



OPEN ACCESS

EDITED BY

Serban Moldoveanu,
Independent Researcher, Greensboro, NC,
United States

REVIEWED BY

Xiaofang Xie,
Fujian Agriculture and Forestry University, China
Yog Raj,
CSIR-Central Drug Research Institute, India
Yanlin An,
Moutai Institute, China
Xuebin Zhao,
China Tobacco Henan Industrial Co., Ltd., China

*CORRESPONDENCE

Shuqi Wang,
✉ wangshuqi@bttbu.edu.cn
Yechun Lin,
✉ linyechun@live.cn

RECEIVED 17 April 2025

ACCEPTED 10 September 2025

PUBLISHED 01 October 2025

CITATION

Zhong R, Sun Z, Feng L, Chen H, Wang S, Lin Y,
Sun J, Zhang N, Zhang H and Wang F (2025)
Decoding honey-sweet flavored flue-cured
tobacco from Guizhou with data science and
flavoromics by volatile and cell
wall components.
Front. Chem. 13:1613828.
doi: 10.3389/fchem.2025.1613828

COPYRIGHT

© 2025 Zhong, Sun, Feng, Chen, Wang, Lin, Sun,
Zhang, Zhang and Wang. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](#).
The use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Decoding honey-sweet flavored flue-cured tobacco from Guizhou with data science and flavoromics by volatile and cell wall components

Risheng Zhong¹, Zhenchun Sun², Liang Feng¹, Haitao Chen¹,
Shuqi Wang^{1*}, Yechun Lin^{2*}, Jie Sun¹, Ning Zhang¹,
Huiying Zhang¹ and Feng Wang²

¹Beijing Advanced Innovation Center for Food Nutrition and Human Health, Beijing Key Laboratory of Flavor Chemistry, Beijing Technology and Business University (BTBU), Beijing, China, ²Tobacco Science Institute of Guizhou Province, Guiyang, Guizhou, China

Nicotiana tabacum L. is often called tobacco. The aroma of flue-cured tobacco (FCT) varies according to the origin and grade. In this study, volatiles and plant cell wall components (CWC) were used to differentiate aroma types and grades of FCT, with a focus on the honey-sweet flavored FCT from Guizhou, China. Volatiles were analyzed by headspace solid-phase microextraction gas chromatography/mass spectrometry, while CWC (cellulose, hemicellulose, pectin, lignin) were quantified. Results indicated that upper leaf (Grade B) tobacco contained higher volatile levels than middle leaf (Grade C). Multivariate analyses-Principal component analysis (PCA), orthogonal partial least squares discriminant analysis (OPLS-DA), and logistic regression (LR) identified 27 key volatiles contributing to aroma differentiation in FCT origin. By combining the screened volatiles with odor activity value, the most important key aroma compounds that distinguish Guizhou honey-sweet flavored from other origins were β -cyclocitral and 1-nonanal. The CWC showed significant variation across origins or grades. Machine learning models (e.g., LR with 96.5% accuracy) effectively distinguished the origin of FCT. This study pioneers the integration of machine learning with molecular sensory science to decode the unique honey-sweet flavor of Guizhou flue-cured tobacco, addressing a critical gap in linking volatile biomarkers to regional terroir. This methodology provides a way to evaluate tobacco quality and aroma characteristics.

KEYWORDS

flavor, volatile, flue-cured tobacco, aroma type, cell wall components, analysis

1 Introduction

Nicotiana tabacum L., commonly referred to as tobacco, is a common perennial plant belonging to the genus *Nicotiana*, family Solanaceae (Philip, 2014). Tobacco is one of the most widely grown cash crops in the world (Peedin, 2011). In addition to its use as a raw and processed material in the manufacture of cigarettes, tobacco has been employed as an insecticide, anesthetic, sedative, diuretic and analgesic (Hu Z. et al., 2021; Lei et al., 2013). China is the world's largest producer of tobacco, in terms of both area cultivated and quantity produced (Martins-da-Silva et al., 2022). Tobacco leaves undergo a series of

processes including roasting, leaf beating, re-roasting, and tobacco mellowing. The final product is a roasted tobacco with a unique aroma (Hu B. et al., 2021). The flavor of tobacco varies markedly among different production regions. In China, flue-cured tobacco (FCT) is categorized into eight distinct types based on their olfactory characteristics, with each type corresponding to a specific geographical area of production. The FCT in Guizhou is distinguished by its distinctive aroma type, which is characterized by a honey-sweet quality. The main characteristics are a prominent honey-sweet aroma, with the main notes being hay-like and honey-sweet. The auxiliary notes include a mellow-sweet, woody, fresh-sweet aroma, among others (Luo, 2019). Aroma can reveal the quality of the tobacco. Cell wall components (CWC) also significantly influence tobacco quality. The most prevalent CWC include cellulose, hemicellulose, pectin, and lignin (Zhao et al., 2019). A positive correlation has been demonstrated between CWC and the maturity, structure, identity, oil content and coloration of tobacco (Chen et al., 2011). It is evident that cellulose, pectin and lignin have a significant impact on the quality of tobacco (Liu, 2015). The CWC and volatiles present in tobacco have a significant impact on its usability and quality as a raw material for the subsequent processing of products. The detection of aroma and CWC collectively serves as an indicator of tobacco quality (Johannes and Wolf-Dieter, 2020).

The aroma of tobacco is typically derived from the degradation of terpenes, terpene alcohols, esters, lipids, alkaloids and other precursors, which are commonly found in phenylalanine degradation products, carotenoid degradation products and cedarane-like degradation products, amongst others (Fan et al., 2022; Geng et al., 2023; Liu A. et al., 2022). A range of methods exists for studying the volatiles of tobacco. Different pretreatment and detection methods are employed in different situations. In order to achieve optimal extraction outcomes, the process of solid-phase microextraction (SPME) was enhanced by employing Microwave-assisted deep eutectic-SPME and Microwave-assisted water eutectic-SPME, leading to the identification of 31 and 13 volatiles, respectively, which were not detected when SPME was utilized in its conventional form (Nie et al., 2017). A total of 81 volatiles were detected in Canadian and Yunnan-roasted tobacco samples. These were analyzed using conventional extraction methods in conjunction with distillation extraction (Tie et al., 2024). However, the vast majority of volatile substances are detected by headspace (HS)-SPME-GC/MS, which is the preferred method for volatile compounds and complex samples due to its solvent-free, cost-effective, convenient, rapid, relatively stable, multi-functional and high-throughput sample preparation characteristics (Jalili et al., 2020). For instance, a total of 58 volatiles were detected by SPME-GC-MS in tobacco from eight different regions (Wang et al., 2024), 56 volatiles were detected by SPME-GC-MS in 14 different types of tobacco from the Fujian production area (Gong et al., 2023), and a total of 40 volatiles were detected by SPME-GC-MS for volatiles during the fermentation of roasted tobacco leaves (Zhang et al., 2024).

Furthermore, molecular sensory science can be used to investigate the significance of various volatiles on aroma, frequently employing the odor activity value (OAV) to ascertain whether the concentration attains the olfactory threshold, thereby determining the compound's effect on the aroma (Guadagni et al.,

1966). This method is also used in tobacco aroma research (Zhao et al., 2025). In the field of data analysis, the integration of statistics and computing technology has enhanced the efficiency of data utilization, facilitating the identification of associations between variables such as geographical origin and sensory characteristics. The incorporation of chemometrics and machine learning algorithms into data processing has become a prevalent practice in the analysis of tobacco aroma. The evaluation model developed by Ye, X.F. et al. based on aroma quality and principal component analysis can more objectively, directly and accurately reflect the aroma quality of tobacco (Ye et al., 2011). Support Vector Machine (SVM) algorithm was used together to select 22 chemical compounds by Relief F-Particle Swarm Optimization to classify the tobacco leaves (Gu et al., 2016). Partial least squares regression (PLSR) analysis has the capacity to reveal that aroma precursors were positively correlated with the sensory quality characteristics of flue-cured tobacco (Li et al., 2024). These tools and methods have facilitated the identification of volatile compounds among the numerous volatiles that are more likely to have a significant impact on aroma. Despite extensive research on tobacco volatiles, the mechanistic link between region-specific aroma signatures and molecular drivers remains poorly understood.

The aim of this study is to identify the volatile compounds responsible for the honey-sweet flavor profile in FCT. Samples were collected from 11 production regions in Guizhou Province, China, and three neighboring provinces. A total of 28 samples were collected from each region, comprising upper leaves (Grade B) and middle leaves (Grade C). Volatile compounds were extracted from the tobacco leaves using HS-SPME/GC-MS technology and the CWC content was measured simultaneously. The study examined the discriminatory power of CWC content across different growing regions and grades. Chemometric methods and machine learning logistic regression were employed to investigate how the types and levels of volatile compounds distinguish origin and grade. The study integrated chemometrics, machine learning and molecular sensory science to identify the characteristic aroma compounds that distinguish Guizhou tobacco from that of other regions. The results suggest that the distinctive honey-like flavor profile of Guizhou tobacco is determined by specific volatile compounds.

2 Materials and methods

2.1 Materials preparation

The FCT were collected from 14 origins, with each origin contained upper leaf (Grade B) and middle leaf (Grade C), a total of 28 samples. The B and C grades are from different parts of the leaf, and all the color grades of the leaf are orange-yellow. All samples were provided by the Guizhou branch of China National Tobacco Corporation (CNTC). Samples were collected from representative farmers in key monitored townships, selecting households characterized by medium-to-high production levels, medium-sized operations, and stable tobacco cultivation. Three households were chosen from each monitoring area. The sampled variety was Yunyan 87, which was cut into fine shreds and uniformly blended. All samples can be categorized into two major origins, those within Guizhou province and those outside

Guizhou province. Among the origins, HBNS, HNCZ and YNDL were tobacco leaves from outside Guizhou province, and the rest of the origins were samples from inside Guizhou province, as shown in [Supplementary Tables SA1, SA2](#).

