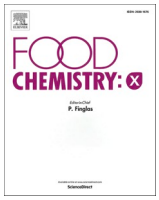
Food Chemistry: X 18 (2023) 100716

Contents lists available at ScienceDirect

Food Chemistry: X

journal homepage: www.sciencedirect.com/journal/food-chemistry-x



Tracing the geographical origin of tobacco at two spatial scales by stable isotope and element analyses with chemometrics

Lili Cui [a,d,e], Huan Chen [a,d], Yuwei Yuan [b,c], Fengpeng Zhu [a], Jing Nie [b,c], Shulei Han [a,d], Ya'ning Fu [a,d], Hongwei Hou [a,d,*], Qingyuan Hu [a,d,*], Zengping Chen [e,*]

[a] China National Tobacco Quality Supervision and Test Center, Key Laboratory of Tobacco Biological Effects, Zhengzhou 450001, China
[b] Institute of Agro-product Safety and Nutrition, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, China
[c] Key Laboratory of Information Traceability for Agricultural Products, Ministry of Agriculture and Rural Affairs of China, Hangzhou 310021, China
[d] Beijing Life Science Academy, Key Laboratory of Tobacco Biological Effects and Biosynthesis, Beijing 100101, China
[e] State Key Laboratory for Chemo/Biosensing and Chemometrics, College of Chemistry and Chemical Engineering, Hunan University, Changsha 410082, China

ARTICLE INFO

Keywords:
Geographical origin
Stable isotopes
Elements
OPLS-DA modeling
Tobacco

ABSTRACT

Tobacco is a widely cultivated cash crop, but it is often smuggled and sold illegally. Unfortunately, there is currently no way to verify the origin of tobacco in China. In an effort to address this issue, we conducted a study using stable isotopes and elements from 176 tobacco samples at both provincial and municipal scales. Our findings revealed significant differences in $\delta^{13}C$, K, Cs, and $^{208/206}Pb$ at the provincial-level, and Sr, Se, and Pb at the municipal level. We created a heat map at the municipal level, which showed a similar cluster classification to geographic grouping and provided an initial assessment of tobacco origins. Using OPLS-DA modeling, we achieved a 98.3% accuracy rate for the provincial scale and 97.6% for the municipal scale. It is worth noting that the importance of rankings of variables varied depending on the spatial scale of the evaluation. This study offers the first traceability fingerprint dataset of tobacco and has the potential to combat mislabeling and fraudulent conduct by identifying the geographical origin of tobacco.

1. Introduction

Tobacco (*Nicotiana tabacum L.*) is one of the most widespread cash crops globally (Tang et al., 2020; Wu, Tang, Yang, Liu, & Guo, 2013). It is widely consumed for pleasure, such as in edible oils, and utilized as a treatment for postoperative pain (Matthews, Fu, Dana, & Chou, 2016) as well as agricultural pesticides (Simon-Delso et al., 2015). Tobacco has long been cultivated and produced in China. The Southwest Tobacco region, including Yunnan (YN), Guizhou (GZ), and Sichuan (SC) provinces is the most renowned and primary production area, known for the fragrant aroma, golden color, mild effect, and pure taste of the produced tobacco (Li et al., 2019). YN tobacco accounts for 50% of Chinese production, with 750,000 tons of flue-cured tobacco grown on 320,000 ha of land each year (G. Liu & Liu, 2010). There are concerns related to tobacco smuggling, illicit sales, and inferior or counterfeit tobacco entering the marketplace (Nguyen, Giang, & Pham, 2020). The distinction of flue-cured tobacco by appearance and smell only allows criminals to mislabel the origin of tobacco, increasing profits to the detriment of consumers. It is therefore of paramount importance to uncover ways to effectively profile flue-cured tobacco, as well as determine and authenticate its geographical origin.

