature of $230\text{ }^{\circ}\text{C}$. Full MS scans were performed at the mass range of 40–300 m/z ²⁴.

Chromatographic analyses. The obtained MS spectra were matched to the National Institute of Standards and Technology reference spectra (NIST 08 and Wiley 7). The retention index (RI) of putative compounds is further compared to previous reports (Table S2). The normalized peak area was used for calculating the compound concentration. Data from three biological replicates were reported (Table S3).

Sample	Origin	Number	Cultivar
Wuchang rice	Changbao	CB	Wuyoudao 4
	Yingchengzi	YCZ	Wuyoudao 4
	Longfengshan	LFS	Wuyoudao 4
	Lalin	LL	Wuyoudao 4
	Minle	ML1	Wuyoudao 4
	Minle	ML2	Wuyoudao 4
	Anjia	AJ1	Wuyoudao 4
	Anjia	AJ2	Wuyoudao 4
	Xiangyang	XY	Wuyoudao 4
	Non-wuchang rice	Jiansanjiang	JSJ
Zhaoyuan		ZY	Songjing 22
Xiangshui		XS	Shangyu 397
Jilin		JL	Jijing 83
Panjin		PJ	Liaojing 9
Sheyang		SY	Huaidao 5
Hubei		HB	Liangyoupeijiu
Jiangxi		JX	Jinyou 463
Henan		HN	Yuandao 109
Korea		HG	Dongjin
Thailand	TG	Jasmine 105	

Table 1. The origins of the 20 rice samples.

Clustering analysis. Prior to statistical analysis, the raw MS intensities were first subjected to log transformation and Pareto scaling. Feature selection was performed using standard methods such as fold change (FC), volcano plot, t-test and empirical Bayesian analysis of microarray (EBAM). Features with a relative standard deviations (RSD) > 30% were removed from further analysis. A feature was considered to be differentially expressed with the following criteria: (1) $FC < 0.8$ or $FC > 1.2$; (2) $p\text{-value} < 0.05$ and (3) false discovery rate (FDR) < 0.1.

Euclidean distance-based hierarchical cluster analysis (HCA) was used to explore the variance distribution and classification patterns among samples with the Ward method²⁵. Discriminant analysis of the rice samples was performed with principal component analysis (PCA), partial least square-differential analysis (PLS-DA), cross validation and Heatmap. Further evaluation of the accuracy of these predictions were performed by random forest analysis²⁶. Finally, putative biomarker candidates to discriminate Wuchang rice from other varieties were chosen based on (1) a variable importance in the projection (VIP) score ≥ 0.8 ; and (2) differentially expressed in the univariate analysis²⁷. All the data were analyzed were conducted by R 4.2.2.

Results and discussion

The shape and VOCs profile of the rice grain. The morphology of rice showed in Fig. 1. Some of the rice have different shapes with Wuyoudao4, but some cultivar has similar length and shape such as JSJ. Although Wuyoudao4 have the characteristics with long grains, but it is also hard to tell whether it is Wuyoudao4 only depend on the shapes by naked eye. To prevent mislabeling and adulteration, we developed a method to profile VOCs from Wuchang rice and non-Wuchang rice. The representative total ion chromatography was shown Fig. 2 with the peaks annotated. In total, 22 VOCs were identified from our samples. These VOCs covered compounds with a diverse range of chemical compositions including eight aldehydes, five alcohols, two ketones, two heterocyclic compounds, and five hydrocarbon compounds (Table 2). In addition to VOCs, siloxane derivatives were also observed due to the presence of PDMS in the SPME fibers (Table S2), which is consistency with the previous report^{28,29}.

Chemometric analysis of biomarkers for rice screening. To identify putative biomarker in determining the geographic origin of rice, we performed both single factor and multivariate analyses on the VOC dataset. First, differential analysis revealed that the concentration of eight features were significantly different among samples ($p < 0.05$, $FDR < 0.1$, and $FC < 0.8$ or $FC > 1.2$) showed in Venn diagram (Fig. 2). This indicates that the profile of VOCs could be used to distinct rice varieties.

The eight features showing differential concentration between Wuchang and non-Wuchang rice were further evaluated by the variable importance in the projection (VIP) score. All of them passed the 0.8 cutoff and were thus considered as putative biomarkers. They consisted of three aldehydes (heptanal, octanal, (E)-2-decenal), 3 alcohols (1-heptanol, (E)-2-decen-1-ol, 3,7,11-Trimethyl-3-dodecanol), one ketone (3-octene-2-one), and one heterocyclic compound (2-acetyl-1-pyrroline, or 2-AP). Details of the these compounds were summarized in Table S1. Compared to non-Wuchang rice, Wuchang rice showed a significantly lower level of 1-heptanol and 3-octene-2-one ($P < 0.05$), and a much higher level for the remaining six putative biomarkers (Fig. 3a). Among