2.2 Chemicals

Normal alkanes C_6 – C_{30} were purchased from Sigma-Aldrich Company, USA. deuterated hexanol-D13 was purchased from Beijing Quanpinsu Biotechnology Co., Ltd. Dichloromethane was purchased from Beijing InnoChem Science & Technology Co., Ltd. Helium was purchased from Beijing Beihe Pufen Gas Industry Co., Ltd. The reagents for CWC assay were purchased from Sinopharm Chemical Reagent Co.

2.3 Measurement of CWC

2.3.1 Pectin measurement

The sample (0.1 g) was weighed and transferred into a centrifuge tube. The tube was then filled with steel beads and 1 mL of anhydrous ethanol. The contents of the tube were subsequently agitated using a grinding apparatus. The tube was placed into a heating block and subjected to a temperature of 90 °C for a duration of 10 min. The contents of the tube were subjected to a centrifugal force of 10,000 rpm for a duration of 3 min. Remove the upper layer of the mixture and leaving the precipitate. Repeat the following steps three times: add 70% ethanol at 80 °C, rinse the precipitate, centrifuge, remove the supernatant and leave the precipitate. Add 1 mL of sulfuric acid solution (pH = 0.5), heat at 90 °C for 1 h, and centrifuge at 10,000 rpm for 3 min. Aspirate 100 μ L of supernatant and add 275 μ L of concentrated sulfuric acid. Mix well and water bath at 90 °C for 15 min. Add 25 μ L of carbazole ethanol solution and mix well. Then, react in the dark for 30 min. Aspirate 200 μ L of supernatant on top of the enzyme plate, and then measure the absorbance value at 525 nm. Absorbance value was measured at 525 nm. At the same time, different concentrations of galacturonic acid solution was configured, and the standard curve was plotted ([NY/T 2016-2021](#), [2025](#)).

2.3.2 Lignin measurement

A dried and sieved sample (0.05 g) was added to 1.5 mL of 80% ethanol. The mixture was then vortex-shaken and mixed. Subsequently, the mixture was placed in a 50 °C water bath for 20 min. The mixture was allowed to cool, and then subjected to 12,000 rpm centrifugation for 10 min. The precipitate was removed. Add 1 mL of 80% ethanol to the precipitate and mix it with shaking for 2 min. Then, put it in a 50 °C water bath for 20 min. Next, cool it, centrifuge it at 12,000 rpm for 10 min, discard the liquid part, leave the solid part, and dry the solid part at 95 °C. After that, add 0.1 mL of a 30% acetyl bromide solution, mix well, and place in a water bath at 60 °C for 1 h. After that, add 0.4 mL of 2 mol/L NaOH solution and 0.5 mL of ice-acetic acid solution, mix well, centrifuged at 12,000 rpm for 3 min, take the supernatant, put into the ultraviolet spectrophotometer in the 280 nm colorimetry ([Johnson et al., 1961](#)).

2.3.3 Cellulose measurement

A dried sample (0.05 g) was weighed and placed in a 2-mL centrifuge tube. Then, 1 mL of 60% H_2SO_4 was added to the tube, and it was digested in an ice bath for 30 min. The tube was spun at 10,000 rpm for 3 min, and the liquid on top was collected. Dilute the supernatant 50 times, add 30 μ L of 2% anthrone to 120 μ L of the solution to be measured. Add 250 μ L of concentrated H_2SO_4 along the wall of the tube. Mix well, and then immediately put it into a boiling water bath for 10 min, take it out and cool it down. Pipette 200 μ L into a 96-well enzyme plate, and then measure the absorbance at 620 nm.

2.3.4 Hemicellulose measurement

Hemicellulose assay. The sample was sieved and weighed 0.05 g, added to 1 mL of calcium nitrate solution, boiled water bath for 10 min, centrifuged at 10,000 rpm for 3 min and the precipitate was left. The precipitate was rinsed 3 times with hot distilled water and centrifuged at 10,000 rpm to leave the precipitate. Add 500 μ L hydrochloric acid solution (2 mol/L), boil water bath for 1 h, add 500 μ L sodium hydroxide solution (2 mol/L), centrifuge at 10,000 rpm for 3 min, aspirate 50 μ L of supernatant and add 150 μ L of DNS reagent to mix well and then boil water bath for 5 min, cool down and then add 800 μ L of distilled water to add 200 μ L of supernatant to enzyme plate after mixing well and then aspirate 200 μ L of supernatant to the enzyme plate, and then measure the absorbance at 540 nm. Absorbance was measured at 540 nm ([Xiong et al., 2005](#)).

2.4 Isolation of volatiles by HS-SPME

The extraction of the FCT volatile compounds was carried out by the HS-SPME method using a 50/30 μ m DVB/Carboxen/PDMS fiber StableFlex/SS (1 cm) (Sigma Aldrich Supelco). 0.3 ± 0.01 g of tobacco and 6 μ L of internal standard deuterated hexanol-D13 at a concentration of 0.948 mg mL^{-1} (dissolved in dichloromethane) were added to a 20 mL headspace flask. The equilibrium time was set at 20 min, and the equilibrium temperature was set at 60 °C in the agitator of Thermo Scientific TriPlus RSH (Thermo Fisher Scientific, Waltham, MA, United States). After the equilibrium, the volatile compounds were adsorbed at the top of the vial by inserting the SPME fiber. The adsorption time was 30 min. After the adsorption, the SPME fiber was desorbed for 5 min while maintaining the GC-MS injection port at 250 °C.

2.5 GC–MS analysis

GC–MS analysis was performed using a Thermo Fisher Trace 1,300 gas chromatograph combined with a Thermo Fisher mass spectrometer (both Thermo Fisher Scientific, Waltham, MA). Separation was performed using a TG-Wax column (30 m $\times$ 0.25 mm $\times$ 0.25 μ m; Thermo Fisher Scientific). The chromatographic conditions of helium as the carrier gas were as follows: column flow rate of 1.2 mL/min, injection temperature, transmission line and ion source temperature of 250 °C, 240 °C and 250 °C, respectively. The column heating program was maintained at 50 °C for 2 min, then increased to 120 °C at a rate of 5 °C/min and

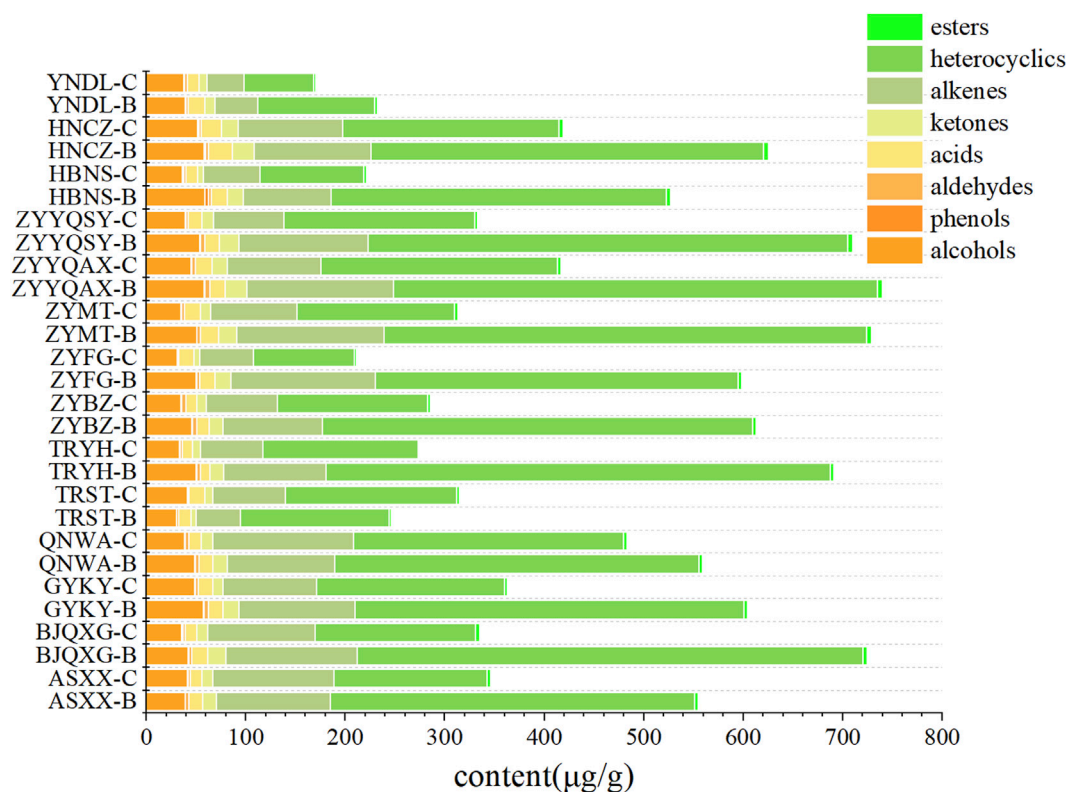


FIGURE 1
Bar stacking plots of eight types of volatile compounds from different origins.

kept for 4 min, then increased to 200 °C at a rate of 4 °C/min and then increased to 230 °C at a rate of 8 °C/min and kept for 4 min.

The electron ionization (EI) mode was used. The electron energy was −70 eV; the mass spectrometry data were collected in full scan mode with a mass range of m/z 40–350 and a solvent delay time of 3 min.