Recently, stable isotopes combined with multi-element analyses have been used to determine the geographical origins of various agricultural products based on the correlation of isotopes and elements with ecological environments (Opatić, Nečemer, Budič, & Lojen, 2018; C. Wang et al., 2020). Geographical authentication of crops (Wadood, Nie, Li, Rogers, Zhang, & Yuan, 2022; Wang, Chen, Zhang, Zhao, Yang, & Chen, 2020), fruit (Muoz-Redondo, Bertoldi, Tonon, Ziller, & Moreno-Rojas, 2021), tea (Wang et al., 2022; Xia et al., 2021), garlic (Nie et al., 2022), Chinese herbal medicine (Hua et al., 2020), and aquatic products (Luo, Jiang, Chen, Zheng, Liu, & Yang, 2019) has allowed for an essential guarantee of safety and quality. Carbon isotope ($\delta^{13}C$) levels in plants are strongly correlated with temperature and light conditions associated with photosynthesis (Hu, Cheng, & Tao, 2016; Xia et al., 2022). The isotopic composition of nitrogen ($\delta^{15}N$) is impacted by fertilizers and soil contributions (Liu, Liu, Qian, Song, & Yuan, 2020). Hydrogen and oxygen isotopes in crops primarily rely on precipitation and water-use efficiency. The geographical location (altitude, latitude)

* Corresponding authors.
E-mail addresses: qsfctc@163.com (H. Hou), huqy1965@163.com (Q. Hu), zpchen@hnu.edu.cn (Z. Chen).

https://doi.org/10.1016/j.fochx.2023.100716
Received 10 February 2023; Received in revised form 13 May 2023; Accepted 15 May 2023
Available online 18 May 2023

exerts a significant impact on plants. Dansgaard *et al.* reported that $\delta^{2}H$ and $\delta^{18}O$ values decrease alongside increasing latitude, altitude, and distance inland (Dansgaard, 1964). Several key element contents (Liu et al., 2019) and isotope ratios, including $^{207/206}Pb$ (Wang et al., 2020), and $^{87/86}Sr$ (Petrini, Sansone, Slejko, Buccianti, Marcuzzo, & Tomasi, 2015) are assimilated by plants and can serve as efficient traceability tools to identify their geographical origin. Additionally, chemometric statistical analyses, such as Principal Component Analysis (PCA) (Abdi & Williams, 2010) and Orthogonal Partial Least Squares Discrimination Analysis (OPLS-DA) (Boccard & Rutledge, 2013), have been successfully utilized to comprehensively extract feature variables from geographic fingerprint information.

A limited number of papers have been published on stable isotope tracing to identify the geographic origin of tobacco. A previous study (Jamin, Naulet, & Martin, 1997) tentatively interpreted the multi-element and multi-site isotopic analysis of nicotine from tobacco leaves which indicated the influence of environmental factors on nicotine isotopic parameters. Similarly, a combination of $\delta^{2}H$ and $\delta^{15}N$ values may be suitable for distinguishing foreign counterfeits of Canadian cigarettes (Binette, Lafontaine, Vanier, & Ng, 2009). Elemental and isotopic compositions combined with chemometrics have been addressed in one study focused on preserving the quality and geographic traceability of Umbrian tobacco (Bontempo, Bertoldi, Franceschi, Rossi, & Larcher, 2021). To the best of our knowledge, there is currently no published research related to tracing the origin of flue-cured tobacco in China. Additionally, the accuracy of authentication or variable importance of agricultural product discrimination is necessarily related to the spatial scale considered (Xiong et al., 2022). Investigation of more spatial scales is particularly necessary for accurately determining the traceability and authenticity of tobacco.

Here we aim to fill this gap in tobacco research and explore the characterization and classification of stable isotopes and elements in tobacco grown in China. Additionally, we explored authentication and differentiation at provincial and municipal scales. We profiled YN, GZ, and SC as the main producing provinces, as well as five cities in YN: Puer (YNPE), Qujing (YNQJ), Yuxi (YNYX), Wenshan (YNWS), and Chuxiong (YNCX) (Fig. 1 and Table S1), to construct a model usable at both provincial and municipal scales, and to compare accuracy and importance of different variables. This study provides a technical basis for enhancing the quality and traceability of tobacco.

