Pak. J. Bot., 57(3): DOI: http://dx.doi.org/10.30848/PJB2025-3(3)

REGULATORY EFFECTS OF DIFFERENT GROWTH HABITATS ON THE MICROBIAL COMPOSITION AND CHEMICALS PRODUCED IN K326 TOBACCO

SHEN HUANG[1], KE WANG[1], MENGYUE CHEN[1,2], AAMIR RASOOL[3], ROBINA MANZOOR[4], DUOBIN MAO[1], JIZHONG WU[5*] AND CHANGHE CHENG[5*]

[1]*College of Tobacco Science and Engineering, Zhengzhou University of Light Industry, Zhengzhou 450002, China*
[2]*Shanghai Tobacco Group Beijing Cigarette Factory Co., Ltd., Beijing 101121, China*
[3]*Institute of Biochemistry, Balochistan University, Quetta 87300, Pakistan*
[4]*Department of Biotechnology and Bioinformatics, Lasbella University of Agriculture, Water and Marine Sciences, Uthal 90150, Pakistan*
[5]*Technology Center, China Tobacco Henan Industrial Co., Ltd., Zhengzhou, 450000, China*
**Corresponding author's email: wujizhong@zjtobacco.com, chengch@zjtobacco.com*

Abstract

K326 has excellent botanical and economic properties, good smoke appearance quality, orange color, moderate and strong gloss, medium and more oil content, loose leaf structure, moderate and thin leaf thickness, suitable nitrogen to base ratio, and high industrial availability. The quality of flue-cured tobacco depends on the types of chemicals produced in its leaves. In one tobacco plant, the total amount of furans, aldehydes and organic acids in middle tobacco leaves was significantly higher than that in upper tobacco leaves, but there was no significant difference between middle tobacco leaves and lower tobacco leaves. The total amount of phenols and organic acids in the pyrolysis products of the lower tobacco leaves was the highest and significantly higher than that of the upper tobacco leaves, but the difference was not significant compared with the middle tobacco leaves. There was significant negative correlation between total alkaloids, total nitrogen, total sugars and reducing sugars in all three parts. The different habitats effectively regulate the production of chemicals in tobacco by controlling the composition of the microbiome of its leaves. Nicotine content is different in different growing environments, for example, altitude and nicotine content are negatively correlated. Therefore, the effect of seven different habitats, including Sichuan, Hunan, Yunnan, Henan, Fujian, and Guizhou, on the composition of the K326 tobacco microbiome was investigated in this study using high throughput 16S rRNA Illumina MiSeq sequencing technology in this study. The results demonstrated that the abovementioned growth habitats strongly regulated the composition of the microbiome of K326 tobacco, which resultantly affected the production of chemicals. The following bacterial genera, such as *Allorhizobium, Neorhizobium, Pararhizobium,* and *Rhizobium,* displayed significantly positive correlations with the potassium and nitrogen-nicotine contents of K326 tobacco leaves. *Methylobacterium* influenced the production of 5-Methylfuranaldehyde and furfural. *Halomonas* positively correlated with the production of key neutral aroma compounds (NACs) in K326 tobacco leaves, such as geraniol, 4-Oxyisofluorone,Phenyl ethanol and Damascusone. Our findings show that the aroma of K326 tobacco can be engineered by controlling the microbiome composition by growing tobacco in selective habitats.

Key words: 16S rRNA sequencing, Bacterial community, Neutral aroma compound, Routine chemical compound，Tobacco leaf.

Introduction

Tobacco is unique among many other plants because it houses several thousand chemical compounds essential for flavor, aroma, and physiological effects. The composition and proportion of compounds in tobacco play a crucial role in the quality and flavor of tobacco. By adjusting the content of each compound, the smoking quality of tobacco can be controlled, such as improving the aroma and reducing the irritation. The total sugar, reducing sugar, nicotine, total nitrogen and other chemical components of tobacco leaves and tobacco can be analyzed to test the quality of different tobacco leaves. The aroma of tobacco leaves reflects the integrated response of these compounds, which smokers enjoy through smell, taste, and feel (Villanti *et al*., 2013; Masugi *et al*., 2016; Zhang *et al*., 2016). Among them, K326 has excellent botanical and economic properties, good tobacco appearance quality, orange color, moderate to strong gloss, medium to more oil content, loose leaf structure, moderate to thin leaf thickness, suitable nitrogen to base ratio, and high industrial availability (Zhu *et al*., 2024). The K326 tobacco variety is suitable for strong cultivation in most parts of China.

Microbiomes play an essential role in the aging of tobacco leaves, and thus have a strong influence on the quality of tobacco (Jiaxi, 2020). The microbiota not only improve the sensory profile of the tobacco leaves by transforming the biomolecules of tobacco into NACs but also detoxifies the nicotine and nitrosamines into harmless compounds (Liu, 2015). For instance, *Pseudomonas* spp. has been shown to degrade nicotine into various harmless and aroma enhacing compounds (Zgadza *et al*., 2016). Additionally, *Bacillus* spp. also has been reported to produce small aromatic compounds (dihydroactinodiolide, 3-Oxoionol, and isophorone) from complex molecules, such as carotene (Segata *et al*., 2011).

Neutral aroma compounds (NACs) such as volatile and semi-volatile aldehydes, ketones, alcohols, esters, and alkenes play a vital in shaping the landscape of the sensory profile of tobacco (Jenkins *et al*., 2000; Xiao *et al*., 2016). Numerous studies have investigated the generation of NACs in various flue-cured tobaccos cultivated across diverse agroecological regions (Weeks *et al*., 1992; Jian-Jun *et al*., 2009; Fei *et al*., 2013; Yang *et al*., 2015). A large number of studies have reported that the production of different chemicals, including NACs in the tobacco leaves

is significantly affected by the altitude (Wang *et al*., 2011a, 2015b). However, very little research has been done focusing on the exploration of the distribution of NACs in the K326 tobacco variety.

Therefore, a study in which the impact of various habitats on the microbial composition and chemical profile of K326 tobacco leaves is assessed would be a valuable endeavor. This study will significantly enhance our understanding of how aroma characteristics of tobacco are influenced by different environments.

Material and Method

Collection and processing of tobacco leaves: Tobacco leaf samples were obtained from various sampling sites at different altitudes and latitudes in China, including Sichuan, Hunan, Yunnan, Henan, Fujian, and Guizhou. The cultivation system and fertilization methods applied to the K236 test sample were identical across all locations. The collected samples consisted of upper and middle leaves from the K326 tobacco variety cultivated in the habitats listed in (Table 1).

In the production of tobacco leaves, the upper and middle leaves are usually used, and the lower leaves have more defective products and are not often used. The upper and middle leaves were selected for the experiment. The upper leaves of the K326 variety were labeled as K, R, F, C, P, I, and H. In contrast, the middle leaves were designated as L, Q, E, D, O, J, and G. Following the harvest and drying process, the tobacco leaves were promptly packed into 1.3m cartons (density: 200 kg/carton) for transportation to Henan (HN) for storage. To prepare the samples for analysis, all tobacco samples were dried for 30 minutes in an oven set at 60°C. Subsequently, the dried leaves were ground to achieve a specific granularity using a whirlwind grinding instrument. The tobacco leaves' microbiome composition and chemical compounds, specifically focusing on neutral aroma compounds (NACs), were evaluated using various analytical methods.

Table 1. The habitats of different sampling sites.