2.6 Data statistics and analysis

The results of peak resolution were matched with the mass spectral information in the database of the National Institute of Standards and Technology (NIST, v1.6). The retention indices of the compounds were calculated according to the retention time of *n*-alkane, and then compared with the retention indices in the database (NIST Chemistry WebBook) to verify the results of the characterization. Finally, the relative quantification of the compounds was performed based on the ratio of the peak areas of the characterized compounds to the peak areas of the internal standards. The data were then subjected to principal component analysis (PCA), orthogonal partial least squares discriminant analysis (OPLS-DA) using SIMCA-P version 13.0 software package (Umetrics, Umea, Sweden). Bar stacking plots were generated using Origin. The data were analyzed to generate heat maps using TBtools. LR is a linear model used to solve classification problems, especially binary classification problems. It is one of the most basic, commonly used, and important algorithms in machine learning. The [Supplementary Code SA1](#), developed with the

assistance of ChatGPT, is provided in the appendices. Odor Activity Value (OAV) was calculated by the following expression: $OAV = C_i/OT_i$, where the C_i is the concentration and the OT_i is the threshold value of the target aromatic components from the book ([Gemert, 2025](#)) and previous literature ([Liu, 2021](#)). The OAVs are frequently used to evaluate the aromatic compounds. The compound with the $OAV > 1$ is significant for the aroma characteristics. IBM SPSS Statistics was used to perform analysis of ANOVA and *t*-test ($P < 0.05$) for significance assessment.

3 Results and discussion

3.1 Analysis of volatile components by HS-SPME-GC-MS

A total of 28 different origins of FCT were analyzed after removal of plasticizers and low-match volatiles. The analysis revealed that all samples possessed the same types of volatile compounds, with variations occurring solely in their composition. A total of 67 volatile components were identified, which belonged to 8 major groups of volatile compounds, including 9 alcohols, 4 phenols, 6 aldehydes, 6 acids, 24 ketones (including four isomers of megastigmatrienone), 2 olefins, 8 heterocycles, and 8 esters ([Supplementary Tables SA1, SA2](#)). The nine most abundant compounds in all samples were nicotine, neophytadiene, menthol, acetic acid, benzyl alcohol, nicotyrine, 6,8-Nonadien-2-one, 8-methyl-5-(1-methylethyl) -, (E)-, phenethyl alcohol, and 2-

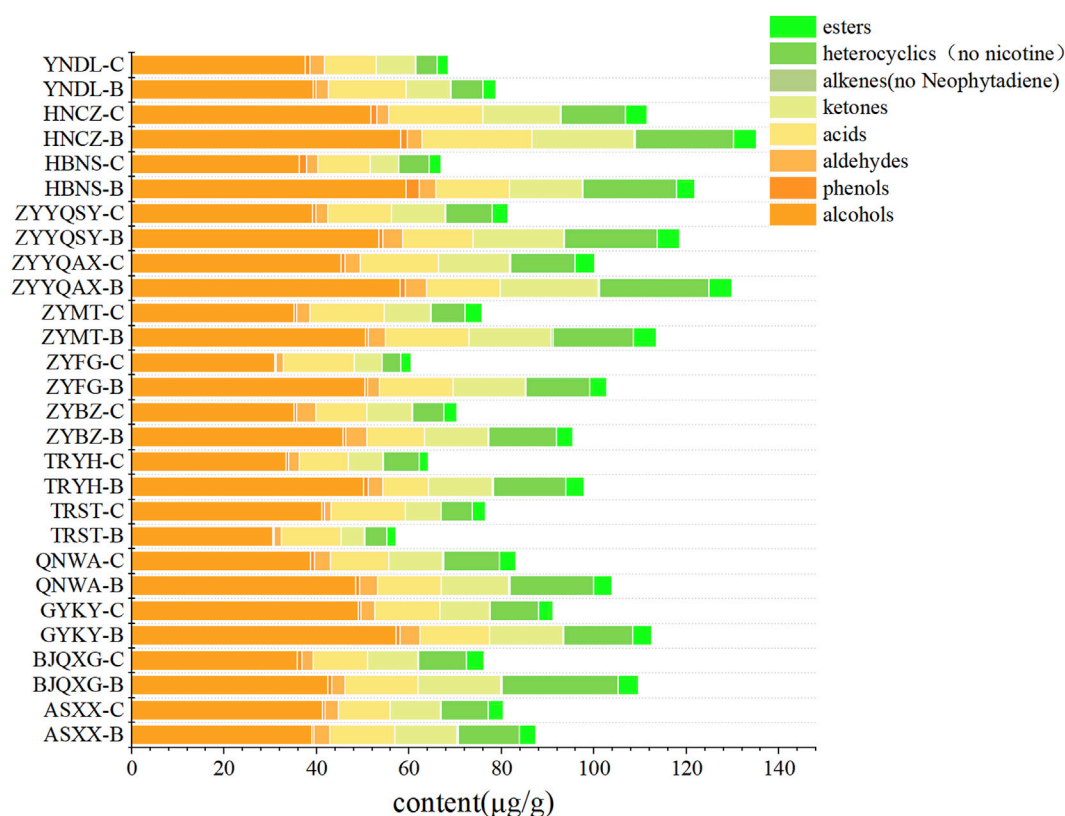


FIGURE 2
Bar stacking plots of eight types of volatile compounds from different origins (Exclusion of nicotine and neophytadiene).

ethylhexanol. Nicotine was the most abundant volatile constituent in all samples, accounting for more than 40% in all samples. Nicotine, a toxic and addictive volatile compound present in tobacco, is characterized by its pungent odor. It is important to note that the detected nicotine content may be inaccurate due to the large peak area in the chromatogram. The more accurate nicotine content would require a separate measurement with a reduced injection volume. Neophytadiene was the second most abundant volatile component in all samples, with a content of more than 14% in all samples. Neophytadiene is a non-pigmented terpene in tobacco, formed by the degradation of chlorophyll during maturation and modulation to form chlorohydrin and then dehydrated. Neophytadiene has a clear aroma and is less irritating, and is itself a precursor of low molecular mass aroma components (Shi et al., 2011). It is evident that neophytadiene confers a distinctive contribution to the aroma of tobacco. The two most abundant volatile components accounted for more than 50% of the total content. The categorization of the volatile constituents with higher contents revealed the presence of benzyl alcohol and menthol among the alcohols, phenol among the phenols, benzaldehyde among the aldehydes, acetic acid among the acids, 6,8-nonadien-2-one, 8-methyl-5- (1-methylethyl)-, (E)- and megastigmatrienone among the ketones, nicotine and nicotyrine among the heterocycles, and 2-ethylhexyl acetate among the esters.

As demonstrated in Figure 1, the highest levels of volatile compounds in grade B were identified in ZYYQAX, ZYMT, and

BJQXG, while the lowest levels were found in YNDL, TRST, and HBNS. The highest levels of volatile compounds in grade C were found in QNWA, HNCZ, and ZYYQAX, and the lowest levels were found in YNDL, HBNS, and ZYFG. The levels of volatile substances in ZYYQAX were generally high, while HBNS had generally low levels of volatile compounds. With the exception of TRST samples, all other samples exhibited higher levels of volatile substances in grade B compared to grade C. The predominant volatile compounds in all samples were identified as heterocyclic and olefinic compounds, with nicotine being the predominant compound in heterocyclic and only two olefinic compounds, primarily contributed by neophytadiene, both of which accounted for more than 90% of the content in their respective categories. To visualize the compounds with lower content, after excluding nicotine and neophytadiene (Figure 2), the content of volatiles detected was higher in class B than in class C for all samples except those with TRST origin. The most abundant types of compounds were alcohols, with the highest levels of menthol and benzyl alcohol. Menthol has cooling, anesthetic and analgesic properties that moderate the irritation and pungency of tobacco. Due to its physiological effects, menthol contributes to the organoleptic quality of smoke and influences the preference of tobacco (Yerger and McCandless, 2011). The most diverse group was ketones, containing four conformations of megastigmatrienone. However, ketones were not among the top three classes of compounds in terms of content, and the relative content was not high. The samples with the highest content of volatiles in the grade B