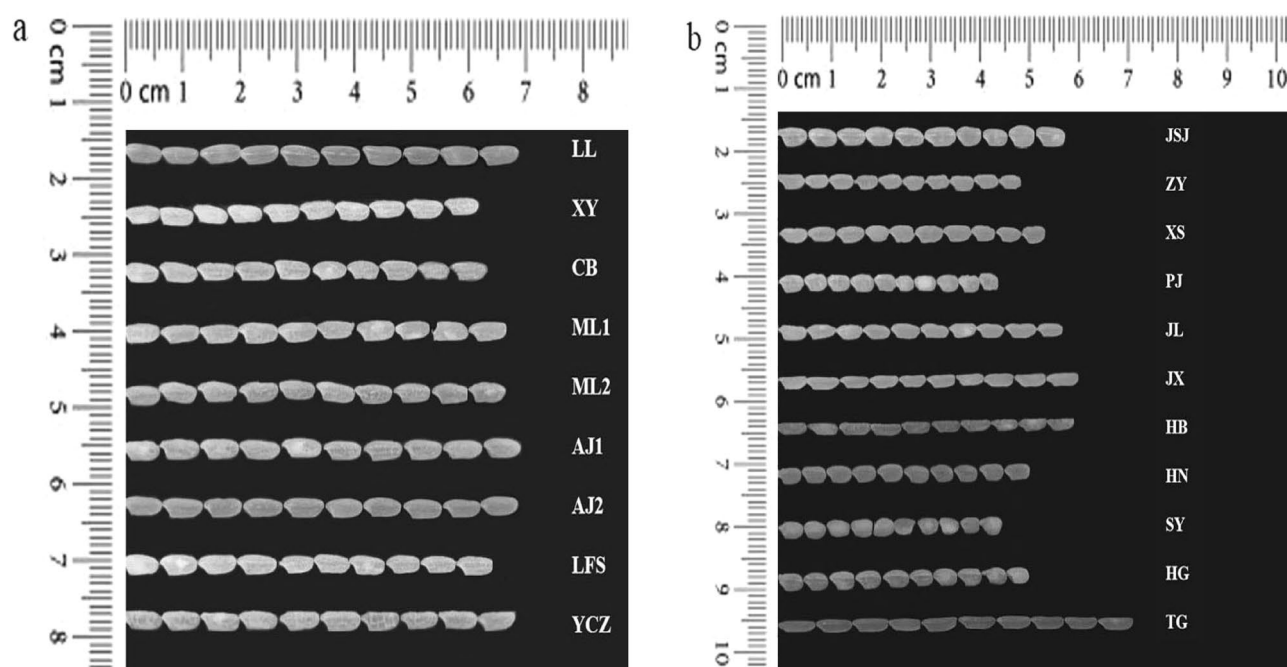


Figure 1. The morphological characteristics of difference rice grains. (a) The wuyoudao4 grown in 11 sites of Wuchang (b) The cultivars from Non-wuchang origins.