	Liangshan	Nanyang	Chenzhou	Sanming	Yuxi	Honghe	Tongren
Elevation	1717m	186m	300m	379m	1817m	1800m	823m
Longitude	E102°32′25″	E111°42′36″	E112°45'46"	E117°09′52″	E102°39′17″	E102°52'00"	E107°59'20.01"
Latitude	N26°36′47″	N32°52′52″	N26°13'30"	N26°50′50″	N24°08′51″	N23°30'10"	N28°08'6.76"

Determination of common chemicals of tobacco: The quantification of total sugar and reducing sugar in the samples of tobacco leaves was performed using the continuous flow method, following the measurement protocol outlined in the tobacco industry standard (YC/T 159-2002). We also estimated the water-soluble sugar content in tobacco and related products and nicotine, total nitrogen, total potassium, and total chlorine levels in tobacco leaves following the tobacco industry standard (YC/T 160-2002). The content of alkaloids in the tobacco and related products was determined through the continuous flow method, following the measurement guidelines specified in the tobacco industry standard (YC/T 161-2002).

Estimating total nitrogen content in tobacco and related products was conducted using the continuous flow method, following the measurement procedure specified in the standard YC/T 173-2003. The quantification of potassium in tobacco and related products using the flame photometric method, per the tobacco industry standard (YC/T 162-2002). The continuous flow method assessed chlorine content in tobacco and related products.

The experiments were conducted in triplicate, and the average values from these replicates were used for the analysis.

Simultaneous distillation extraction (SDE): The SDE was done with a Likens-Nickerson apparatus. The tobacco leaves of K326, sourced from different habitats, were crushed into tobacco powder by a multifunctional pulverizer and sieved through a 40-mesh sieve. We weighed 30.0 g of tobacco powder in a 1L round-bottomed flask, then added zeolite and 0.4L of pure water. The mixture was boiled with an electric heating sleeve, and a simultaneous distillation and extraction device was installed. On the other side of the device, 0.1L of dichloromethane was added to a 0.25 L eggplant bottle. The dichloromethane extraction solution was obtained after 2.5 h after the first reflux. Simultaneous distillation and extraction were performed to obtain the extraction solution of neutral aroma compounds. The neutral aroma compounds (NACs) extract was refrigerated overnight with an appropriate amount of anhydrous Na_2SO_4 and then concentrated to 1 mL at room temperature.

GC-MS analysis: The volatile aromatic (VA) compounds were analyzed using an Agilent 7890 Series gas chromatograph (Agilent, Santa Clara, CA) equipped with FID. A fused-silica column (HP-5) with dimensions of 60 m × 0.25 mm (inner diameter) and a film thickness of 0.25μm was used to separate VA compounds. The temperature of the GC injector was set at 280°C, and a 10:1 split ratio was used for the autosampler. The helium, approximately 99.999% pure, was used as the carrier gas, and the flow rate was set at 1.0 mL/min. The oven temperature was ramped at a rate of 4°C/min from 50°C to 280°C and the final temperature of 280°C was sustained for 10 min.

An Agilent 5973 mass selective detection system coupled with an HP 5980 gas chromatography instrument was utilized for gas chromatography-mass spectrometry analysis of VA compounds. The transfer line temperature of gas chromatography-mass spectrometry was sustained at 280°C, and the instrument operated in EI mode with a voltage of 70 eV. The scanning for detection was set as m/z 35 to 550, and the electronic multiplier tube voltage was set at 230 volts.

Aromatic compounds were detected and identified by matching their mass spectra with the NIST 20 library (Agilent). The identified compounds' retention indices and mass spectra were compared to further verify the qualitative analysis. The quantification of essential neutral

aroma compounds (NACs) was conducted using an external standard method.

Considering their significant influence on tobacco quality according to existing literature, ketones, alcohols, esters, and other neutral flavor compounds were selected as the primary investigation indices in this study.

Extraction of gDNA and sequencing of 16S rRNA gene: The gDNA from the collected samples was isolated using the OMEGA Soil DNA Kit (M5635-02) (Omega Bio-Tek, Norcross, GA, USA). The extracted gDNA samples were then kept at -20°C. The concentration of each gDNA sample was determined with the NanoDrop NC2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), and the DNA samples' purity was verified through DNA gel electrophoresis.

The V3-V4 region of 16S rRNA genes of bacteria was amplified by using the following primers: F (5'-GGAAGTAAAAGTCGTAACAAGG-3') and R (5'-GCTGCGTTCTTCATCGATGC-3'). We added the sample-specific 7 bp sequences to the primers to perform the multiplex sequencing. The PCR reaction mixture included 5× buffer (5 μL), 5 U/μL Fast pfu DNA polymerase (0.25 μL), 2.5 mM dNTPs (2 μL), 10 μM F-primer (1 μL) and 10 μM R-primers (1 μL), DNA template (1 μL), and ddH2O (14.75 μL). The PCR program was designed as follows; the initial denaturation temperature was 98°C and maintained for 5 min, then PCR machine was run for 25 cycles where the denaturation temperature for each cycle was set as 98°C for 30s, annealing temperature was set at 53°C for 30s, and extension temperature was at 72°C for 45s. The final extension step was performed at 72°C for 5 min.

The amplified DNA fragments were purified using Vazyme VAHTSTM DNA Clean Beads (Vazyme, Nanjing, China). After purification, the amplicons were quantified using the Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen, Carlsbad, CA, USA). The sequencing was performed on the Illumina NovaSeq platform using the NovaSeq 6000 SP Reagent Kit (500 cycles) at Shanghai Personal Biotechnology Co., Ltd. in Shanghai, China. The sequencing was conducted as pair-end 2x250 bp reads.

Sequence analysis

The microbiome bioinformatics analysis was performed with slight modifications using QIIME2 (Bolyen *et al*., 2018). Firstly, the raw sequence data were demultiplexed using the demux plugin. Subsequently, the cutadapt plugin was utilized for primer cutting (Martin, 2011). Quality filtering, denoising, sequence merging, and removal of chimeras were carried out using the DADA2 plugin.

The non-singleton amplicon sequence variants (ASVs) were aligned using Mafft and FastTree2 to construct a phylogeny. Alpha-diversity metrics, such as Shannon diversity index (Shannon 1948), Simpson index (Simpson,1949), Good's coverage (Good, 1953), Pielou's evenness (Pielou,1966), Chao 1 richness (Chao, 1984) and Faith's phylogenetic diversity (Faith, 1992) were estimated using the diversity plugin. Sequences were standardized to 4798 for each sample to ensure comparability. In addition, beta-diversity metrics, including weighted UniFrac, unweighted UniFrac, Jaccard distance, and Bray-Curtis dissimilarity, were also calculated to evaluate the diversity between samples.

Taxonomic assignments of the ASVs were conducted using the classify-sklearn naive Bayes taxonomy classifier within the feature-classifier plugin, employing the SILVA Release 132 database. The resulting taxonomic assignments provide information about the taxonomic identity of the ASVs, aiding in the characterization of the microbial composition.

Metagenomics analysis was performed using PICRUSt2: Metagenomics analysis was performed using PICRUSt2, which stands for "Phylogenetic Investigation of Communities by Reconstruction of Unobserved States. This method involves imputing functional profiles based on phylogenetic information. It allows for predicting microbial communities' gene content and functional potential based on 16S rRNA gene sequencing data.

The PICRUSt script was used to normalize the copy number and reckon the variations in the copy numbers of each operational taxonomic unit (OUT) of 16S rRNA gene. The data for this purpose was used from U.S Department of Energy Joint Genomic Institute's Integrated Microbial Genomes (IMG) database. Functional predictions, specifically KEGG orthology annotations, were performed using the PICRUSt2 scripts such as predict_metagenomes.py and metagenome_contributions.py.

To address differences in read numbers across samples, the output representations from PICRUSt2 were adjusted by normalizing occurrences per 100,000 reads. This normalization allowed for more accurate comparisons and assessments of functional potential within the microbial communities.