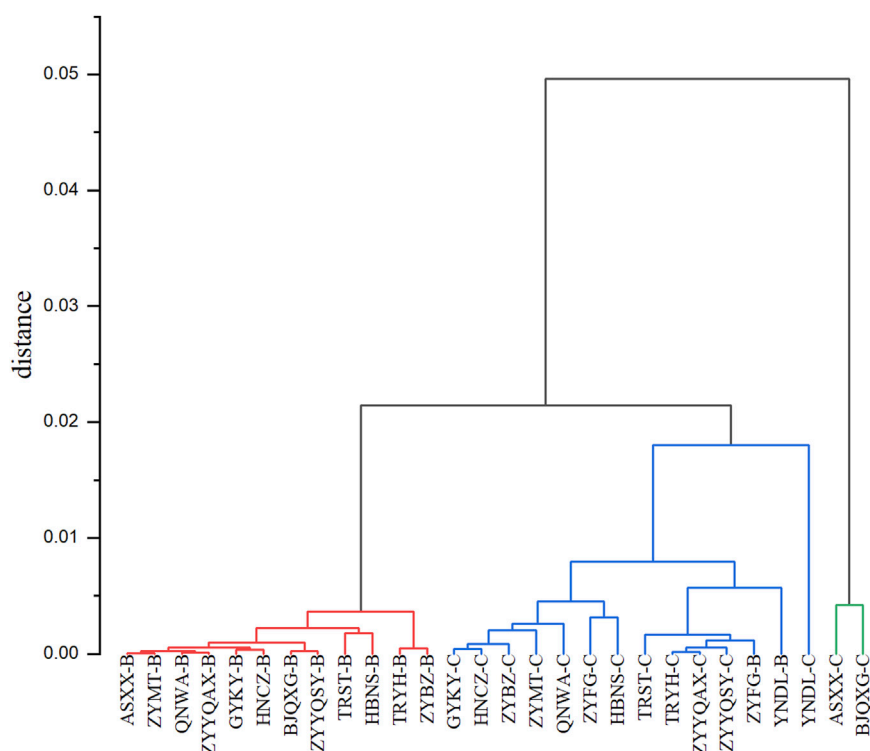


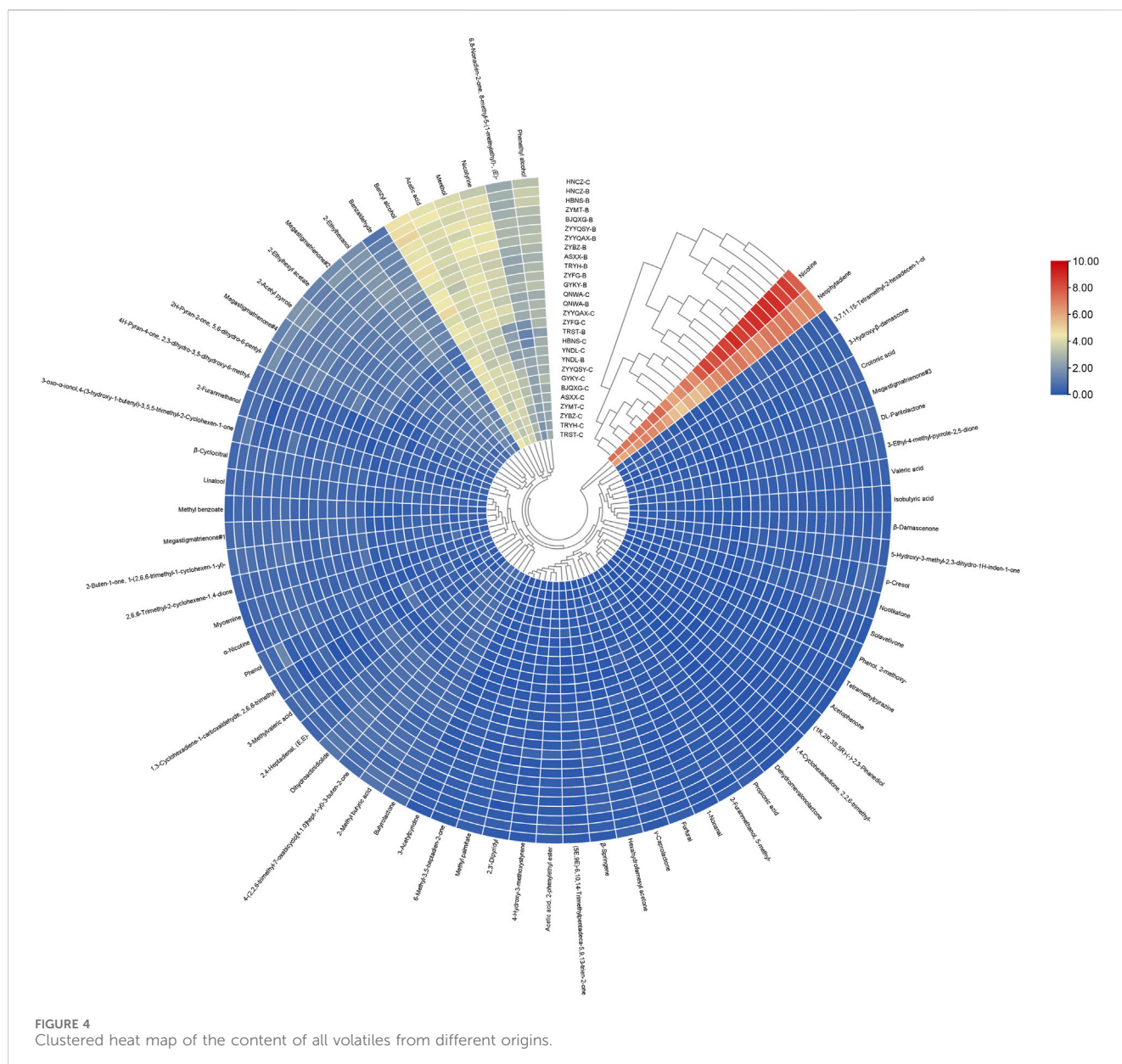
FIGURE 3
Systematic clustering of different origins.

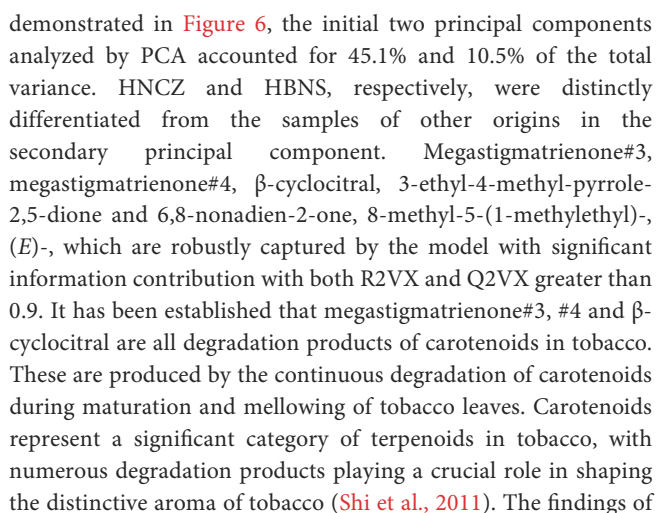
-samples were HNCZ, ZYYQAX, and HBNS, and those with the lowest content were YNDL, ASXXH, and TRST. The samples with the highest content of volatiles in the grade C samples were HNCZ, ZYYQAX and GYKY, and the lowest contents were TRYH, ZYFG and HNNS. It is notable that HNCZ and ZYYQAX exhibited higher concentrations of volatile compounds. The distinction between the two grades lies in the specific leaf region: grade B is designated for the upper leaf, while grade C is allocated for the middle leaf. The *t*-test was used to divide the data into two groups based on grade. There were significant differences in volatile substance content between the two groups. The differences in alcohols, ketones, heterocyclics, heterocyclics without nicotine, and alkenes without neophytadiene were significant. This discrepancy may be attributed to the differing levels of light exposure between the upper and middle leaves within a single plant, which may result in greater accumulation of volatile substances in the upper leaf. Light intensity and quality significantly influence the biosynthesis of secondary metabolites, which are precursors to many volatile aroma compounds (Zhang et al., 2021). However, the presence of more volatile substances does not necessarily guarantee a more significant contribution of aroma substances. The *t*-test was used to divide the data into Guizhou and non-Guizhou categories. There was no significant difference between the two categories of volatile substances.

As demonstrated in the systematic clustering diagram (Figure 3), all the samples are divided into three classes, of which ASXX-C and BJQXG-C are distinct from all the samples of the clustering. The red clustering part of all the grade B samples. The blue clustering part of the samples is mainly grade C samples. The vertical axis represents

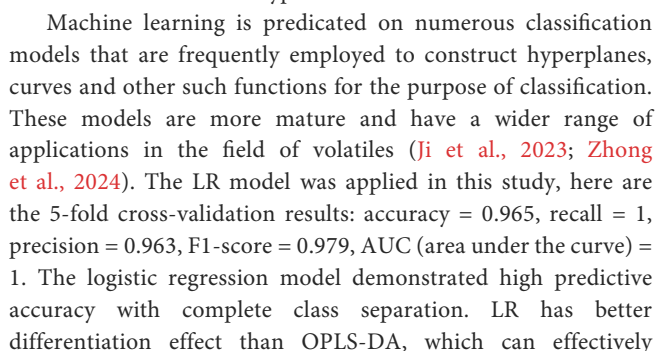
the distance between the different clusters, which are relatively small. This suggests that the clustering of different samples was also small, and that the systematic clustering can differentiate to a certain degree the levels of the tobacco origin. However, the origin cannot be effectively distinguished.

To facilitate a more thorough investigation of the effect of different compounds on clustering, a clustering heat map was produced (Figure 4). The analysis revealed a consistent overall trend in the content of all volatile compounds across all origins, with minor variations in the content of volatile components between origins. From the standpoint of cluster analysis, nicotine and neophytadiene exhibited a tendency to congregate, likely attributable to their preponderance in abundance relative to other compounds. The second most abundant clusters comprised phenethyl alcohol, nicotine, menthol, acetic acid, benzyl alcohol and 6,8-nonadien-2-one, 8-methyl-5-(1-methylethyl)-, (E)-, respectively. Nicotine has a pungent odor and is a volatile with a negative correlation to sensory quality (Chen et al., 2021a). Nicotine has been reported to be a metabolite of nicotine, is a minor tobacco alkaloid (Jalili et al., 2020). Phenethyl alcohol has a typical rose-like odor, benzyl alcohol has a slight almond flavor, all these odors may be directly related to the honey-sweet scent. The third most abundant clusters were seven compounds with two conformations of megastigmatrienone, respectively. The third most abundant clusters were benzaldehyde, 2-ethylhexanol, megastigmatrienone#1, 2-acetyl pyrrole, 2-ethylhexyl acetate, megastigmatrienone#2, and 2H-pyran-2-one,5,6-dihydro-6-pentyl-. The differences characteristics of these compounds led to the categorization of the different origins into distinct clusters.