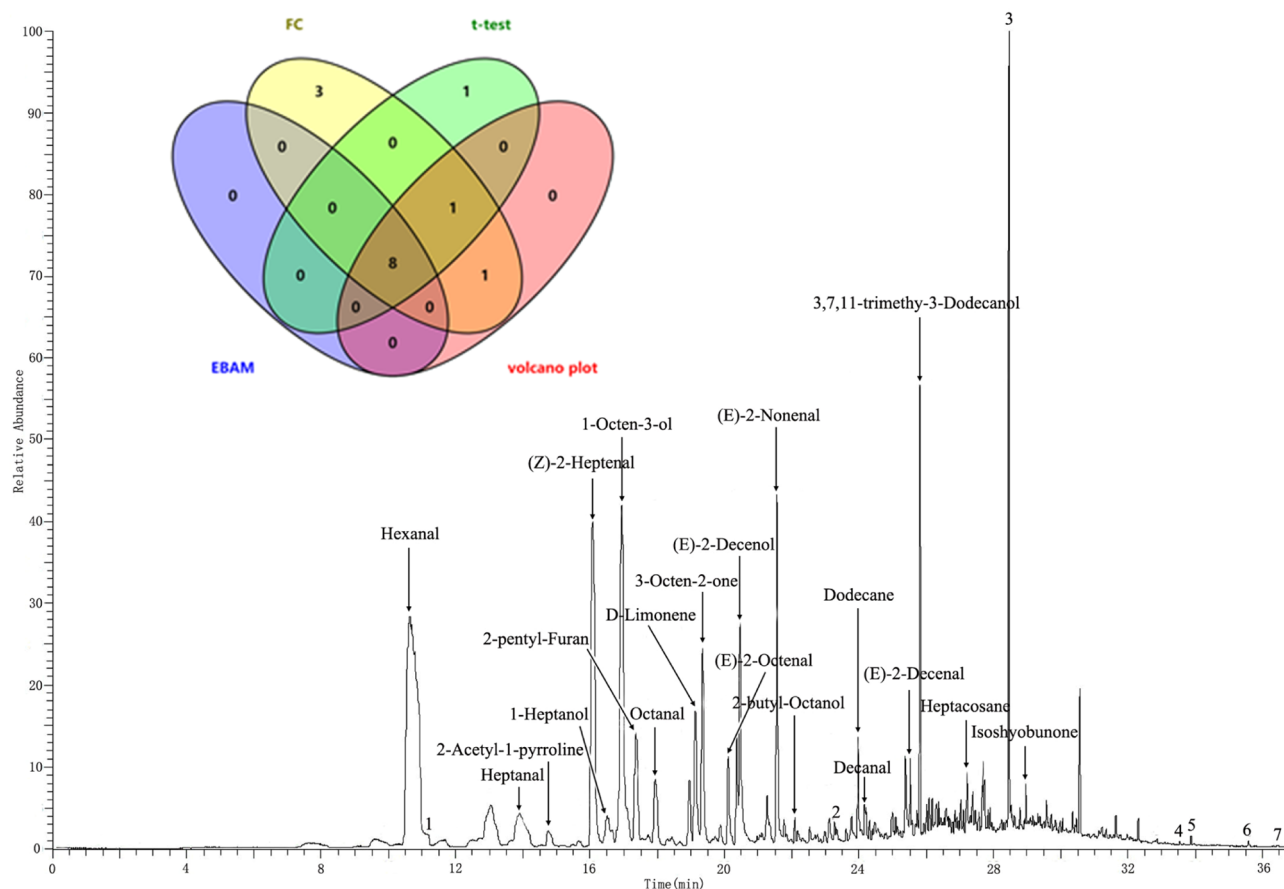


Figure 2. VOCs profiling of Wu Chang rice obtained by GC–MS using CAR/PDMS SPME.

Retention time (min)	Compound name	Formula	NIST match (%)	Retention index (RI) ^a			VIP Score	T-test		FC
				Experiment	Reference	RSD (%)		P-value	FDR	
14.06	Heptanal	C ₇ H ₁₄ O	93	839	844	28.93	2.33	0.032	0.035	2.33
14.76	Acetyl-1-pyrroline	C ₆ H ₉ NO	96	897	901	5.22	3.41	0.0002	0.004	8.69
16.53	1-Heptanol	C ₇ H ₁₆ O	92	739	754	27.97	0.99	0.028	0.017	0.76
17.98	Octanal	C ₈ H ₁₆ O	93	907	941	5.81	0.89	0.043	0.037	1.25
19.35	3-Octen-2-one	C ₈ H ₁₄ O	92	800	810	27.96	0.98	0.049	0.034	0.59
20.38	(E)-2-Decen-1-ol	C ₁₀ H ₂₀ O	90	778	790	10.36	1.34	0.016	0.013	1.47
25.42	(E)-2-Decenal	C ₁₀ H ₁₈ O	91	804	856	28.35	0.88	0.028	0.024	1.24
25.82	3-Dodecanol,3,7,11-trimethyl-	C ₁₅ H ₃₂ O	90	780	817	19.15	0.89	0.036	0.035	1.22

Table 2. 8 typical biomarkers characteristics of Wuchang rice. ^aRI indicates retention index. Reference RI values were obtained from <https://www.nist.gov/>.

them, 2-AP has been reported as an important chemical trait to differentiate fragrant rice from non-fragrant rice. Other non-wuchang rice have a relative lower amount of 2-AP amount except TG (Jasmine 105). The identification of 2-AP (Fig. 3b) was validated which the robustness of our method, but also aligned well with previous findings.

Despite we observed a significant difference in the 2-AP concentration between Wuchang and non-Wuchang rice, we next sought to see whether other volatiles show a similar pattern. Thus, we examined the correlations between 2-AP and other compounds. We observed a negative correlation for (E)-2-octenal, (Z)-heptenal, 1-octene-3-ol, dodecane, and heptadecane, and a positive correlation for the rest of 22 VOCs (Fig. 3c).

Evaluation of biomarker efficiency. Previous studies reported the use of 2-AP in determining fragrant rice and non-fragrant rice³⁰. However, the usefulness of other biomarkers remain unclear. To this end, we assessed the discrimination efficiency of putative markers identified here using several multivariate analysis techniques.