Bioinformatics analysis

The sequence data analysis primarily involved using QIIME2 and R packages (v3.2.0). To assess alpha diversity at the ASV-level, various indices were calculated using the ASV table in QIIME2, including, Shannon diversity index (Shannon 1948), Simpson index (Simpson,1949), Good's coverage (Good, 1953), Pielou's evenness (Pielou,1966), Chao 1 richness (Chao, 1984) and Faith's phylogenetic diversity (Faith, 1992). ASV-level ranked abundance curves were generated to compare the richness and evenness of ASVs across different samples.

Beta diversity analysis was conducted to evaluate the structural diversity among microbial communities in tobacco samples from different habitats. Jaccard metrics, Bray-Curtis metrics, and UniFrac distance metrics were utilized for this analysis. The results were visualized through principal coordinate analysis (PCoA) and nonmetric multidimensional scaling (NMDS) hierarchical clustering. Furthermore, a principal component analysis (PCA) was performed using genus-level compositional profiles.

To compare taxa abundances at the family level among samples or groups, statistical analysis was conducted using MetagenomeSeq. The results were represented as Manhattan plots, visualizing significant differences in taxonomic abundances (Zgadza *et al*., 2016).

R programe was used to analyze the relationship between conventional flavor components and neutral behavioral components in tobacco leaves and microorganisms in different growing environments.

Overall, these analyses provided insights into the alpha and beta diversity of microbial communities in tobacco samples and the taxonomic composition at different levels, facilitating the understanding of microbial dynamics and differences among the samples.

Results

Production of Conventional Chemical Components in Tobacco Leaves Grown in Different Habitats. It is evident from (Table 2) that there is a wide range of variations in the chemical production of K326 tobacco leaves sourced from different habitats.

The chlorine and nicotine variation coefficients were recorded as the highest (55.49%) and lowest nitrogen variation coefficient was reckoned to be the highest (75.49%) and lowest (13.99%), respectively, compared with other chemicals among middle leaves.

The tobacco leaves obtained from different habitats displayed distinct chemical compositions; for example, the highest total sugar and reducing sugar content in upper leaves were recorded as 36.06% and 28.07%, respectively, whereas in middle leaves, the highest total sugar and reducing sugars content were reckoned as 38.19% and 30.45% respectively. The upper and middle leaves were obtained from Guizhou. The sugar content of tobacco leaves obtained from Sichuan Liangshan Huidong, Fujian Sanming, and other areas was ≤ 30%. On the other hand, upper leaves obtained from Honghe produced the lowest total sugar (23.45%) and reduced sugar (18.63%) content.

The highest total nitrogen contents (2.70%) were recorded in upper leaves sourced from Honghe, Yunnan, and the lowest (1.86%) in leaves obtained from Sanming, Fujian. On the other hand, middle leaves originating from Nanyang, Henan demonstrated highest total nitrogen content (2.46%), whereas leaves obtained from Tongren, Guizhou displayed the lowest (1.79%). The tobacco leaves obtained from Yunnan and Henan exhibited strong flavor and strong vigor, which is actually due to their higher total nitrogen content (Shi *et al*., 1998).

The order of nicotine content recorded in upper leaves grown under different habitats is shown as follows: Honghe > Chenzhou > Yuxi > Tongren > Liangshan > Sanming > Nanyang. Whereas the order of nicotine content recorded in middle leaves grown under different habitats is given as follows: Honghe > Nanyang > Yuxi > Tongren, Chenzhou > Sanming > Liangshan.

The results of the chemical composition ratio depicted that the coefficient of variation of K / Cl ratio was 64.92% and a minimum 60.61% for upper and middle leaves, respectively. The more significant difference in the sugar-base ratio indicates the production of high-quality tobacco leaves (Wang, 2011).

Production of Key NACs in tobacco leaves grown in different habitats: The key neutral aroma compounds (NACs) of tobacco leaves directly affect the smoking quality of tobacco leaves (Table 3) (Yu *et al*., 2009). The brown reaction products formed from the reactions between amino acids and sugars have both aroma and color; therefore, the quality of aroma also depends upon these substances (Wang *et al*., 2015).

It can be seen in Table 3 that the contents of linalool, phenylethanol, damascene, and geraniol in upper leaves obtained from Yuxi, Honghe, and Chenzhou were relatively higher than in other areas; particularly, the content of damascone was significantly higher compared with other areas. On the other hand, Tongren tobacco leaves produced significantly higher furfural than other areas. The tobacco leaves from Liangshan and Nanyang produced the lowest key neutral aroma compounds.

The middle leaves grown in Tongren, Nanyang, and Yuxi areas exhibited higher contents of benzylalcohol, phenylethanol and phenylacetaldehyde than the leaves from other areas, particularly the content of benzyl alcohol (2 mg / mL) were highest among leaves from different areas (Table 3). The contents of Damascusone in tobacco leaves obtained from Liangshan and Nanyang were higher than those of other key neutral aroma compounds compared with leaves from different areas.

Composition of bacterial community on the tobacco leaves: This study investigated the bacterial community composition on K326 tobacco leaves cultivated in seven distinct habitats using high throughput 16S rDNA sequencing. Alpha-diversity analysis was performed, and rarefaction plots demonstrated that species diversity reached saturation at an OTU level with a 0.5% abundance threshold (Fig. 1). This indicates that even samples with lower read depths were adequately sampled, providing reliable measurements of species diversity. Notably, leaf samples from Honghe, Sanming, Liangshan, and Nanyang exhibited higher overall species diversity than those from Yuxi, Tongren, and Chenzhou.

The generated plots illustrate the estimated abundance of Operational Taxonomic Units (OTUs) in subsamples that were progressively rarefied from the original dataset (n = 50). Error bars represent the standard deviation. These alpha-rarefaction plots were created using QIIME, utilizing the observed species metric to estimate the within-sample alpha diversity. The plots provide insights into the diversity of OTUs present in the samples as the subsample size increases.

Nonmetric multidimensional scaling (NMDS) was employed to construct a distance matrix to evaluate beta diversity across the different samples (Fig. 2). The principal components analysis (PCA) of the distance matrix revealed that the distance between replicates of the same product was significantly smaller than the distance between different products. This suggests that the bacterial communities on tobacco leaves from the same habitat displayed greater similarity than those from other habitats. These findings provide insights into the variation and similarity of bacterial community compositions among different habitats and replicate samples within the same habitat.

It visually displayed the variability of replicates, samples, and types, allowing for examining differences and similarities between them.

Table 2. Percentage of chemicals produced in K326 tobacco leaves sourced from different habitats

Chemicals	k	R	F	C	P	I	H	avg1	CV_1	L	Q	E	D	O	J	G	avg2	CV_2
Total sugar	33.02	27.46	24.3	31.49	32.11	23.45	36.06	29.69	15.9	37.2	22.02	31.56	31.62	35.9	27.99	38.19	32.06	17.85
Reducing sugar	23.9	23.81	21.9	28.41	26	18.63	28.07	24.38	14.23	26.01	19.21	26.94	27.95	24.72	21.31	30.45	25.22	15.36
Total nitrogen	2.46	2.06	2.14	1.86	2.36	2.7	2	2.22	13.19	1.89	2.46	1.87	1.85	2.25	2.42	1.79	2.07	13.99
Nicotine	2.87	2.69	3.56	2.85	3.15	3.77	3.09	3.14	12.57	2.11	3.34	2.64	2.56	2.71	3.65	2.64	2.8	18.43
Potassium	1.68	1.53	1.95	2.19	1.34	2.1	1.75	1.79	17.14	2.2	1.56	2.19	2.4	1.63	2.39	1.88	2.03	17.07
Chlorine	0.52	1.02	0.22	0.58	0.78	0.41	0.2	0.53	55.49	0.11	0.84	0.27	0.29	0.23	0.22	0.25	0.31	75.49
Nitrogen-nicotine ratio	0.86	0.77	0.6	0.65	0.75	0.72	0.65	0.71	12.2	0.9	0.74	0.71	0.72	0.83	0.66	0.68	0.74	11.33
Sugar-nitrogen ratio	8.33	8.85	6.15	9.97	8.25	4.94	9.08	7.93	22.23	12.33	5.75	10.2	10.92	9.12	5.84	11.53	9.38	28.24
Potassium-chlorine ratio	3.23	1.5	8.86	3.78	1.72	5.12	8.75	4.7	64.92	20	1.86	8.11	8.28	7.09	10.86	7.52	9.1	60.61

*The upper leaves of the K326 variety were labeled as K, R, F, C, P, I, and H ; the middle leaves were designated as L, Q, E, D, O, J, and G.