2. Material and methods

2.1. Sample collection and preparation

A total of 176 tobacco samples were collected in August 2020 from YN Province, GZ Province, and SC Province. The YN samples were collected between 21°8′–29°15′ N and 97°31′–106°11′ E (Fig. 1), at cultivation altitudes up to 2500 m. The GZ samples were collected from 24°37′–29°13′ N and 103°36′–109°35′ E, at altitudes ranging from 500 to 1800 m. The sites of SC samples were between 26°3′–29°18′ N and 100°4′–103°53′ E, with altitudes ranging from 1158 to 2758 m. Among the samples, 52 were harvested from GZ, 40 from SC, and 84 from YN. The 84 from YN were further divided into 18 samples from YNPE, 22 from YNQJ, 10 from YNYX, 18 from YNWS, and 16 from YNCX. The samples were numbered according to their locations. The collected

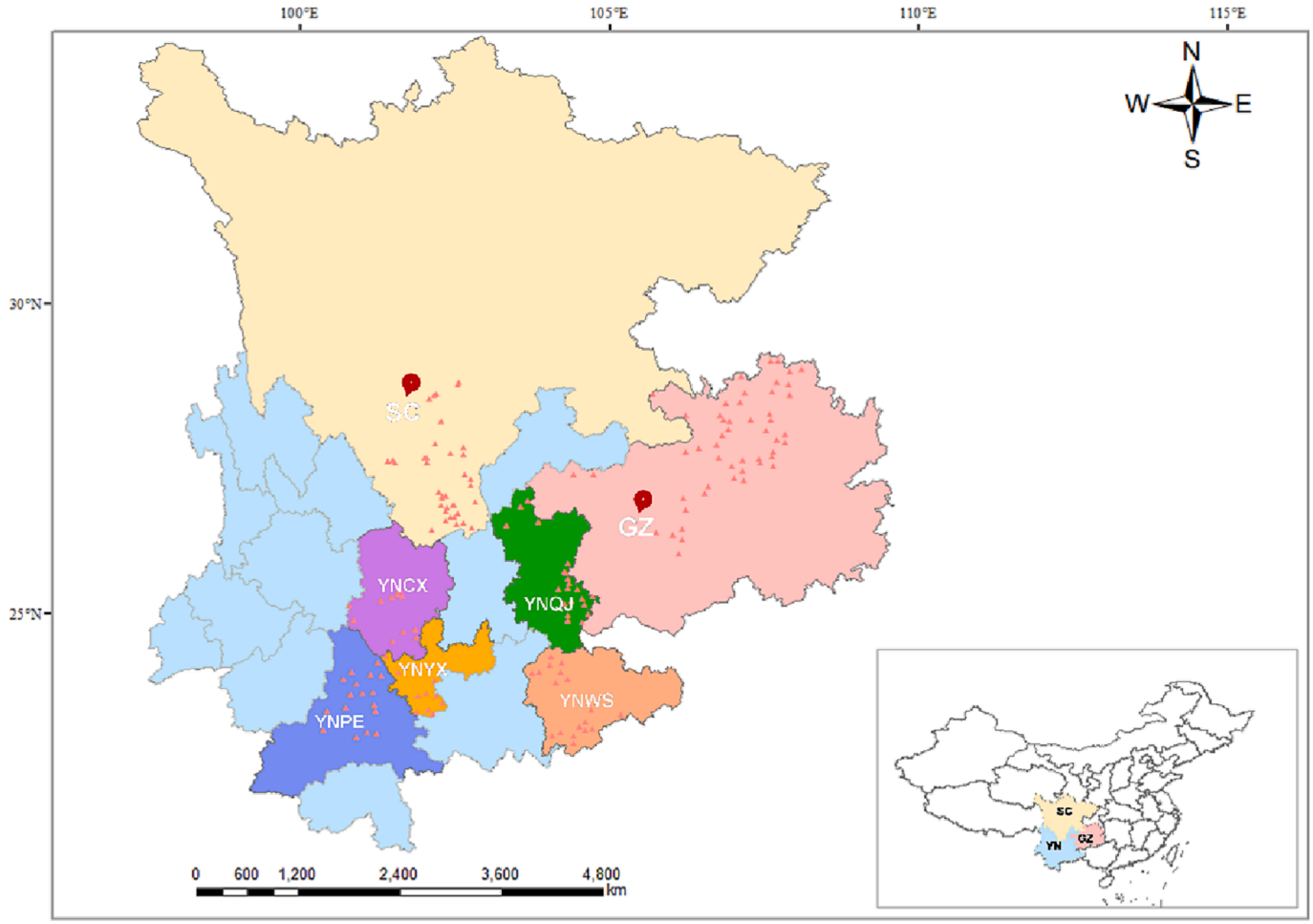


Fig. 1. Sampling regions.

tobacco leaves were initially roasted according to the market requirements for finished tobacco leaves (GB2635-1992) before being ground into a powder with a mortar for preservation. The samples were subsequently oven-dried at 70℃ to constant weight before experimental analysis.

2.2. Stable isotope analysis

All stable isotope analyses were carried out using a Flash 2000 HT elemental analyzer connected to the 253 Plus IRMS instrument (Thermo Fisher, USA) via a GC IsoLink combustion or pyrolysis interface (Thermo Fisher, USA). Approximately 1.5 mg of powdered samples were packed into capsules for analysis of $\delta^{13}C$ and $\delta^{15}N$, and the combustion tubes were heated to 960 °C. Additionally, 0.5 mg of powdered samples were packed into capsules for $\delta^{2}H$ analysis and a high-temperature cracking tube was placed at 1,380 °C for analysis. High-purity helium gas was used as the carrier gas and high-purity CO_2, N_2, and H_2 were used as reference gasses, respectively.