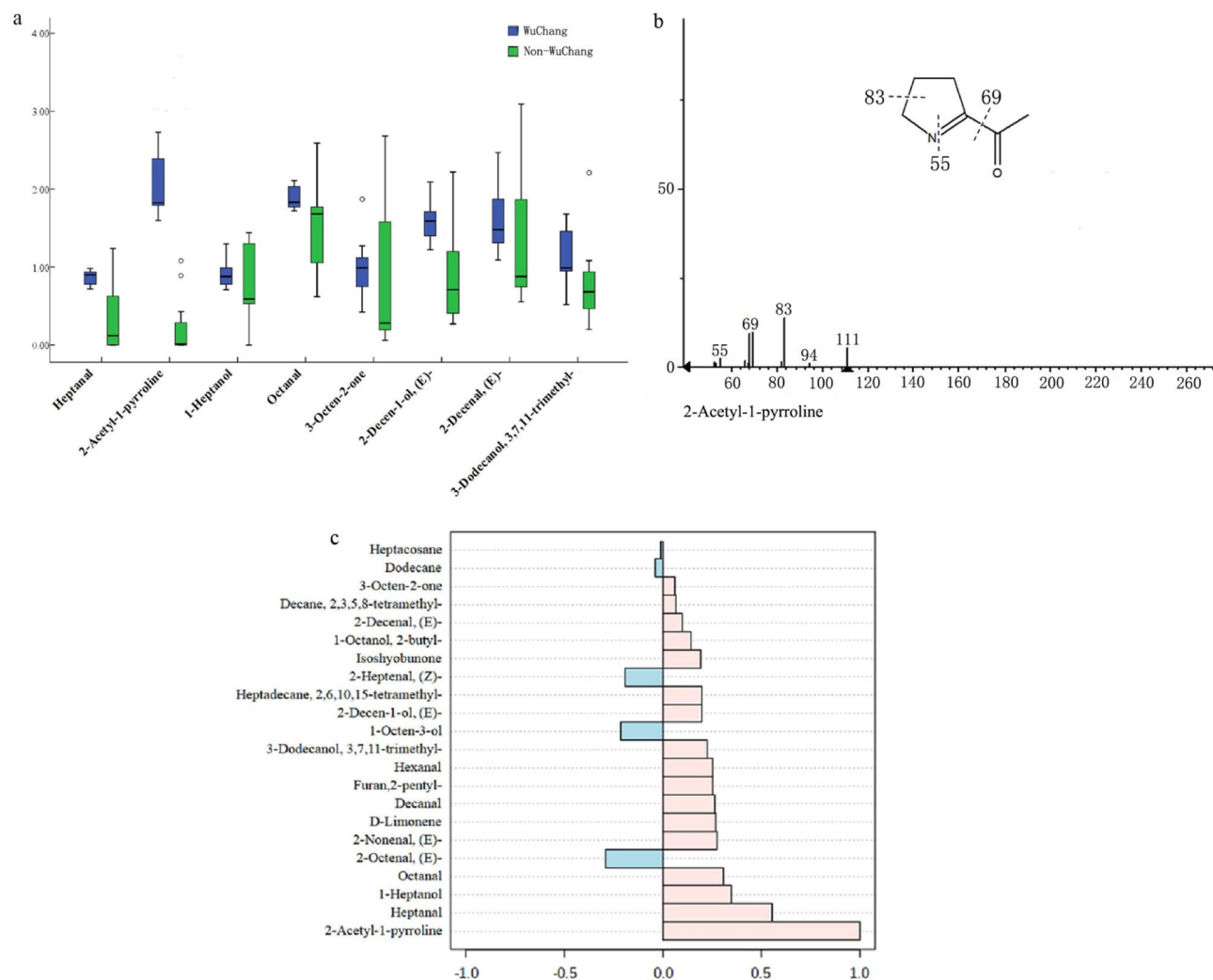


Figure 3. Box plots and 2-AP analysis of eight chemical biomarkers in rice samples. **(a)** Box diagram of eight typical biomarkers for rice samples from WuChang and other regions; **(b)** 2-AP ion map; **(c)** Correlation analysis of 2-AP with other volatile compounds.

Our results showed that a clear separation of rice samples by the VOC data. Finally, random forest (RF) analysis further revealed a high accuracy of these markers in distinguishing rice samples from different sources.

Clustering and heat map analysis. HCA clustering classified the samples into two groups neatly with respective to their geographic origins: Wuchang rice and non-Wuchang (Fig. 4a). A large distance between the two groups was also observed (Euclidean distance > 20), supporting the significant difference in grain quality between them. By contrast, the Euclidean distances among the 8 Wuchang samples are less than 5, suggesting that subtle variations exist for rice grown in the same area.

Next, heatmap analysis was performed to visualize the correlation of eight putative VOCs (Fig. 4b). The 20 samples were separated into two groups. A high degree of correlation was found for the first group consisted of 9 samples, including all 8 varieties from Wuchang. Notably, samples in this group were also high in 2-AP, contribution to a better rice flavor. By contrast, a lower level of correlation in term of VOC content was found for the second group of 11 samples.

Unsupervised PCA was used to process the original data into orthogonal components. The results showed a percentage of variance of 33.9%, 26.8% and 11.1% for the first three PCs respectively (Fig. 5a). Consistent with the clustering analysis, a clear separation of Wuchang rice from other varieties was also observed using the first three PCs. It is notable that nine samples of Wuchang rice concentrated in the three-dimensional PCA plot, supporting that no statistical significant difference were found among them ($p > 0.05$). Conversely, a wide distribution was observed for the 11 other rice samples, reflecting the fact that they were grown in a diverse range of geographic locations.

Similar to clustering and PCA, the PLS-DA analysis showed a clear aggregation among Wuchang varieties (Fig. 5b), indicating that the composition and content of these nine samples were similar. By contrast, disperse