Table 3. Content of key NACs produced in K326 tobacco leaves sourced from different habitats.

Components (mg/ml)	k	R	F	C	P	I	H	avg1	L	Q	E	D	O	J	G	avg2
Furfuryl alcohol	-	0.0688	-	-	0.1715	0.1208	0.3946	0.1889	-	0.0642	0.0231	0.1568	-	-	0.2938	0.1345
Furfural	0.2142	0.4273	0.3765	0.8897	0.7200	0.4246	1.5247	0.6539	0.5044	0.5169	0.3671	0.5586	0.1770	0.2812	2.6145	0.7526
Benzaldehyde	0.0275	0.0516	0.0626	0.0475	0.0605	0.0479	0.0292	0.0467	0.0409	0.0839	0.0304	0.0445	0.0250	0.0563	0.0719	0.0520
Benzylmethanol	0.5861	1.0980	1.1255	0.7440	1.0669	1.5040	0.5442	0.9527	1.2491	1.7220	0.6965	0.4775	0.6041	2.1459	1.3891	1.1725
phenylacetaldehyde	0.1144	0.2012	0.2737	0.1654	0.2841	0.1237	0.1048	0.1810	0.1594	0.2767	0.1562	0.1195	0.0843	0.1461	0.2621	0.1742
2-acetylpyrrole	0.0358	0.0462	0.1028	0.0767	0.1279	0.0343	0.0477	0.0673	0.0850	0.0627	0.0700	0.0496	0.0499	0.0275	0.1855	0.0742
Linalool	0.0416	0.0391	0.0790	0.0442	0.0985	0.0972	0.0321	0.0617	0.0561	0.0505	0.0277	0.0251	0.0295	0.0466	0.0505	0.0383
Phenylethanol	0.1760	0.3070	0.4979	0.2930	0.3689	0.4150	0.1483	0.3152	0.3248	0.5391	0.1906	0.1828	0.1782	0.3849	0.3412	0.3028
4-Oxyisofluorone	0.0021	0.0037	0.0063	0.0050	0.0044	0.0035	0.0017	0.0038	0.0038	0.0053	0.0024	0.0031	0.0019	0.0027	0.0068	0.0037
2-Methyltetrahydrofuran-3-one	-	-	0.1051	0.1664	0.1116	-	0.1510	0.1335	-	-	0.0439	0.0728	-	-	0.2754	0.1307
2-acetylfuran	0.0010	0.0005	0.0022	0.0025	0.0030	0.0010	0.0046	0.0021	0.0007	0.0007	0.0004	0.0013	-	0.0003	0.0073	0.0020
5-Methylfuranaldehyde	0.0006	0.0005	0.0031	0.0131	0.0036	0.0013	0.0469	0.0099	0.0006	0.0017	0.0005	0.0109	-	0.0004	0.0665	0.0160
Farnesyl acetone	0.0270	0.5156	0.0992	0.0551	0.0987	-	0.0746	0.1450	0.0801	0.1092	0.0673	0.0693	0.0298	0.0452	0.0675	0.0647
geraniol	0.0094	-	0.0269	0.0157	0.0288	0.0255	0.0066	0.0188	0.0132	-	0.0023	0.0092	0.0067	0.0111	0.0106	0.0080
Damascusone	0.3974	0.6541	1.1236	0.6434	1.0101	0.7719	0.4746	0.7250	1.0243	1.1060	0.8823	0.8017	0.6351	0.6647	0.9401	0.8383

* "- "Indicates data not detected

*The upper leaves of the K326 variety were labeled as K, R, F, C, P, I, and H; the middle leaves were designated as L, Q, E, D, O, J, and G.

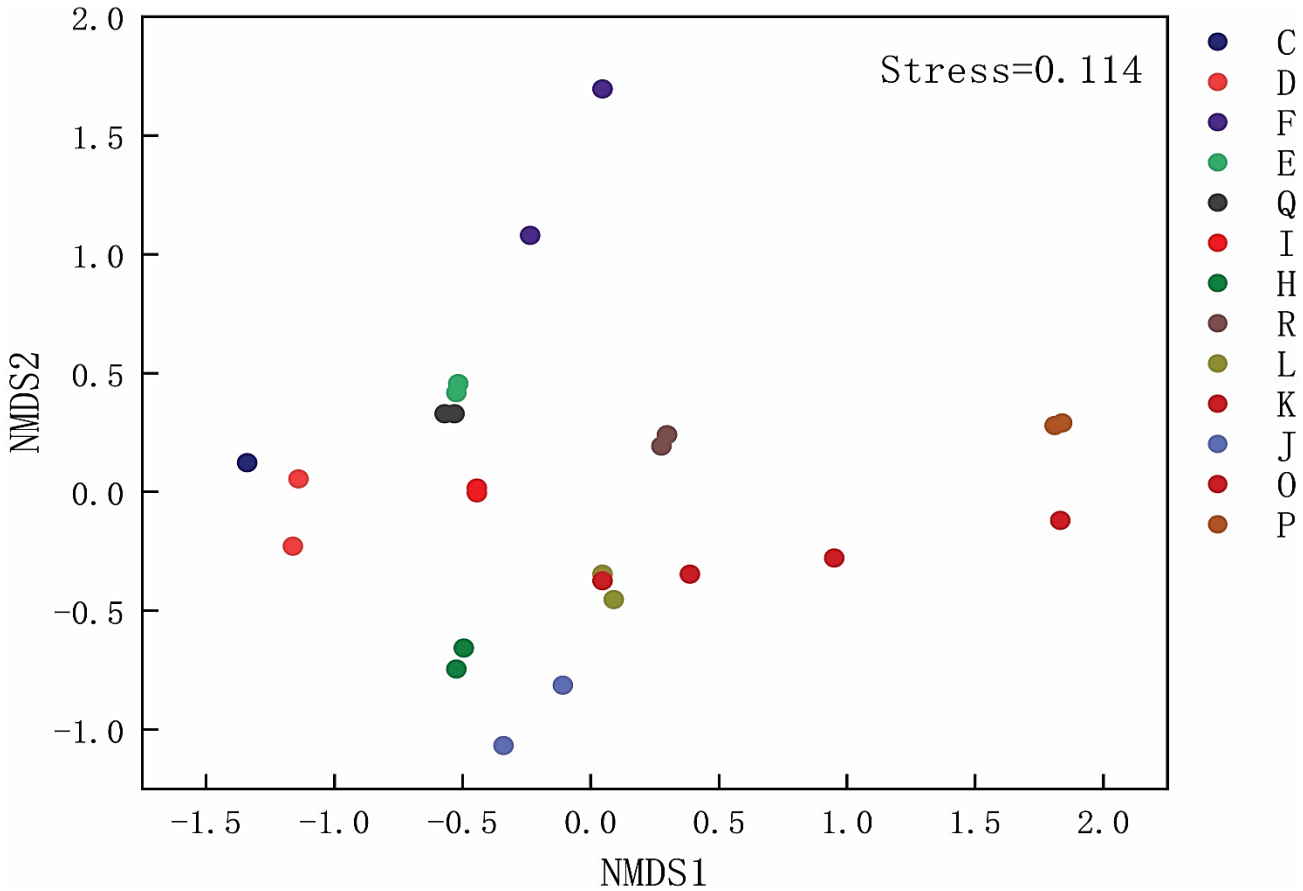

Fig. 1. Alpha-diversity measurement by using QIIME.