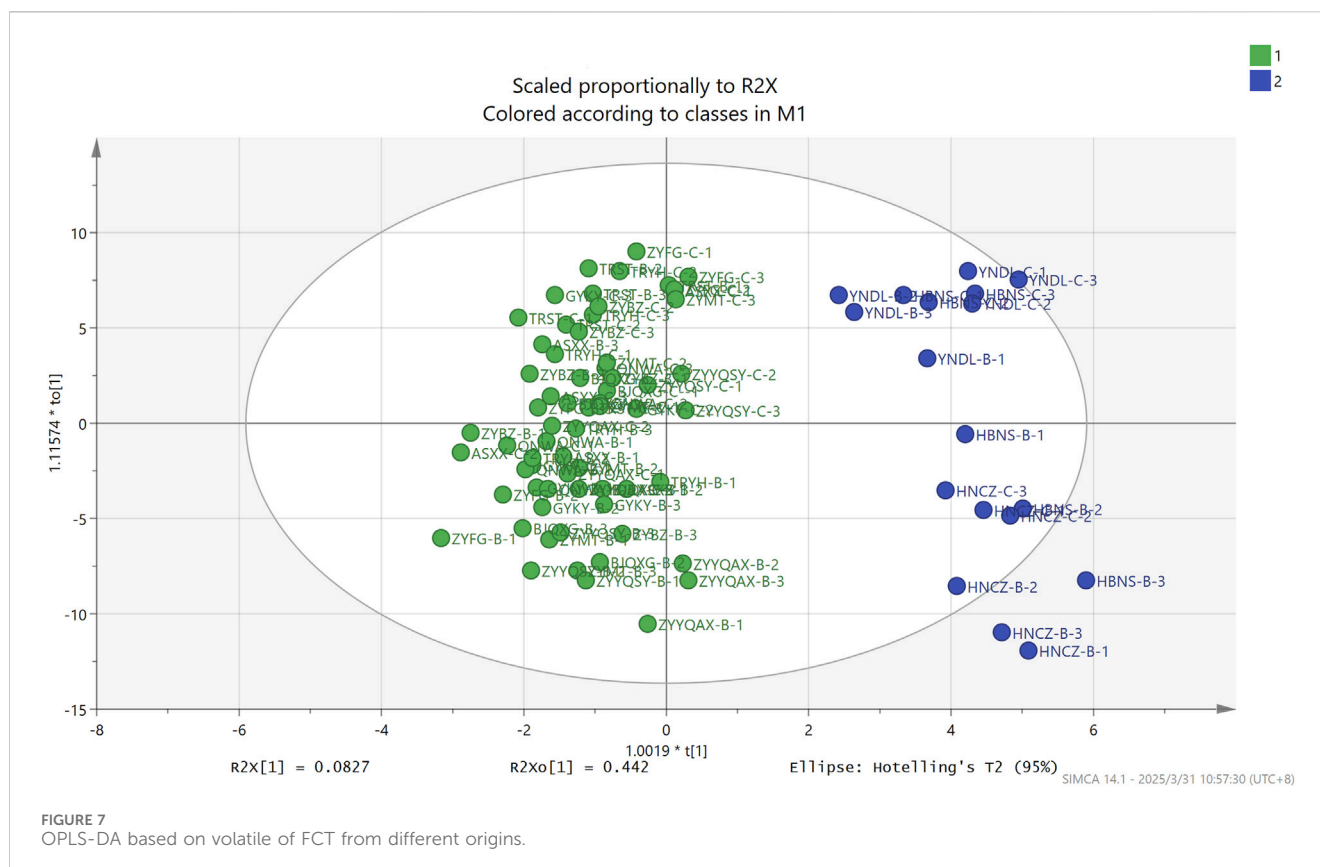




OPLS-DA, a supervised learning approach, facilitates the distinction between different categories determined by which factors. As illustrated in **Figure 7** through the OPLS-DA analysis, it can effectively differentiate tobacco from different origins. In this analysis, the independent variable fitting index R2X is 0.617, the dependent variable fitting index R2Y is 0.882, and the model



differentiate the samples from Guizhou and non-Guizhou provinces. The effect is better, it may be due to the small amount of data itself, which is an easier task for machine learning. The feature importance was then ranked, and the ten compounds with the highest contribution were identified as phenol, 2-methoxy-, 1-nonanal, phenol, furfural, crotonic acid, 1,3-cyclohexadiene-1-carboxaldehyde, 2,6,6-trimethyl-1,3-cyclohexadiene-1-carboxaldehyde (safranal), p-cresol, 2-furanmethanol, 2-acetyl pyrrole and neophytadiene. The volatiles 2-furanmethanol, 2-acetylpyrrole, and neophytadiene are likely responsible for the honey-sweet flavored of FCT in Guizhou, which differs from the aroma of tobacco from other regions. These volatiles are the key factors contributing to the sweet aroma of FCT in Guizhou. Phenol, 2-methoxy- in plants has been reported to be a degradation product of lignin (Zhang et al., 2018). The aroma is smoky and slightly sweet, with a stress and anxiety-relieving effect (Takemoto et al., 2022). Phenol, 2-methoxy-, a major aroma compound in plant samples that have been heated above 40 °C, such as barley and coffee (Joung et al., 2018; Ochiai et al., 2014). The smokiness and sweetness of the aroma may have a direct influence on its overall honey-sweetness. P-Cresol and phenol have a distinctive phenolic aroma and a pungent odor. Furfural has a sweet, toasty aroma that is significantly more affected by climate than the other volatiles (Deng et al., 2010). Collectively, these findings suggest

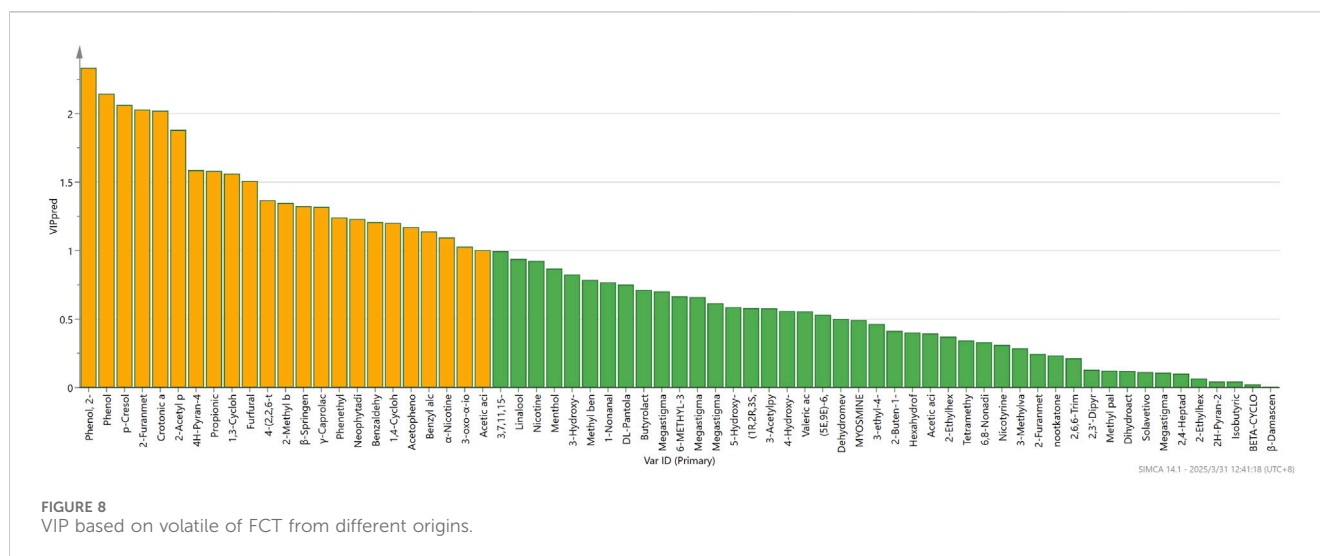


that furfural is an important volatile compound for distinguishing origin and type of aroma. 1,3-Cyclohexadiene-1-carboxaldehyde, 2,6,6-trimethyl-also known as safranal, has a tobacco aroma and pungent flavor. 2-Furanmethanol has also been reported as an important volatile compound in other honey-sweet flavored related studies (Li et al., 2024). 2-Furanmethanol exerts a substantial influence on sensory quality (Song et al., 2012). The analysis revealed the presence of phenol, 2-methoxy-, p-cresol, 2-furanmethanol, crotonic acid, phenol, safranal, and furfural in both statistical methods. These compounds are regarded as key volatile compounds, as they are employed to differentiate between origin and aroma. Furthermore, these compounds are believed to be the volatile compounds that are most likely to influence the differentiation of Guizhou honey-sweet tobacco from other aroma types. The results of system clustering demonstrate that system clustering is incapable of distinguishing between sources. This approach is more influenced by the actual content of the compounds themselves, as the distinction is made based on the differences in compound content between grades B and C. OPLS-DA and LR handle the effects of compound types and content more reasonably, enabling better differentiation between the two. Furthermore, OPLS-DA and LR are supervised learning methods that provide labels in advance during data analysis, allowing for the selection of dimensions that aid in differentiation. PCA, as an unsupervised learning method, also yields relatively average results. The employment of compound content and types to differentiate between Guizhou origins has been demonstrated to yield superior outcomes when utilizing supervised learning methodologies.