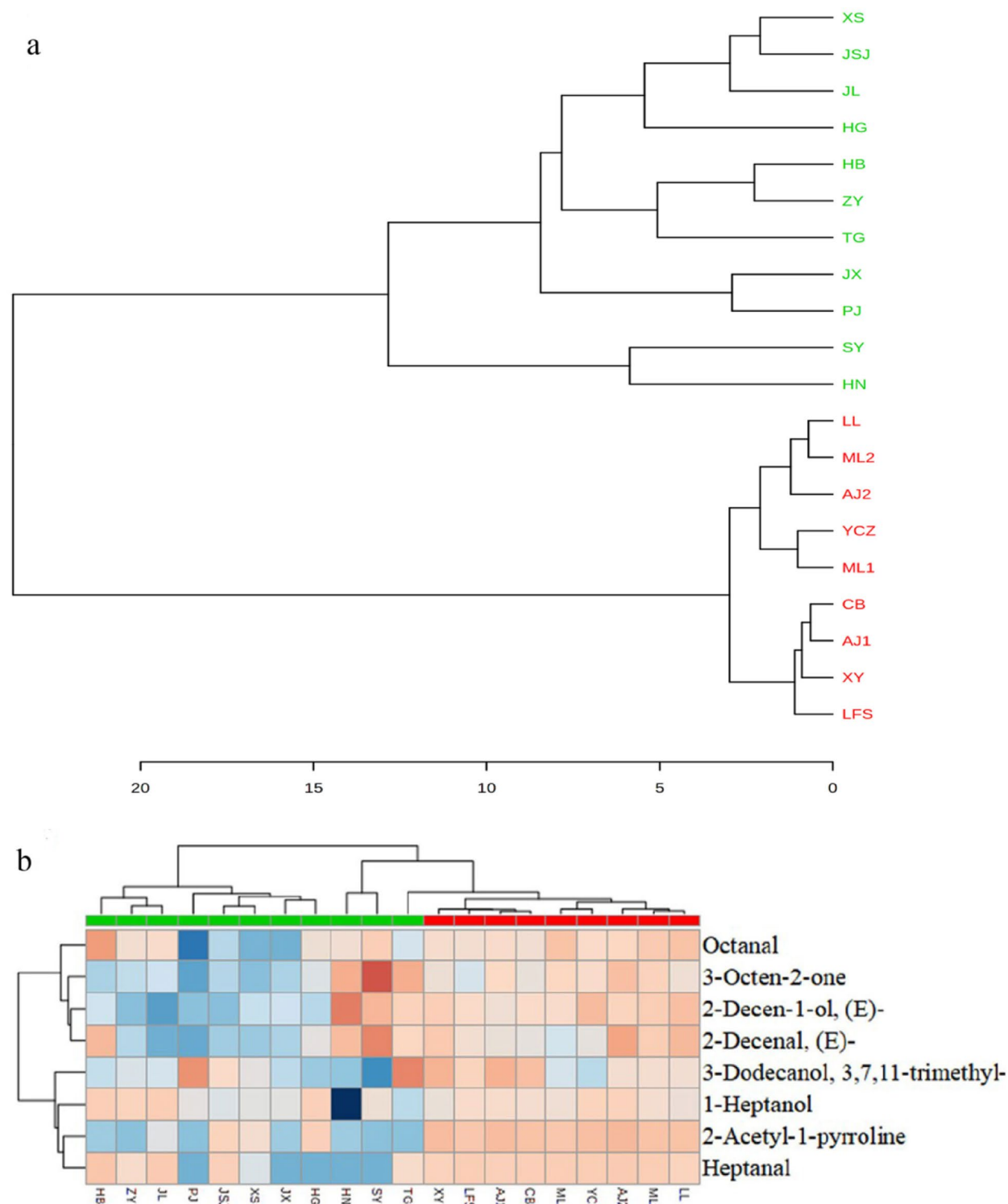


Figure 4. Clustering analysis of rice from different regions using eight VOC typical biomarkers. **(a)** Dendrogram **(b)** heatmap, weak blue and white represent weak correlation, pink and brown represent strong correlation.

of the other 11 samples suggested a large difference in composition and content among them. Significantly, the PLS-DA also revealed the discrimination capacity of the eight biomarkers with a fitness R^2 of 0.90, prediction goodness Q^2 of 0.85, and displacement test p value of 0.01.

To further evaluate the ability of the eight putative biomarkers in rice classification, we built a model based on the random forest algorithm (RF). RF is not only suitable for feature selection, but also provides useful information such as OOB (out-of-bag) error, variable importance measurement, and outlier measurement³¹.