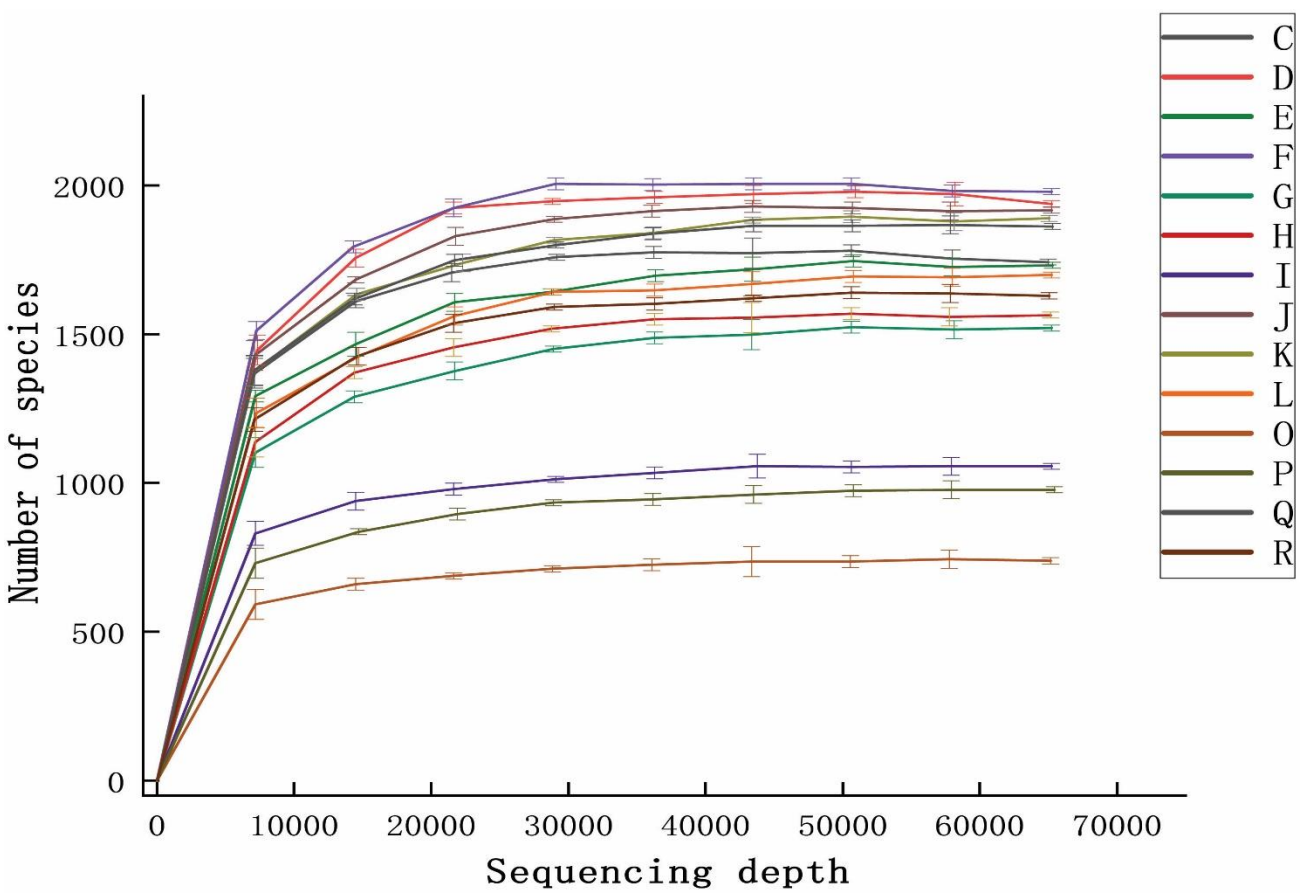

Fig. 2. NMDS principal component analysis.

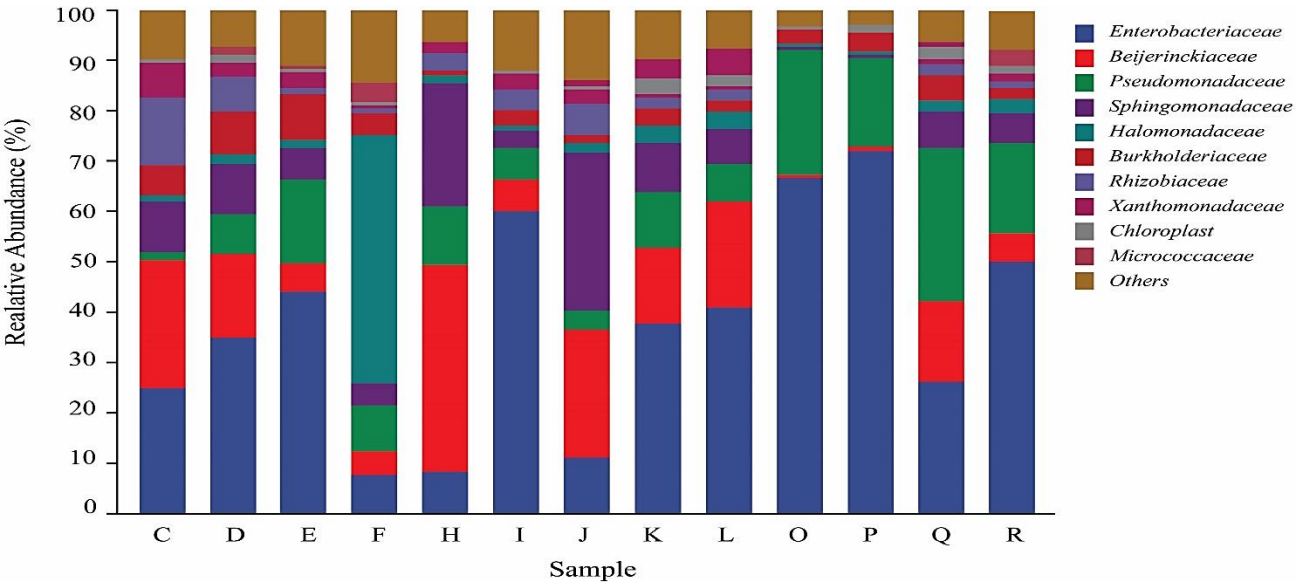

Fig. 3. Relative abundance of taxonomy at the family level.