3.3 OAV to identify key aroma compounds for differentiating honey-sweet flavored tobaccos

Statistical analysis was employed to identify key volatiles that could be included in honey-sweet flavored FCT in Guizhou, which is different from other producing regions. 23 of them were selected from OPLS-DA and the 10 with the highest contribution from the machine learning model, for a total of 27 volatile compounds screened by all methods. To identify the key aroma substances among them, the psychophysics of olfaction was combined with the concentration and odor threshold of the compounds to determine the compounds that contribute to the overall aroma profile. OAV is a reliable parameter that can be used to assess the contribution of a single compound to the overall aroma of a plant. Volatile compounds with an OAV > 1 are usually considered to be aroma-active compounds that contribute significantly to the overall aroma. In the present study, OAV was applied to quantitatively analyze the key aroma compounds that distinguish Guizhou honey-sweet flavored tobacco from other aroma types.

The analysis revealed a total of nine compounds with OAVs greater than 1 (Supplementary Table SA3). 1-nonanal had the highest OAV, which has a greenish, slightly sweet, slightly beeswax floral aroma, and 1-nonanal is one of the major volatile compounds in cigar tobaccos, where the main flavors are sweet, light and grassy, with more emphasis on fermentation (Wu et al., 2023). 1-Nonanal, a volatile substance, has been shown to have a higher correlation with tobacco pests. This correlation may ultimately affect the flavor of FCT indirectly. However, the relationship between these two substances has not yet been reported (Li et al., 2020; Liu Z. et al., 2022; Sun et al.,



2012). The fungal community *s_Penicillium_vulpinum* showed a significant positive correlation with Nonana (Wu et al., 2023). The second highest OAV is β -cyclocitral, a chemical compound with a floral and fruity aroma. It is the main volatile compound in pineapple nectar, which is characterized by a pronounced honey-sweet flavored (Barros-Castillo et al., 2022). This may result in a close relationship with honey-sweet FCT. Benzaldehyde has an OAV of less than 1 only in ZYFG-C and HBNS-C. γ -Caprolactone has an OAV greater than 1 only in ZYMT-B, ZYMT-C, and ZYYQAX-C. 6,8-Nonadien-2-one, 8-(methyl-5-(1-methylethyl)-, (E)- is only present in TRST-B. Phenol, 2-methoxy- has an OAV of less than 1 only in TRST-B. Phenol, 2-methoxy- has an OAV less than 1 only in ZYMT-C. Phenol, 2-methoxy- has an OAV less than 1 only in ZYYQAX-C. The compounds of these origins correspond to OAVs that are different from those of other origins. The present study examines the correlation between disparate origins and diverse aroma types. In this context, OAV = 1 is important as a threshold value, and there is an effect of subthreshold olfaction on blends of odors. Furthermore, subthreshold olfaction has been shown to have an effect on sweetness (Labbe et al., 2007). The key aroma compounds that distinguish Guizhou-cured tobacco from other tobaccos in terms of honey-sweet flavored are 1-Nonanal, Benzaldehyde, β -cyclocitral, 6,8-Nonadien-2-one, 8-methyl-5-(1-methylethyl)-, (E)-, Phenol, 2-methoxy, Benzyl alcohol, Phenethyl alcohol, p-Cresol. The aroma of tobacco is the result of a combination of internal factors, such as genetics and botany, and external factors, such as ecology, cultivation, physiological and biochemical metabolism, synthesis and degradation of precursors, roasting and microorganisms (Chen et al., 2021b; Gui et al., 2024; Li et al., 2022; Zhang et al., 2024; Zhao et al., 2024; Zhengfeng et al., 2024; Zou et al., 2023).

3.4 Analysis of cell wall components

The plant cell wall is a network of interacting polysaccharides, proteins, small polyphenol molecules and water, and the main load-bearing element is cellulose, which is embedded in a mixture of hemicellulose, pectin and lignin (Pedersen et al., 2023). As can be

seen in Figure 9, the application of ANOVA revealed that there is an obvious pattern of the four CWCs (Supplementary Table 1). The samples with the highest cellulose content are ASXX-C, TRYH-C and YNDL-B, and the lowest are ZYBZ-B, ZTFG-B and HNNS-B. The samples with the highest hemicellulose content are ZYYQAX-C, ZYYQAX-B, and ZYBZ-C, and the lowest are BJQXG-C, BJQXG-B, and ASXX-B. The highest levels of lignin were observed in ZYYQSY-B, BJQXG-B, and TRYH-B, while the lowest levels were detected in YNDL-B, HBNS, and ZYYQAX-C. Conversely, the highest pectin contents were observed in ZYYQAX-C, QNWA-C, and QNWA-B, while the lowest contents were found in BJQXG-C and TRYH-B. The cellulose content of the vast majority of origins is higher for grade B than for grade C. Conversely, QNWA and YNDL exhibited higher lignin content compared to grade B. ZYYMT, ZYYQAX, and HNNS showed the opposite trend. No discernible patterns were observed for hemicellulose and pectin. The factors influencing the quantity of CWC include genes, growth and development, and origin, among others (He et al., 2020). In this study, the tobacco varieties are constant, as are the intrinsic factors (e.g., genes), harvesting conditions, and developmental stages of the tobacco leaves. The primary factors that can exert a substantial influence on the content of CWC are the variations in natural environmental conditions between distinct production areas, such as altitude, climate, and other factors. The application of *t*-test revealed that there were no significant differences in pectin and hemicellulose between samples of grades B and C. However, a significant divergence was observed in cellulose and lignin. A notable distinction was observed in the lignin composition of samples originating from the Guizhou region and those from other regions. However, no significant variations were detected in the pectin and cellulose content.

As plants undergo the process of growth, it is required to release the cell walls to allow for expansion, usually by breaking down the binders called pectin, which are acidic polysaccharides containing galacturonic acid residues (Rui et al., 2017). This process includes the degradation of pectin. In a similar manner, cellulose and hemicellulose are macromolecular polysaccharides composed of glucose. The degradation of these polysaccharides produces volatile organic compounds, with aldehydes and ketones being

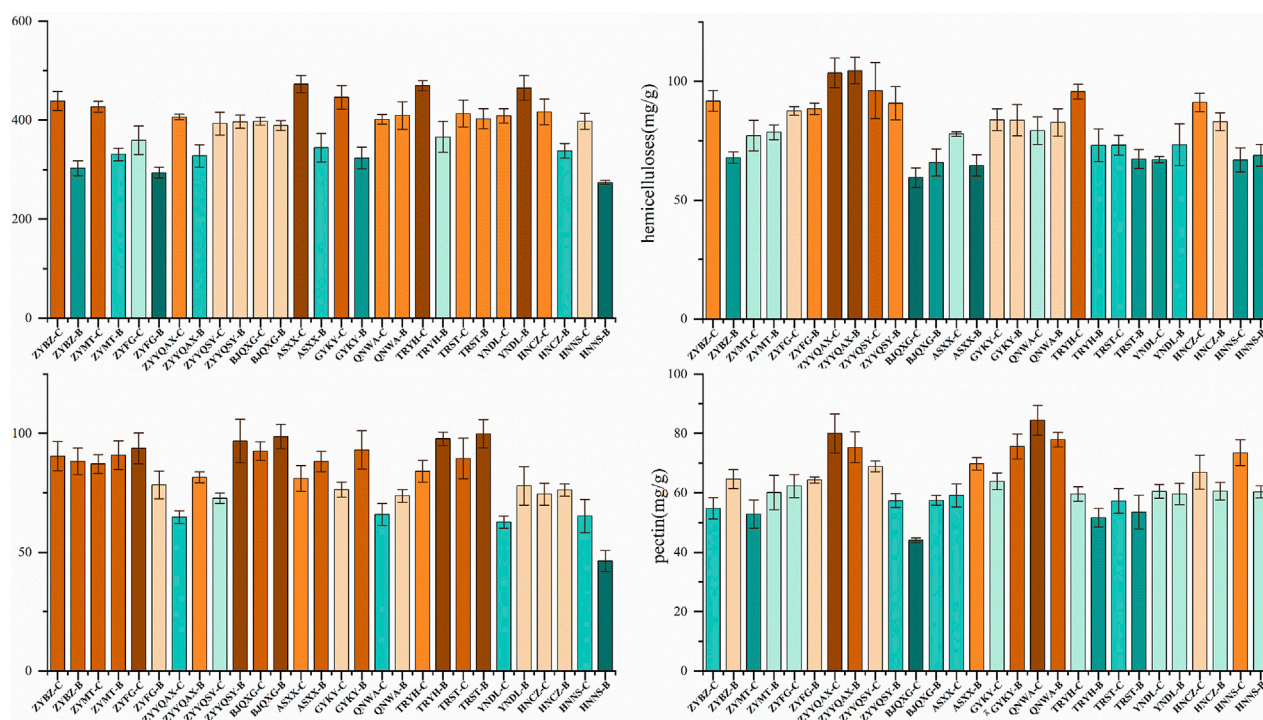


FIGURE 9
CWC content of FCT from different origins.

the most common degradation products (Zhai and Edgar, 2024). In this study, ketones were found to be both abundant and significant. It is possible that they are affected by these CWC. Lignin, a hydrophobic polyphenolic compound formed by phenylalanine derivatives, has been shown to be closely related to the degradation of phenolic compounds (Chen et al., 2021a). The most abundant phenolic compound observed in lignin was 2-methoxy-4-methylphenol, eugenol and phenol, 2-methoxy, respectively (Başakçılardan Kabakcı and Tanış, 2021). Phenylalanine derivatives play a major role as aroma precursors and can form polyphenolic secondary metabolites via phenylpropanoids in the flavonoid pathway (Shi et al., 2011). The study identified four phenolic compounds, of which phenol, 2-methoxy, phenol and p-cresol are important aroma compounds. Furthermore, it has been established that phenylalanine can undergo degradation via the STREKER pathway, resulting in the formation of phenylacetaldehyde and subsequently phenethyl alcohol. The latter is recognized as a critical aroma compounds within tobacco (Song et al., 2012). The significant differences in lignin content suggest that these related volatile compounds are more likely to be responsible for the characteristic aroma of Guizhou FCT, which warrants further validation and research.