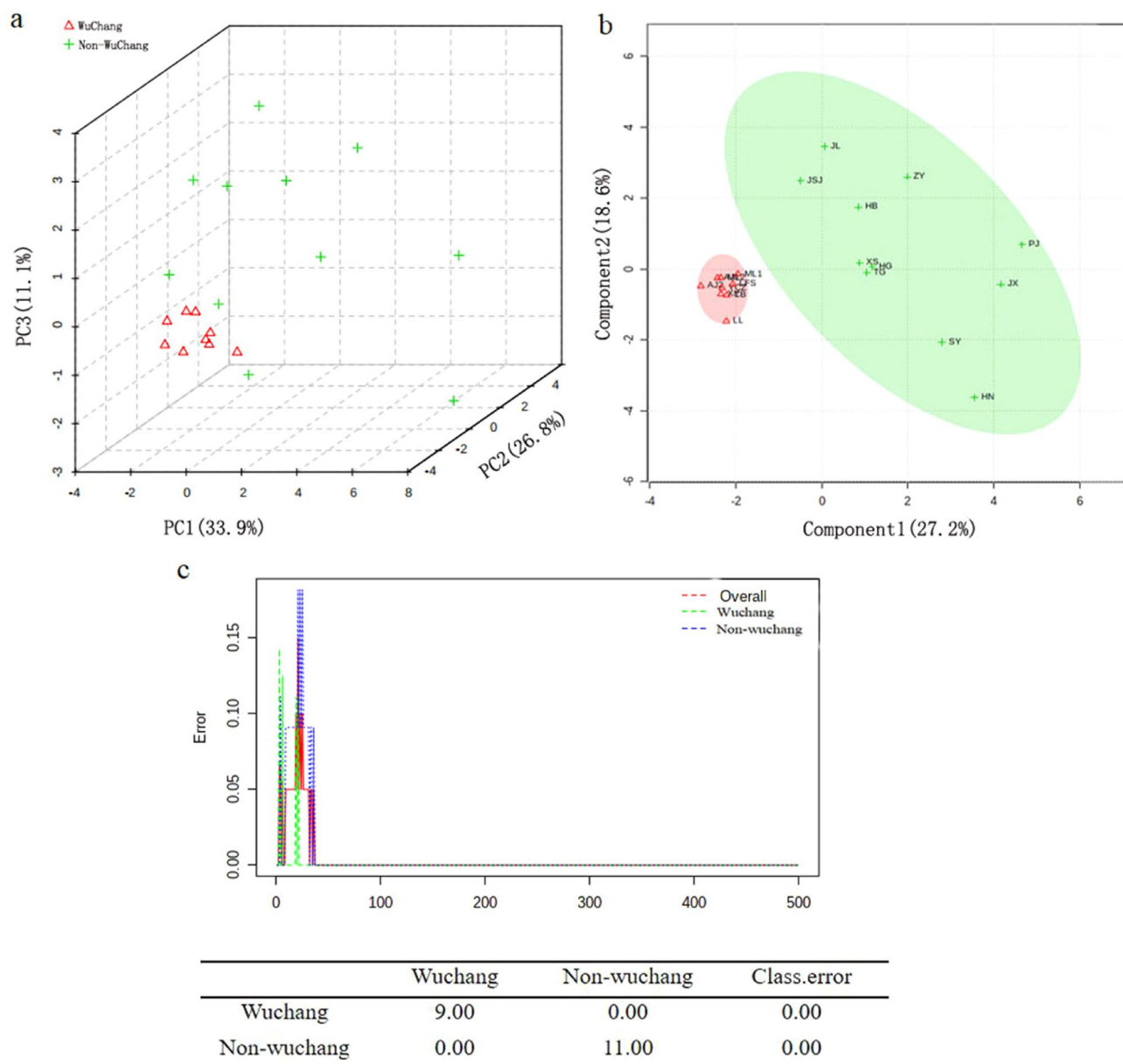


Figure 5. Different statistical analyses of the chemical biomarkers. **(a)** 3D score plot between the selected PCs in PCA. **(b)** PLS-DA models **(c)** Cumulative error rates by Random Forest classification.

We found an excellent classification ability of the eight putative biomarkers, indicated by a classification error of essentially 0 (Fig. 5c).

Our study has demonstrated the feasibility of using HS-SPME/GC-MS to profile the VOCs in rice grains. In addition, we found that Wuchang rice can be unambiguously distinguished from non-Wuchang rice by eight putative biomarkers. Thus, this work represents an important contribution to the field of non-destructive geographical authentication of rice. Application of the method in industrial settings would bring a broad impact on the production and management of Wuchang rice, commercial rice trading and consumer's satisfaction. Putative biomarkers include 2AP, consistent with its contribution to rice aroma and previous findings that 2AP can serve as a biomarker to distinguish fragrant rice from non-fragrant ones. More importantly, our untargeted metabolomic approach also identified other putative biomarkers of diverse chemical properties. In the following section, we discuss the relevance of these compound in distinguishing rice varieties.

Among the eight putative biomarkers, 2-AP is the most important source of flavor in rice, followed by aldols and ketones. As the most important aroma compound in fragrant rice, 2-AP promotes appetite and improves human metabolism^{32–34}. The odor of 2-AP at 0.05 ppm aroma was described as popcorn, which is positively correlated with “butter” and “corn. The odor threshold for 2-AP in water and air had been determined to be 0.1 nL/L, and 0.02 ng/L, respectively³⁵. The extremely low threshold makes 2-AP an important source of food aroma. In rice, 2-AP is mostly produced by the plants, although it was originally thought as a product of the Maillard

reaction during rice cooking^{36,37}. In line with this, we detected 2-AP by GCMS in the rice grain with the highest concentration observed in the Wuchang samples. These results are also in consistent with previous reports that different levels of 2-AP are detected in distinct rice varieties^{38,39}. For instance, a 10 times lower level of 2-AP was found in non-fragrant rice compared to that of fragrant rice. Together with other compounds such as hexanal, (E,E)-2,4-decanedialdehyde, (E)-2-nonanedialdehyde, 4-vinyl-2-methoxy-phenol and hydrazine, 2-AP has been identified as a common aromatic compound among three fragrant rice varieties⁴⁰. Thus, the high level of 2-AP in Wuchang rice suggests that 2-AP is a major factor in determining the pleasant flavor of Wuchang rice.

Aldehydes were odorous compounds found in many plants and foods, especially in fragrant rice. They are generated either by the oxidation of free fatty acids or the decomposition of linoleic acid⁴¹. We found three aldehydes that show differential levels between Wuchang and non-Wuchang rice: heptanal, octanal, and (E)-2-decenal. Previous studies have shown that heptanal, among other saturated aldehydes such as valeraldehyde, hexanal, heptaldehyde, octanal, and furfural, are odor-generating compounds in rice. Interestingly, heptanal and octanal give rice a pleasant fresh grassy scent and a light fruit scent at low concentration, while it produces a disgusting rancid taste if the concentration is too high. In addition, (E)-2-nonenal have a pleasant orange aroma at an extremely low concentration. It will be of interest to explore the contribution of these compounds, either alone or combined with other compounds, in rice flavor.

Three alcohols, 1-heptanol, (E)-2-decen-1-ol, 3,7,11-Trimethyl-3-dodecanol, were identified in our samples. These volatile alcohols produce a milder odor, e.g., the turf taste of 1-heptanol. Most alcohol compounds have higher thresholds, and only higher contents or a non-saturated alcohol such as (E)-2-nonenol and 3,7,11-trimethyl-3-dodecyl alcohol will have a great influence on the flavor.