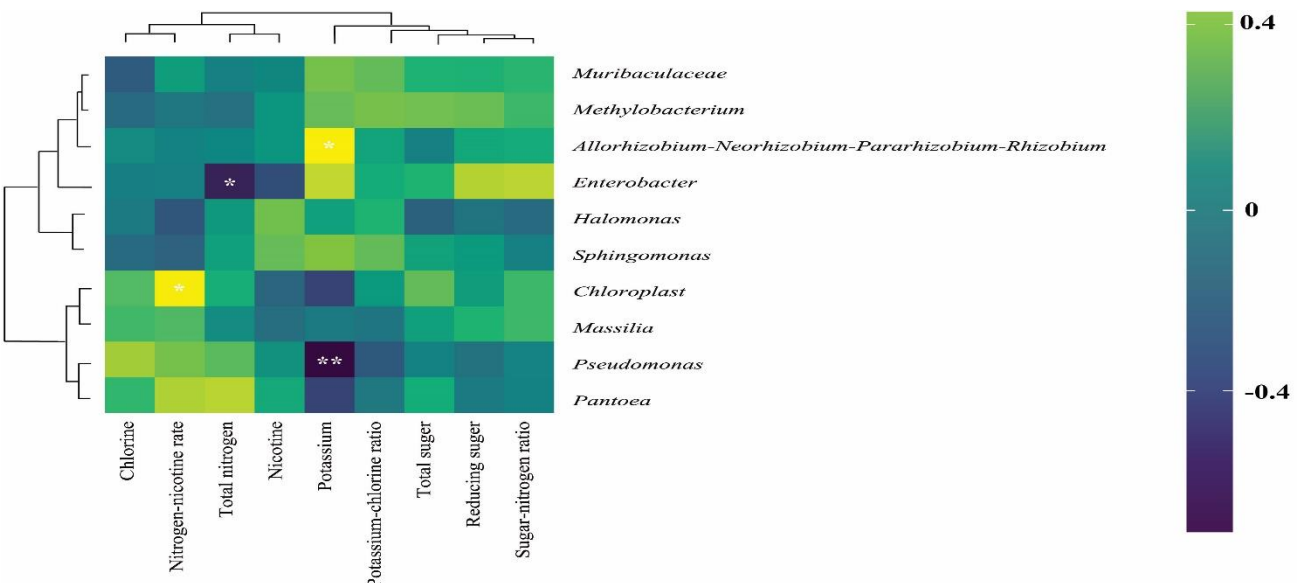

Fig. 4. Cluster heat map of microorganisms and conventional chemical compounds of K326 tobacco leaves from different habitats.

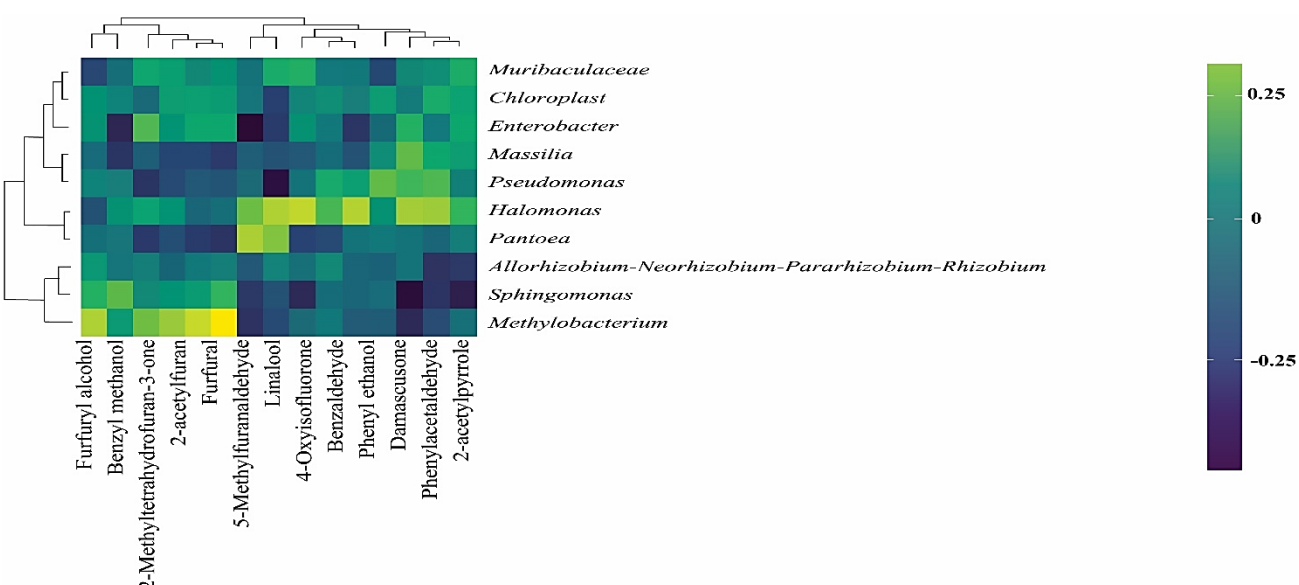

Fig. 5. Cluster heatmap of microorganisms and key NACs of K326 tobacco leaves sourced from different habitats.

Relative abundance of microbial communities: The F samples exhibited a wide range of taxonomic diversity at the family level. In this study, we identified a total of 104 bacterial families belonging to four phyla-*Actinobacteria* (1 family), *Firmicutes* (2 families), *Proteobacteria* (97 families), and *Bacteroidetes* (4 families) with an abundance of at least 0.1% in at least one sample.

In the C samples, the number of families ranged from 59 to 61. The D samples exhibited a range of 85 to 94 families. The E samples displayed a range of 71 to 81 families. The G samples showed a range of 75 to 88 families. The H samples ranged from 77 to 80 families. The I sample had a range of 39 to 41 families. The J samples exhibited a range of 69 to 73 families. The K samples displayed a range of 76 to 88 families. The L samples ranged from 74 to 79 families. The O samples exhibited a range of 29 to 31 families. The P samples displayed a range of 31 to 37 families. The Q samples ranged from 68 to 71 families. Finally, the R samples exhibited a range of 81 to 99 families.

Among the 15 products analyzed in this study, only three families-*Enterobacteriaceae*, *Beijerinckiaceae*, and *Pseudomonadaceae*-were found to be consistently present at an abundance of 0.1% or higher in all samples (Fig. 3).

The O and P samples exhibited a relatively low number of bacterial families, with the majority of reads (>95% in moist samples) dominated by *Proteobacteria*, specifically *Enterobacteriaceae*, *Pseudomonadaceae*, *Burkholderiaceae*, *Chloroplast*, and *Sphingomonadaceae*. In contrast, the analysis of D samples revealed greater diversity, with communities consisting of multiple phyla, primarily *Proteobacteria*, *Bacteroidetes*, and *Actinobacteria*. Among the D samples, five bacterial families were found to predominate: *Enterobacteriaceae* (>13%), *Staphylococcaceae* (>14%), *Burkholderiaceae* (>12%), *Beijerinckiaceae* (>12%), and *Rhizobiaceae* (>11%).

Several bacterial genera, such as *Bacillus*, *Corynebacterium*, *Staphylococcus*, *Pseudomonas*, and *Tetragenococcus*, identified in this study have also been previously reported in tobacco leaves. These genera are primarily associated with soil and plants, although some of them can also act as opportunistic pathogens. The diversity of microbial species residing in K326 tobacco leaves is influenced by various factors, including predominant soil bacteria, microbial populations inherited from seeds/ seedlings, human-associated microbes introduced during harvesting and preproduction stages, resident microbial populations in processing environments (including curing and aging facilities), bacterial populations in fermentation vats, and potentially added fermentation bacteria by manufacturers. Additionally, variations in temperature and humidity across different altitudes and latitudes during tobacco growth also impact microbial diversity (Du *et al*., 2023, Song *et al*., 2024).

While our primary focus has been on bacterial families, it is important to note that many OTUs can be identified at the genus or species level. The bacteria identified in this study primarily represent the bacterial species on K326 tobacco leaves across seven habitats.