The results of this study are presented in δ notation according to international standards, and the values were calculated as follows:

$$\delta^{i}E = \left(\frac{R_{sample}}{R_{reference}} - 1\right)$$

where $\delta^{i}E$ is the delta value of the sample for element E (H, N, C) with heavier mass number ^{i} in parts per thousand (‰), and R_{sample} and $R_{reference}$ are the molar ratios of heavy to light isotopes, respectively in the sample and the international reference (Nie, Shao, Zhang, Li, & Yuan, 2021). Among them, $\delta^{13}C$ was associated with the international standard Vienna-Pee Dee Belemnite (VPDB); $\delta^{15}N$ was relative to atmospheric AIR; $\delta^{2}H$ was correlated with Standard Mean Ocean Water (SMOW). Reference materials (IAEA-International Atomic Energy Agency, Vienna, Austria) were used to calibrate the reference gas through two-point normalization, including USGS 40 (L-glutamic Acid, $\delta^{13}C_{VPDB\text{-}LSVEC} = -26.39 \pm 0.04$ ‰, $\delta^{15}N_{AIR} = -4.52 \pm 0.06$ ‰) and USGS41a (L-glutamic Acid, $\delta^{13}C_{VPDB\text{-}LSVEC} = +36.55 \pm 0.08$ ‰, $\delta^{15}N_{AIR} = +47.55 \pm 0.15$ ‰), USGS84 (Sicilian olive oil, $\delta^{2}H_{VSMOW\text{-}SLAP} = -140.4 \pm 3.1$ ‰, $\delta^{18}O_{VSMOW\text{-}SLAP} = +26.36 \pm 0.50$ ‰, $\delta^{13}C_{VPDB\text{-}LSVEC} = -28.80 \pm 0.09$ ‰), and USGS86 (Tropical Vietnamese peanut oil, $\delta^{2}H_{VSMOW\text{-}SLAP} = -207.4 \pm 4.5$ ‰, $\delta^{18}O_{VSMOW\text{-}SLAP} = +18.76 \pm 1.03$ ‰, $\delta^{13}C_{VPDB\text{-}LSVEC} = -30.63 \pm 0.09$ ‰). The analytical precision was lower than ± 0.2‰ for C, ± 0.2‰ for N, and ± 2.0‰ for H, respectively.