4 Conclusion

In this study, the HS-SPME/GC-MS technique was employed to evaluate the volatile compounds of tobaccos from 14 origins, each containing upper leaf (Grade B) and middle leaf (Grade C),

which can be categorized into Guizhou producing regions and non-Guizhou producing regions. The analysis yielded the identification of 27 volatile compounds that distinguished Guizhou honey-sweet flavored FCT from other tobaccos through the application of data science methodologies. The OAV of molecular sensory science identified eight key aroma active substances: 1-nonanal, benzaldehyde, β -cyclocitral, 6,8-nonadien-2-one, 8-methyl-5-(1-methylethyl)-, (E)-phenol, 2-methoxy-, benzyl alcohol, phenethyl alcohol, p-cresol, which are the key aroma compounds to distinguish Guizhou FCT from other tobaccos, and the most important ones are β -cyclocitral and 1-nonanal. Guizhou FCT was found to differ significantly from tobacco from other regions in terms of volatile substances, as confirmed by machine learning. The analysis enabled the distinction between Guizhou FCT and FCT from nearby production areas based on volatile substances. Finally, the CWC present in different production areas were detected, including cellulose, hemicellulose, pectin, and lignin, which are precursors of certain volatile substances. The CWC were not effective differentiation indices between different production areas and grades. In summary, the integration of data science and flavoromics with volatile compound data emerges as a promising approach for the identification of the origin and aroma of FCT. This approach facilitates the exploration of variations in volatile matter between different production areas, the reflection of aroma differences across diverse production areas, and the establishment of a foundation for the quality indicators of different aroma tobacco and product research and development.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary Material](#), further inquiries can be directed to the corresponding authors.

Author contributions

RZ: Conceptualization, Investigation, Methodology, Visualization, Writing – original draft, Writing – review and editing, Data curation. ZS: Data curation, Writing – original draft. LF: Writing – original draft. HC: Resources, Supervision, Writing – original draft, Writing – review and editing. SW: Writing – original draft, Writing – review and editing, Conceptualization, Investigation, Supervision. YL: Supervision, Writing – original draft, Writing – review and editing. JS: Supervision, Writing – review and editing. NZ: Supervision, Writing – review and editing. HZ: Supervision, Writing – original draft. FW: Supervision, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the China National Tobacco Corporation Guizhou Company (2023MX42).

Acknowledgements

The authors are thankful to Zongliang Ji for his valuable assistance and advice in the machine learning section of this paper.

References

- Barros-Castillo, J. C., Calderon-Santoyo, M., Garcia-Magana, M. de L., Calderon-Chiu, C., and Ragazzo-Sanchez, J. A. (2022). Volatile compounds released by acid hydrolysis in jackfruit (*Artocarpus heterophyllus* Lam.). A comparative study by using SDE and HS-SPME techniques. *J. Food Compos. Analysis* 113, 104701. doi:10.1016/j.jfca.2022.104701
- Başakçılardan Kabakcı, S., and Tanış, M. H. (2021). Pretreatment of lignocellulosic biomass at atmospheric conditions by using different organosolv liquors: a comparison of lignins. *Biomass Conv. bioref.* 11, 2869–2880. doi:10.1007/s13399-020-00677-2
- Chen, H., Ren, X., Yang, Y., and Yu, J. (2011). Relationship between cell wall matter content in sichuan flue-cured tobacco and its appearance and sensory quality. *Tob. Science& Technol.*, 9–12.
- Chen, J., He, X., Zhang, X., Chen, Y., Zhao, L., Su, J., et al. (2021a). The applicability of different tobacco types to heated tobacco products. *Industrial Crops Prod.* 168, 113579. doi:10.1016/j.indcrop.2021.113579
- Chen, J., Li, Y., He, X., Jiao, F., Xu, M., Hu, B., et al. (2021b). Influences of different curing methods on chemical compositions in different types of tobaccos. *Industrial Crops Prod.* 167, 113534. doi:10.1016/j.indcrop.2021.113534
- Deng, X.-H., Xie, P.-F., Peng, X.-H., Yi, J.-H., Zhou, J.-H., Zhou, Q.-M., et al. (2010). Effects of soil, climate, and their interaction on some neutral volatile aroma components in flue-cured tobacco leaves from high quality tobacco planting regions of Hunan Province. *Ying yong sheng tai xue bao = J. Appl. Ecol.* 21, 2063–2071.
- Fan, X., Zi, W., Ao, J., Li, B., Qiao, J., Wang, Y., et al. (2022). Analysis and application evaluation of the flavour-precursor and volatile-aroma-component differences between waste tobacco stems. *Heliyon* 8, e10658. doi:10.1016/j.heliyon.2022.e10658
- Gemert, L. J. van (2025). *Odour thresholds*. Edition 2011. Netherlands: Oliemans Punter & Partners BV.
- Geng, Z., He, P., Gao, H., Liu, J., Qiu, J., and Cai, B. (2023). Aroma precursors of cigars from different tobacco parts and origins, and their correlations with sensory characteristics. *Front. Plant Sci.* 14, 1264739. doi:10.3389/fpls.2023.1264739
- Gong, Y., Li, J., Deng, X., Chen, Y., Chen, S., Huang, H., et al. (2023). Application of starch degrading bacteria from tobacco leaves in improving the flavor of flue-cured tobacco. *Front. Microbiol.* 14, 1211936. doi:10.3389/fmicb.2023.1211936
- Gu, L., Xue, L., Song, Q., Wang, F., He, H., and Zhang, Z. (2016). Classification of the fragrant styles and evaluation of the aromatic quality of flue-cured tobacco leaves by machine-learning methods. *J. Bioinform. Comput. Biol.* 14, 1650033. doi:10.1142/S0219720016500335
- Guadagni, D. G., Buttery, R. G., and Harris, J. (1966). Odour intensities of hop oil components. *J. Sci. Food Agric.* 17, 142–144. doi:10.1002/jsfa.2740170311
- Gui, Z.-Q., Yuan, X.-L., Yang, J., Du, Y.-M., and Zhang, P. (2024). An updated review on chemical constituents from *Nicotiana tabacum* L.: chemical diversity and pharmacological properties. *Industrial Crops Prod.* 214, 118497. doi:10.1016/j.indcrop.2024.118497
- He, C., Chen, R., Ren, K., Zhao, G., He, C., Hu, B., et al. (2020). A predictive model for the sensory aroma characteristics of flue-cured tobacco based on a back-propagation neural network. *SN Appl. Sci.* 2, 1867. doi:10.1007/s42452-020-03656-4

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The authors declare that this study received funding from China National Tobacco Corporation Guizhou Company. The funder had the following involvement in the study: the study design and data collection.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fchem.2025.1613828/full#supplementary-material>