Ketones can be generated from a variety of pathways in plants, including oxidative or thermal degradation of polyunsaturated fatty acids, amino acid degradation or microbial oxidation⁴². Ketones give rice a pleasant aroma. Daygon et al. reported that compounds such as 2-heptanone, 2-hexanone, and 3-octene-2-one give fruity/floral aroma in rice, contributing to rice flavor⁴³. We observed a narrow distribution of 3-octene-2-one content in Wuchang rice (Fig. 3a), indicating a potential of 3-octene-2-one in determining the geographic origin of rice. Interestingly, 3-octene-2-one is considered to be the most active ketone in rice, with a rose aroma and excellent long-lasting flavor, leading to the most intense flavor of rice.

Conclusion

A total of 22 volatile compounds were identified by HS-SPME-GC-MS. Among the identified compounds, aldehydes, alcohols, ketones, hydrocarbon were observed to be predominance compounds. Among them, heptanal, octanal, (E)-2-decenal, 1-heptanol, (E)-2-decen-1-ol, 3,7,11-Trimethyl-3-dodecanol, 3-octene-2-one, and 2-acetyl-1-pyrroline was used to differentiate Wuchang rice from the others using PCA analysis. The amount of 2-acetyl-1-pyrroline was significant higher than the others. 2-AP together with the other 7 compounds constitute the flavor characteristics of Wuchang rice. The methods can be considered to analyze the rice flavor trait of different cultivars.

Data availability

The data that support the findings of this study are available from the corresponding author L M upon reasonable request.

Received: 6 December 2022; Accepted: 8 May 2023

Published online: 17 May 2023

References

- Reig-Valiente, J. L. *et al.* Genetic diversity and population structure of rice varieties cultivated in temperate regions. *Rice* **9**(1), 58–70 (2016).
- Muthayya, S. *et al.* An overview of global rice production, supply, trade, and consumption. *Ann. N. Y. Acad. Sci.* **1324**, 7–14 (2014).
- Zeng, D. *et al.* Rational design of high-yield and superior-quality rice. *Nat. Plants* **3**, 17031 (2017).
- Na, Z. *et al.* Hazard analysis of traditional post-Harvest pperation methods and the loss reduction effect based on five time (5T) management: the case of rice in Jilin province, China. *Agriculture-Basel* **11**(9), 877–898 (2021).
- Vemireddy, L. R. *et al.* Review of methods for the detection and quantification of adulteration of rice: Basmati as a case study. *Journal of Food Sci. Technol.* **52**(6), 3187–3202 (2015).
- Perez-Rodriguez, M. *et al.* Assessing mineral profiles for rice flour fraud detection by principal component analysis based data fusion. *Food Chem.* **339**, 128125 (2021).
- Chunlin, L. *et al.* Geographical origin modeling of Chinese rice using stable isotopes and trace elements. *Food Control* **138**(4), 108997–109005 (2022).
- Meiling, S. *et al.* Predicting isoscapes based on an environmental similarity model for the geographical origin of Chinese rice. *Food Chem.* **397**(12), 133744–133752 (2022).
- Chae, Y. K. *et al.* Discrimination of rice products by geographical origins and cultivars by two-dimensional NMR spectroscopy. *Bull. Korean Chem. Soc.* **37**(10), 1612–1617 (2016).
- Xing, L. *et al.* Determining the geographical origin and cultivation methods of Shanghai special rice using NIR and IRMS. *Food Chem.* **394**(11), 1–9 (2022).
- Zhu, L. *et al.* Identification of rice varieties and determination of their geographical origin in China using Raman spectroscopy. *J. Cereal Sci.* **82**, 175–182 (2018).
- Liyanaarachchi, G. D. *et al.* Volatile profiles of traditional aromatic rice varieties in Sri Lanka. *J. Natl. Sci. Found.* **42**(1), 87–93 (2014).
- Miller, M. E. & Stuart, J. D. Comparison of gas-sampled and SPME-sampled static headspace for the determination of volatile flavor components. *Anal. Chem.* **71**(1), 23–27 (1999).
- Domínguez, R. *et al.* Characterization of volatile compounds of dry-cured meat products using HS-SPME-GC/MS technique. *Food Anal. Methods* **12**(6), 1263–1284 (2019).