2.3. Elemental analysis

Around 0.25 g of flue-cured tobacco samples were weighed and digested in a microwave digester (ETHOS ONE, Milestone, Italy) using HNO_3-H_2O_2 (v:v = 6:1, mL) system, and the digested samples were diluted to 50 mL with ultrapure water (≥18MΩ, Milli-Q Element ultrapure water treatment system, Millipore, USA). A mixture of internal standard solutions (10 ng/mL) of Ge (GSB 04-1728-2004), In (GSB 04-1731-2004), and Bi (GSB 04-1719-2004) were prepared using standards purchased from the National Center of Analysis and Testing for Nonferrous Metals and Electronic Materials (Beijing, China) to monitor and check instrumental drift. 30 mineral parameters, and 5 isotope ratios ($^{84}Sr/^{86}Sr$, $^{87}Sr/^{86}Sr$, $^{88}Sr/^{86}Sr$, $^{207}Pb/^{206}Pb$, and $^{208}Pb/^{206}Pb$), were measured using an inductively coupled plasma-mass spectrometer (ICP-MS, Agilent 7900, Agilent Technologies, Tokyo, Japan). The elements were separated into two groups based on their concentrations. This first group of samples was directly injected for analysis, including Li, Cr, Co, Ni, Cu, Ga, As, Rb, Cd, Cs, Ba, Pb (Multi-elemental calibration standard 2A, Agilent Technologies), Sn, Sb (Multi-elemental calibration standard-3, Agilent Technologies), Y, La, Pr, Nd, Sm, Dy, Er, Tl (Multi-elemental Solution 1, SPEX CertiPrep.), while the second group of samples was diluted 10 times, and the elements in this group included Na, Mg, Al, K, Mn, Zn, Se, and Sr (Initial calibration verification standard, Agilent Technologies). The limits of detection (LODs) for the above elements were 2.06, 0.02, 0.01, 0.05, 0.04, 0.01, 0.03, 0.01, 0.004, 0.004, 0.08, 0.03, 0.01, 0.01, 0.003, 0.004, 0.001, 0.008, 0.005, 0.001, 0.002, 0.001, 0.47, 0.10, 0.45, 0.37, 0.02, 0.21, 0.26, and 0.02 μg/kg, respectively. Instrumental reproducibility (RSDs) for concentration was lower than ± 5.0%. The accuracy of spiking experiments was in the range of 85.8%–110% for all elements. A cation exchange column (AG500W-X8200-400 mesh) was used to separate Sr-Rb, followed by ICP-MS analysis. The ICP-MS experimental parameters consisted of 10 acquisition times, 1000 repeated scans, and an integration time of 5 s. The accuracy of the instrument was lower than 0.1%.

2.4. Statistical analysis and chemometrics methods

Univariate analyses were carried out using SPSS (ver. 23.0, IBM, USA) to determine differences within the samples. Regional differences were assessed by one-way analysis of variance (ANOVA) and *post-hoc* Tukey's HSD tests considering origin discrimination, a p-value ≤ 0.05 was considered statistically significant. Hierarchical Clustering Analysis (HCA) was performed using the Multi Experiment Viewer (MeV, ver. 4.9.0) software.