Relationship between microorganisms and conventional chemical components of K326 tobacco leaves grown under different habitats: The result of cluster analysis displaying the relationship between microorganisms and conventional chemical compounds in 87 tobacco leaves grown under seven different habitats is shown in (Fig. 4).

The transformation of metabolites usually accompanies the progression of microbial communities. The correlation analysis can preliminarily predict the functional potential of microbial communities. The results showed a positive correlation between reducing sugar content and *Enterobacter*, which may contribute directly or indirectly to the formation of carbohydrates. According to the literature, *Enterobacter* contains endoglucanase-related genes that encode enzymes necessary for the degradation of cellulose (Chen *et al*., 2014). The chloroplast and nitrogen-nicotine ratios were significantly positively correlated ($p<0.05$) in *Enterobacter*, due to which it may have a higher ability to utilize nitrogenous substances and consequently reduce the irritation of tobacco leaves. (Chen *et al*., 2024). *Pseudomonas* and *Pantoea* had a significant impact on sugar content (Cui *et al*.,2020). In contrast, *Methylobacterium*, *Muribaculaceae*, and *Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium* had a significant impact on nitrogen-nicotine ratio, total nitrogen, and chlorine content (Zheng, 2021, Wu, 2024).

Relationship between microorganisms and Key NACs of K326 tobacco leaves grown in different habitats: The result of cluster analysis shows the relationship between microorganisms and key NACs of K326 tobacco leaves sourced from different habitats is shown in (Fig. 5).

From figure 5, it is evident that the production of specific type of NAC in tobacco is strongly influenced by the microbiota living on its leaves. For example, *Methylobacterium* strongly influences the production of 5-Methylfuranaldehyde and furfural in tobacco. Through whole-genome sequencing results demonstrated that *Methylobacterium* express the enzymes required for the production of pyrroloquinoline quinine (precursor of NACs) (Zhi *et al*., 2011; Zhao *et al*., 2020). A positive correlation has been discovered between the abundance *of Halomonas* abd the production of NACs, e. g. 4-Oxyisofluorone, geraniol, phenylethanol and damascene. Damascene and its derivatives have a sweet, rose-like aroma that adds sweetness to the smoke, while phenyl ethanol has an elegant floral scent (Lai *et al* 2022). Confirmed that *Halomonas* use aromatic acids as the sole carbon source for their growth and explained how aromatic acid degradation pathways lead to the production of key NACs.

Discussion

Growing environment and tobacco quality: The quality and flavor of K326 tobacco leaves are strongly influenced by the chemical composition and microbiome of the leaves, which are, in turn, affected by the growing environment. This study highlights the significant impact that different habitats have on both the chemical profile and microbial communities of K326 tobacco leaves. Our analysis shows considerable variation in the chemical components of tobacco leaves grown in different regions. For instance, the

highest total sugar and reducing sugar content were found in tobacco leaves from Guizhou, while the lowest were from Honghe. This variation can be attributed to differences in soil composition, climate, and altitude, which influence nutrient uptake and metabolism in the tobacco plants (Jiang *et al*., 2017).

The nicotine content also varied significantly, with the highest levels recorded in upper leaves from Honghe and middle leaves from Nanyang. This aligns with previous studies that suggest altitude and climatic conditions significantly affect nicotine levels, with higher altitudes generally resulting in lower nicotine content due to cooler temperatures and reduced oxygen availability (Du *et al*., 2023).

Microbiome and tobacco quality: The microbiome plays a crucial role in shaping the chemical profile of tobacco leaves. Our study identified several bacterial genera, such as Allorhizobium, Neorhizobium, Pararhizobium, and Rhizobium, that positively correlate with potassium and nitrogen-nicotine contents (Chen *et al*., 2024). These bacteria are known to enhance nutrient uptake and metabolism, thus influencing the production of key chemical compounds (Cui *et al*., 2020). In contrast, *Methylobacterium*, *Muribaculaceae*, and *Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium* had a significant impact on nitrogen-nicotine ratio, total nitrogen, and chlorine content (Zheng, 2021, Wu, 2024).

Microbial diversity and tobacco quality: Alpha-diversity analysis revealed higher species diversity in tobacco leaves from Honghe, Sanming, Liangshan, and Nanyang compared to those from Yuxi, Tongren, and Chenzhou. Greater microbial diversity is often associated with better soil health and plant resilience, which can positively impact tobacco leaf quality (Jia et al., 2022). The beta-diversity analysis further indicated that the bacterial communities in tobacco leaves from the same habitat were more similar to each other than to those from different habitats, underscoring the strong influence of local environmental conditions on microbial composition (Liu, 2024).

Conclusions

The habitats strongly regulate microbiome composition, consequently controlling the production of conventional chemicals and NACs in tobacco. The microbiome is a critical component of tobacco's ageing ecosystem and plays a vital role in producing specific compounds in specific ratios. Generally, the contents and ratios of chemical compounds produced in different tobacco varieties vary. This study observed that the coefficient of variation for reducing sugar content was the highest, while it was the lowest for total nitrogen content. Most studies have suggested that the water-soluble total sugar is the main factor that determines the sweetness and alcohol content of the flue gas. On the other hand, total nitrogen and nicotine content reflect the physiological strength and concentration of flue gas produced by tobacco. The sugar-base and nitrogen-base ratios are important indicators used to evaluate the acid-base balance of flue gas. They are usually used to evaluate the flexibility and fineness of flue gas. Generally, the ratios of sugar-base and nitrogen-base are 8–10 and 0.9–1.0, respectively. The ratio of potassium-chloride in K326 tobacco leaves sourced from the abovementioned habitats was greater than 4, which aligns with the characteristics of high-quality tobacco leaves. The tobacco with a specific flavor and taste can be customized through the growth of K326 tobacco in selective habitats suitable for specific microorganisms, which will produce specific aroma compounds in a defined ratio. Tobacco aroma is the general term used for the aroma emitted by tobacco leaves, and the odor and taste generated after combustion due to the combined action of various compounds with specific aroma characteristics. The high-quality tobacco leaves give off a pure aroma and good taste during combustion. Aroma compounds have received much attention from tobacco scientists and technology workers because they are important indices used to evaluate the quality of raw tobacco and its products. Currently, tobacco research mainly focuses on the influence of aroma compounds on the smoking quality, the comparison of genotypes, differences in different ecological regions, and the regulation of agronomic measures.

In this study, the effect of habitats on the composition of the microbiome and, consequently, on the production of conventional chemicals and key NACs in K326 tobacco leaves was explored, and the sensory quality of K326 tobacco leaves was linked to the different habitats and composition of the microbiota.

Acknowledgment

The fund for the research work comes from the National Natural Science Foundation of China, whose fund number is 81903507.

References

Bolyen, E., J.R. Rideout, M.R. Dillon, N.A. Bokulich and J.G. Caporaso. 2018. QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. *Peer J.*, 12(11): 89-99.

Chao, A. 1984. Nonparametric estimation of the number of classes in a population. *Scand. J. Stat.*, 11(9): 54-59.

Chen, S., Q. Chen, J. Liu, Y. Chen, J. Fan, X. Deng, J, Li and J. Tao. 2024. Community structure of microorganisms on the surface of different alcoholized cigarette raw materials and their degradation function of macromolecular substances. *Tobacco Science & Technology*, 57(03): 34-42.

Chen, Y., J. Pei, Y. Guo, L. Lin, R. Huang and H. Pang. 2014. Cloning and functional verification of an endoglucanase Cel8A from *enterobacter cloacae*. *J. Nanjing Norm. Univ.*, 8(1): 5-11.