- Hu, B., Gu, K., Gong, J., Zhang, K., Chen, D., He, X., et al. (2021a). The effect of flue-curing procedure on the dynamic change of microbial diversity of tobaccos. *Sci. Rep.* 11, 5354. doi:10.1038/s41598-021-84875-6
- Hu, Z., Pan, Z., Yang, L., Wang, K., Yang, P., Xu, Z., et al. (2021b). Metabolomics analysis provides new insights into the medicinal value of flavonoids in tobacco leaves. *Mol. Omics* 17, 620–629. doi:10.1039/d1mo00092f
- Jalili, V., Barkhordari, A., and Ghiasvand, A. (2020). A comprehensive look at solid-phase microextraction technique: a review of reviews. *Microchem. J.* 152, 104319. doi:10.1016/j.microc.2019.104319
- Ji, H., Pu, D., Yan, W., Zhang, Q., Zuo, M., and Zhang, Y. (2023). Recent advances and application of machine learning in food flavor prediction and regulation. *Trends Food Sci. & Technol.* 138, 738–751. doi:10.1016/j.tifs.2023.07.012
- Johannes, N., and Wolf-Dieter, B. (2020). Medicinal, aromatic and stimulant plants.
- Johnson, D. B., Moore, W. E., and Zank, L. C. (1961). The spectrophotometric determination of lignin in small wood samples. Available online at: <https://api.semanticscholar.org/CorpusID:99113043>.
- Joung, W. Y., Kim, S. J., Kim, B. G., Hurh, B. S., and Baek, H. H. (2018). Effects of barley roasting methods on the aroma characteristics of boricha. *Korean J. Food Sci. Technol.* 50, 464–473. doi:10.9721/KJFST.2018.50.5.464
- Labbe, D., Rytz, A., Morgenegg, C., Ali, S., and Martin, N. (2007). Subthreshold olfactory stimulation can enhance sweetness. *Chem. SENSES* 32, 205–214. doi:10.1093/chemse/bjl040
- Lei, B., Zhao, X.-H., Zhang, K., Zhang, J., Ren, W., Ren, Z., et al. (2013). Comparative transcriptome analysis of tobacco (*Nicotiana tabacum*) leaves to identify aroma compound-related genes expressed in different cultivated regions. *Mol. Biol. Rep.* 40, 345–357. doi:10.1007/s11033-012-2067-0
- Li, X., Zhang, X., Xiao, C., Gao, Y., and Dong, W. (2020). Behavioral responses of potato tuber moth (*Phthorimaea operculella*) to tobacco plant volatiles. *J. Integr. Agric.* 19, 325–332. doi:10.1016/S2095-3119(19)62663-8
- Li, X., Bin, J., Yan, X., Ding, M., and Yang, M. (2022). Application of chromatographic technology to determine aromatic substances in tobacco during natural fermentation: a review. *Separations* 9, 187. doi:10.3390/separations9080187
- Li, J., Ma, Z., Dai, H., Li, H., Qiu, J., and Pang, X. (2024). Application of PLSR in correlating sensory and chemical properties of middle flue-cured tobacco leaves with honey-sweet and burnt flavour. *Heliyon* 10, e29547. doi:10.1016/j.heliyon.2024.e29547
- Liu, X. (2015). *The effects of cell wall materials of flue-cured tobacco on tobacco quality in typical ecological area*. Qingdao, China: Chinese Academy of Agricultural Sciences. Available online at: https://kns.cnki.net/kcms2/article/abstract?v=hW9FdoQEerra97jbkutnQTjpq5OSEuXmFwQNad8N06BYWvSMRBgWfCTPaXXw6dC0N4-EAftMJsJtsdbafZwLipKoXI7pgdLL7TH7odAsui0lseHNEfj4d3wVlgMm7EMGAgMWsVOfgpAYH99saN9hUE_TP5Jrw0h3TnfrTksNQ6fOmMa3A9zCibDVGvORyFK2d73KIETt24=&uniplatform=NZKPT&language=CHS.
- Liu, X. (2021). *Odor activity value of tobacco, cigarette, tobacco additive research on the main aroma components of tobacco additives*. Kunming, China: Kunming University of Science and Technology. Available online at: https://kns.cnki.net/kcms2/article/abstract?v=zclOVLBHd2yZX0bXpClAxRRm_cKiZIXRL021GgY2gTjH2baxqFnXwqPvEdQ_qG6F-WT5UGmCB0pF539M88SjXj3W1bZcbpR9UJh5Z7hBBuXVMVVKpekHg=&uniplatform=NZKPT&language=gb.
- Liu, A., Yuan, K., Li, Q., Liu, S., Li, Y., Tao, M., et al. (2022a). Metabolomics and proteomics revealed the synthesis difference of aroma precursors in tobacco leaves at various growth stages. *Plant Physiology Biochem.* 192, 308–319. doi:10.1016/j.plaphy.2022.10.016
- Liu, Z., Chen, W., Zhang, S., Chen, H., Su, H., Jing, T., et al. (2022b). Behavioral responses of *Bemisia tabaci* mediterranean cryptic species to three host plants and their volatiles. *INSECTS* 13, 703. doi:10.3390/insects13080703
- Luo, D. (2019). Explanation of national regionalization of leaves style of flue-cured tobacco. *Acta Tabacaria Sin.* 25 (4). doi:10.16472/j.chinatobacco.2019.218
- Martins-da-Silva, A., Torales, J., Becker, R., Moura, H., Campos, M., Fidalgo, T., et al. (2022). Tobacco growing and tobacco use. *Int. Rev. Psychiatry* 34, 51–58. doi:10.1080/09540261.2022.2034602
- Nie, J., Yu, G., Song, Z., Wang, X., Li, Z., She, Y., et al. (2017). Microwave-assisted deep eutectic solvent extraction coupled with headspace solid-phase microextraction followed by GC-MS for the analysis of volatile compounds from tobacco. *Anal. Methods* 9, 856–863. doi:10.1039/C6AY03076A
- NY/T 2016-2011 (2025). Determination of pectin content in fruit and derived product-spectrophotometry method.
- Ochiai, N., Tsunokawa, J., Sasamoto, K., and Hoffmann, A. (2014). Multi-volatile method for aroma analysis using sequential dynamic headspace sampling with an application to brewed coffee. *J. Chromatogr. A* 1371, 65–73. doi:10.1016/j.chroma.2014.10.074
- Pedersen, G. B., Blaschek, L., Frandsen, K. E. H., Noack, L. C., and Persson, S. (2023). Cellulose synthesis in land plants. *Mol. Plant* 16, 206–231. doi:10.1016/j.molp.2022.12.015
- Peedin, G. F. (2011). “Tobacco cultivation,” in *Specialty crops*.
- Philip, W. (2014). Encyclopedia of toxicology.
- Rui, Y., Xiao, C., Yi, H., Kandemir, B., Wang, J. Z., Puri, V. M., et al. (2017). Polygalacturonase involved in expansion3 functions in seedling development, rosette growth, and stomatal dynamics in *Arabidopsis thaliana*. *Plant Cell* 29, 2413–2432. doi:10.1105/tpc.17.00568
- Shi, H., Liu, G., Yang, H., and Ji, X. (2011). *Tobacco flavoring*. Beijing, China: China Agriculture Press.
- Song, G., Li, D., Dai, Y., Feng, G., and He, R. (2012). Grey dominance evaluation of neutral fragrance components and sensory quality in flue-cured tobacco. *Asian J. Chem.* 24, 4935–4938.
- Sun, J., Huang, L., and Wang, C. (2012). Electrophysiological and behavioral responses of *Helicoverpa assulta* (Lepidoptera: noctuidae) to tobacco volatiles. *Arthropod-Plant Interact.* 6, 375–384. doi:10.1007/s11829-012-9190-7
- Takemoto, H., Saito, Y., Misumi, K., Nagasaki, M., and Masuo, Y. (2022). Stress-relieving effects of sesame oil aroma and identification of the active components. *Molecules* 27, 2661. doi:10.3390/molecules27092661
- Tie, J., Li, S., He, W., Li, Y., Liao, F., Xue, J., et al. (2024). Study of metabolite differences of flue-cured tobacco from Canada (CT157) and Yunnan (yunyan 87). *Heliyon* 10, e32417. doi:10.1016/j.heliyon.2024.e32417
- Wang, Y., Fan, J., Chen, Y., Chen, S., Deng, X., Huang, H., et al. (2024). Correlation between microbial diversity and flavor substances on the surface of flue-cured tobacco. *Anal. Lett.* 57, 1197–1214. doi:10.1080/00032719.2023.2244614
- Wu, X., Hu, Y., Wang, Q., Liu, J., Fang, S., Huang, D., et al. (2023). Study on the correlation between the dominant microflora and the main flavor substances in the fermentation process of cigar tobacco leaves. *Front. Microbiol.* 14, 1267447. doi:10.3389/fmicb.2023.1267447
- Xiong, S., Zuo, X., and Zhu, Y. (2005). Determination of cellulose hemi-cellulose and lignin rice hull. *Cereal & Feed Ind.*, 40–41.
- Ye, X., Liu, G., Liu, H., and Li, S. (2011). Study on model of aroma quality evaluation for flue-cured tobacco based on principal component analysis. *J. FOOD Agric. & Environ.* 9, 501–504.
- Yerger, V. B., and McCandless, P. M. (2011). Menthol sensory qualities and smoking topography: a review of tobacco industry documents. *Tob. Control* 20, ii37–ii43. doi:10.1136/tc.2010.041988
- Zhai, Z., and Edgar, K. J. (2024). Polysaccharide aldehydes and ketones: synthesis and reactivity. *Biomacromolecules* 25, 2261–2276. doi:10.1021/acs.biomac.4c00020
- Zhang, J., Lombardo, L., Gözaydin, G., Dyson, P. J., and Yan, N. (2018). Single-step conversion of lignin monomers to phenol: bridging the gap between lignin and high-value chemicals. *Chin. J. Catal.* 39, 1445–1452. doi:10.1016/S1872-2067(18)63132-8
- Zhang, S., Zhang, L., Zou, H., Qiu, L., Zheng, Y., Yang, D., et al. (2021). Effects of light on secondary metabolite biosynthesis in medicinal plants. *Front. Plant Sci.* 12, 781236. doi:10.3389/fpls.2021.781236
- Zhang, W., Deng, Q., Xiao, D., Chen, Q., Pan, H., Chen, J., et al. (2024). Improving the quality of low-grade tobacco by enzymatic treatment and Co-fermentation with yeast and lactic acid bacteria. *Appl. Biochem. Biotechnol.* 197, 613–630. doi:10.1007/s12010-024-05007-0
- Zhao, Y., Man, Y., Wen, J., Guo, Y., and Lin, J. (2019). Advances in imaging plant cell walls. *Trends Plant Sci.* 24, 867–878. doi:10.1016/j.tplants.2019.05.009
- Zhao, X., Cheng, Q., Luan, M., and Zhang, Y. (2024). Effects of environmental factors on the growth and quality of the tobacco variety zhusha 2: a case study in yunan, China. *Pak. J. Agric. Sci.* 61, 733–741. doi:10.21162/PAKJAS/24.143
- Zhao, X., Chen, X., Wang, E., Ma, S., Pang, Y., Zhang, H., et al. (2025). Unraveling the relationship between key aroma compounds and sensory characteristics of cherry-red tobacco. *ACS Omega* 10, 6092–6104. doi:10.1021/acsomega.4c10431
- Zhengfeng, L., Tian, Q., Yuzhen, X., Le, Z., Luoping, W., Kai, D., et al. (2024). Similarity in the microbial community structure of tobacco from geographically similar regions. *Sci. Rep.* 14, 30933. doi:10.1038/s41598-024-81565-x
- Zhong, R., Ji, Z., Wang, S., and Chen, H. (2024). Bridging odorants and olfactory perception through machine learning: a review. *Trends Food Sci. & Technol.* 153, 104700. doi:10.1016/j.tifs.2024.104700
- Zou, L., Su, J., Xu, T., Ji, X., Wang, T., Chen, Y., et al. (2023). Untargeted metabolomics revealing changes in aroma substances in flue-cured tobacco. *Open Chem.* 21, 20220326. doi:10.1515/chem-2022-0326