15. Dong, W. *et al.* Comparative evaluation of the volatile profiles and taste properties of roasted coffee beans as affected by drying method and detected by electronic nose, electronic tongue, and HS-SPME-GC-MS. *Food Chem.* **272**, 723–731 (2019).
16. Hu, X. *et al.* Volatile compounds, affecting factors and evaluation methods for rice aroma: A review. *Trends Food Sci. Technol.* **97**, 136–146 (2020).
17. Wu, W. *et al.* Impact of contract farming on farmers' income: A case of Wuchang Rice in China. *Jarq-Japan Agric. Res. Q.* **54**(2), 171–177 (2020).
18. Deng, Y. *et al.* Effect of hydrostatic high pressure pretreatment on flavor volatile profile of cooked rice. *J. Cereal Sci.* **58**(3), 479–487 (2013).
19. Feng, Y. *et al.* Research on differential metabolites in distinction of rice (*Oryza sativa* L.) Origin Based on GC-MS. *J. Chem.* **2019**, 1614504 (2019).
20. Qian, L. *et al.* Determination of geographical origin of Wuchang Rice with the geographical indicator by multielement analysis. *J. Food Qual.* **2019**, 8396865 (2019).
21. Song, H. *et al.* Fraud vulnerability quantitative assessment of Wuchang rice industrial chain in China based on AHP-EWM and ANN methods. *Food Res. Int.* **140**, 109805 (2021).
22. Bryant, R. J. & McClung, A. M. Volatile profiles of aromatic and non-aromatic rice cultivars using SPME/GC-MS. *Food Chem.* **124**(2), 501–513 (2011).
23. Lim, D. K. *et al.* Non-destructive profiling of volatile organic compounds using HS-SPME/GC-MS and its application for the geographical discrimination of white rice. *J. Food Drug Anal.* **26**(1), 260–267 (2018).
24. Zeng, Z. *et al.* Flavor volatiles of rice during cooking analyzed by modified headspace SPME/GC-MS. *Cereal Chem.* **85**(2), 140–145 (2008).
25. Li, Y. L. *et al.* Study on the Identification of geographical Indication Wuchang Rice Based on the content of inorganic elements. *Spectrosc. Spectral Anal.* **36**(3), 834–837 (2016).
26. Lim, D. K. *et al.* The integration of multi-platform MS-based metabolomics and multivariate analysis for the geographical origin discrimination of *Oryza sativa* L. *J. Food Drug Anal.* **26**(2), 769–777 (2018).
27. Kim, S. *et al.* Expedient discrimination of four species of the panax genus using direct infusion-MS/MS combined with multivariate statistical analysis. *J. Chromatogr. B* **1002**, 329–336 (2015).
28. Grimm, C. C. *et al.* Screening for 2-acetyl-1-pyrroline in the headspace of rice using SPME/GC-MS. *J. Agric. Food Chem.* **49**(1), 245–249 (2001).
29. Laguerre, M. *et al.* Rapid discrimination of scented rice by solid-phase microextraction, mass spectrometry, and multivariate analysis used as a mass sensor. *J. Agric. Food Chem.* **55**(4), 1077–1083 (2007).
30. Xuan, W. *et al.* Comparison of the sensory properties of fragrant and non-fragrant rice (*Oryza sativa*), focusing on the role of the popcorn-like aroma compound 2-acetyl-1-pyrroline. *Food Chem.* **339**(3), 1–9 (2021).
31. Cutler, D. R. *et al.* Random forests for classification in ecology. *Ecology* **88**(11), 2783–2792 (2007).
32. Du, B. *et al.* Effects of Exogenous alpha-ketoglutaric acid on 2-Acetyl-1-Pyrroline, yield formation and grain quality characters of aromatic rice. *Phyton-Int. J. Experim. Bot.* **90**(2), 437–447 (2021).
33. Routray, W. & Rayaguru, K. 2-Acetyl-1-pyrroline: A key aroma component of aromatic rice and other food products. *Food Rev. Intl.* **34**(6), 539–565 (2018).
34. Wei, X. *et al.* Evaluation of 2-acetyl-1-pyrroline in foods, with an emphasis on rice flavour. *Food Chem.* **232**, 531–544 (2017).
35. Mathure, S. V. *et al.* Quantification of 2-Acetyl-1-pyrroline and other rice aroma volatiles among indian scented rice cultivars by HS-SPME/GC-FID. *Food Anal. Methods* **4**(3), 326–333 (2011).
36. Okpala, N. E. *et al.* The genetics and biosynthesis of 2-acetyl-1-pyrroline in fragrant rice. *Plant Physiol. Biochem.* **135**, 272–276 (2019).
37. Sakhivel, K. *et al.* Genetic and molecular basis of fragrance in rice. *Biotechnol. Adv.* **27**(4), 468–473 (2009).
38. Mahattanatawee, K. & Rouseff, R. L. Comparison of aroma active and sulfur volatiles in three fragrant rice cultivars using GC-olfactometry and GC-PFPD. *Food Chem.* **154**, 1–6 (2014).
39. Ying, X. *et al.* Determination of 2-acetyl-1-pyrroline in aroma rice using gas chromatography-mass spectrometry. *Chin. J. Chromatogr.* **28**(8), 782–785 (2010).
40. Zhao, Q. *et al.* Volatile fingerprints and biomarkers of Chinese fragrant and non-fragrant japonica rice before and after cooking obtained by untargeted GC/MS-based metabolomics. *Food Biosci.* **47**(8), 101764 (2022).
41. Concepcion, J. C. T. *et al.* Quality evaluation, fatty acid analysis and untargeted profiling of volatiles in Cambodian rice. *Food Chem.* **240**, 1014–1021 (2018).
42. Yaylayan, V. A. & Keyhani, A. Origin of 2,3-pentanedione and 2,3-butanedione in D-glucose/L-alanine Maillard model systems. *J. Agric. Food Chem.* **47**(8), 3280–3284 (1999).
43. Daygon, V. D. *et al.* Understanding the Jasmine phenotype of rice through metabolite profiling and sensory evaluation. *Metabolomics* **12**, 63–77 (2016).

Author contributions

S.H. performed the experiments and wrote the manuscript draft; H.R. and F.L. helped with the aroma experiments. Songyong helped with methodology and data analysis. L.Q. and F.Z. help with manuscript revision. L.M. designed the experiments, provided guidance and finalized the manuscript.

Funding

This work was supported by National Key Research and Development Program of China (2017YFC1601905) and Heilongjiang Province University Basic Research Operating Costs Heilongjiang University Special Project (KJCX201817).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-34797-2>.

Correspondence and requests for materials should be addressed to L.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023