Cui, Y., Z. Jin and Y. He. 2020. Carbon source affects the biosynthesis of Garcinia camperin PA1201 in Pseudomonas rhizosphere in rice. Microbiology China, 47(08): 2399-2408.

Du, X., Y. Hu, M. Gao, W. Guo, L Wang and A. Wang. 2023. Analysis of the characteristics and relationship between meteorological factors and chemical composition of tobacco leaves in the main flue-cured tobacco producing areas of Shaanxi. *Acta Agriculturae Boreali-Occidentalis Sinica*, 32(05): 815-823.

Faith, D.P. 1992. Conservation evaluation and phylogenetic diversity. *Biol. Conserv.*, 61(1): 1-10.

Fei, Y., G. Liu, H. Shi and X. Yang. 2013. Interactive effects of light intensity and nitrogen supply on the neutral volatile aroma components and organic acids of flue-cured tobacco. *Int. J. Food Sci. Technol.*, 11(1): 1187-1194.

Good, I.J. 1953. The population frequencies of species and the estimation of population parameters. *Biometrik.*, (3-4): 237-264.

Jenkins, R.A., B. Tomkins and M.R. Guerin. 2000. The chemistry of environmental tobacco smoke: composition and measurement. *J. Neurosci*., 35(3): 960-971.

Jia, H., J. Feng, T. Wu, X. Zhang, T. Sun, X. Hu, P. Wang, Z. Li, Y. Cheng, C. Cheng and Z. Li. 2022. Effects of biochar bacteria on tobacco black shank disease and soil microorganisms. *Chinese Tobacco Science*, 43(06): 68-75.

Jiang, J., Y. Lu, Y. Su, L. Pan, X. Tang, X. Chen, F. Niu, L. Zhao, M. Lu and S. Zhu. 2017. Chemical composition characteristics and availability evaluation of flue-cured tobacco in major tobacco producing areas in China. *Zhongguo Yan Cao Xue Bao*., 23(02): 13-27.

Jian-Jun, Y., S. Hui-Fang, L. Yan-Fang, P. Tian-He and M. Xin-Ming. 2009. Relationships between ecosystem factors and contents of megastigmatrienones in flue-cured tobacco leaves in Liangshan, Sichuan Province. *Acta Ecol. Sin.*, 32(7): 23-35.

Jiaxi, Z. 2020. Characterization of the core microbiome in tobacco leaves during aging. *Microbiol. Open*, 9(3): e984-e990.

Lai, X., H. Ling, N. Zhou and Y. Xu. 2022. Isolation and culture of *Halomonas* sp. NyZ771 from abyssal sediments and its degradation of aromatic acids. *Microbiol.*, 49(5): 10-17.

Liu, J. 2015. Nicotine-degrading microorganisms and their potential applications. *Appl. Microbiol. Biotechnol*., 99(9): 3775-3785.

Liu, T. 2024.Study on microbial diversity of tobacco rhizosphere soil and antagonistic microorganisms of Tobacco black shank. Hunan Agricultural University.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *Embnet. J*., 17(1): 17-25.

Masugi, E., R. Kindt, K. Mitsu, K. Sandra and A. Ortiz. 2016. Lipidomics of tobacco leaf and cigarette smoke. *J. Chrom. A.I. Ectroph. & Other Sep. Methods*., 75(10): 118-129.

Pielou, E.C.J. 1966. The measurement of diversity in different types of biological collections. *J. Theor. Biol*., 13(9): 131-144.

Segata, N., J. Izard, L. Waldron and D. Gever. 2011. Metagenomic biomarker discovery and explanation. *Genom. Biol*., 12(6): R60-70.

Shannon, C.E. 1948. A mathematical theory of communication. *Bell Syst. Tech. J*., 27(4): 623-656.

Shi, H., J. Han, G. Liu and Y. Wang. 1998. Study on relationship between carbon-nitrogen metabolism and flavor of flue-cured tobacco. *Acta Bot. Sin.*, 7(2): 8-19.

Simpson, E.H. 1949. Measurement of diversity. *Nat.*, 163(4148): 688-688.

Song, S., J. Yang, W. Duan, S. Chen, W. Wang, L. Zhang, J. Xu, D. Han, Z. Xu, W. Jia. 2024. Effect of altitude on the growth metabolism and post-roasting quality of tobacco leaves. *Journal of Henan Agricultural Sciences*, 53(03): 53-67.

Villanti, A.C., A. Richardson, D.M. Vallone and J.M. Rath. 2013. Flavored tobacco product use among U.S. young adults. *Amer. J. Prev. Med.*, 44(4): 388-391.

Wang, J. 2011. Study on the effect of sugars on sweet taste in cigarette products. *Ocean Univ. China*., 9(11): 34-41.

Wang, Y., J. Chu, B. Xu, H. Liu, C. Du, S. Wang, T. Liu, B. Xin and J. Qiu. 2015. Changes of maillard reaction products during natural alcoholization of flue-cured tobacco and its relationship with sensory quality. *Chin. Tob. Sci.*, 36(4): 6-14.

Weeks, W.W., V.A. Sisson and J.F. Chaplin. 1992. Differences in aroma, chemistry, solubilities, and smoking quality of cured flue-cured tobaccos with aglandular and glandular trichomes. *J. Agric. Food Chem.*, 40(10): 57-67.

Wu, B. 2024. Mechanism of high nitrogen promotion of bacterial wilt outbreak in tobacco. Southwest University.

Xiao, D.S., Y.L. Yan, L.J. Wang, D. Wu, S.J. Zheng, W.M. Wang, Y.G. Tong, Y. Wang and H.Z. Jian. 2016. Capturing tobacco specific N-nitrosamines (TSNA) in industrial tobacco extract solution by ZnO modified activated carbon. *Microporous Mesoporous Mater.*, 222(9): 160-168.

Yang, Y., J. Zhang, J. Wang, X. Feng and H. Cui. 2015. Effects of biochar on plastid pigment metabolism in flue-cured tobacco. *Tobacco Sci. Technol.*, 49(12): 8-13.

Yu, J., H. Yang, Q. Bi, H. Wang, G. Li, T. Pang and W. Guo. 2009. Canonical correlation analysis of neutral aroma components and aroma quality of flue-cured tobacco. *J. Gansu Univ. Technol*., 44(2): 6-13.

Zgadzaj, R., R. Garridoter, D.B. Jensen, A. Koprivova and S. Radutoiu. 2016. Root nodule symbiosis in Lotus japonicus drives the establishment of distinctive rhizosphere, root, and nodule bacterial communities. *PNAS*., 113(49): 67-78.

Zhang, X., S. Song, K. Hayat and K.M.H. Eric. 2016.Tobacco alkaloids reduction by casings added/enzymatic hydrolysis treatments assessed through PLSR analysis. *Regul. Toxicol. Pharmacol.*, 76(9): 70-80.

Zhao, C., H. Zhang, X. Zhang, Y. Dou and X. Bao. 2020. Whole genome sequencing and comparative genomics analysis of a new pyrroloquinoline quinone methylotrophic strain. *Microbiol.*, 47(12): 14-21.

Zheng, S. 2021. Preparation of microbial fertilizer for straw and its effect on tobacco growth. Dalian University of Technology.

Zhi, J., X. Xiong, J. He, J. Wang, T. He and W. Zhang. 2011. Construction of complete genome map of methylotrophic bacteria MP688. *Biotech. Comm.*, 22(4): 4-11.

Zhu, J., J. Tang, J. Xu, S. Ren, F. Guo, S. Guo and F. Guo. 2024. Relationship between single roasted upper leaves leaf weight and sensory quality of tobacco variety K326. *Journal of Anhui Agricultural Sciences*, 52(11): 163-168+178.

(Received for publication 13 February 2024)