Orthogonal partial least-squares discriminant analysis (OPLS-DA), a supervised recognition tool, is an orthogonal exchange correction based on partial least-squares discriminant analysis (PLS-DA), which is able to filter out noise unrelated to the classification signal and improve the analytical effectiveness of the model. Key characteristic variables used to discriminate the origin of samples were identified according to variable loadings and importance for the projection values (VIPs > 1 indicates the importance). Chemometric analyses (OPLS-DA) were carried out with SIMCA 14.1 software (Umetrics, Umeå, Sweden). The drawing of VIP plots based on stable isotope and element content was accomplished using chiplot software (website address: https://www.chiplot.online/).

3. Results and discussion

3.1. Distribution of isotopes and elements in flue-cured tobacco at a provincial scale

Analytical results of three stable isotopes ($\delta^{13}C$, $\delta^{15}N$, and $\delta^{2}H$) in flue-cured tobacco were compared at provincial scales (Fig. S1). According to ANOVA, there were significant differences ($p < 0.05$) in the $\delta^{13}C$ and $\delta^{15}N$ values obtained from tobacco samples across three different geographical origins, as indicated by the superscripts with different letters. The average $\delta^{13}C$ value of the sample from SC was the highest (−25.5‰), followed by that of YN (−28.1‰) and GZ (−29.3‰). This may be due to cooler weather and higher sunlight intensity in SC, which could lead to preferential ^{13}C uptake during photosynthesis and lower CO_2 respiration (Buchmann, Brooks, Rapp, & Ehleringer, 1996; Evans & Von Caemmerer, 2013). The $\delta^{15}N$ values from YN (0.1 ± 1.2 ‰) were highest, significantly higher than SC samples (−1.9 ± 1.3 ‰), indicating that soil nutrients may have an impact on the nitrogen isotope signature. Mean $\delta^{2}H$ values from YN, GZ, and SC were −113.7 ± 12.7‰, −112.8 ± 6.3‰, and – 121.7 ± 15.6‰, respectively. YN and GZ had higher mean $\delta^{2}H$ values which are likely due to exposure to a more tropical (humid and warm) climate than the SC Province. Distinct climates and altitude factors that covary with geographical characteristics impart significant variations in $\delta^{2}H$ value. Due to its wider production range and altitude range (500–2500 m), YN samples are the focus of analysis into the relationship between altitude differences and hydrogen isotope fractionation. As shown in Fig. S2, $\delta^{2}H$ values significantly increased along with altitude increases. More positive values were found in 2000–2500 m (−110.0‰) followed by 1500–2000 m (−112.7‰), 1200–1500 m (−113.1‰), and more negative values were observed in 500–1200 m (−121.1‰). Our findings were consistent with previous research (Dansgaard, 1964).

Table 1 summarizes the content of 35 elements (including 5 isotope ratios) in 176 tobacco samples at the provincial scale. Tobacco from the three regions had significant differences for most elements ($p < 0.05$), except for Mn, Co, Sn, Pb, and $^{84/86}$Sr. The majority of the tobacco from the YN and SC had higher levels of the considered elements than GZ regions. For example, Na, Cs, and Sr were more abundant in YN tobacco, with mean values of 37.7 mg/kg, 62.3 μg/kg, and 276.5 mg/kg, respectively. Higher levels of Ba and Pr were found in SC samples, with mean values of 35.9 mg/kg and 149.1 μg/kg, respectively. However, GZ tobacco showed higher values of Li and Mg (4.1 mg/kg and 4.5. mg/g). It is worth noting that the concentration of rare earth elements (RREs) in GZ tobacco was generally lower than in tobacco from other regions, likely due to the different distributions of RREs in the growth soils. Previous studies have indicated the essentiality of these parameters in food authentication (Bontempo et al., 2011). Strontium and lead isotopes are important lithologic and mineral indicators, characterized by varied geological and soil conditions, and can be used for geographic traceability of agricultural products (Wang et al., 2020). As shown in Table 1, $^{87/86}$Sr, $^{88/86}$Sr, $^{207/206}$Pb, and $^{208/206}$Pb ratios in tobacco harvested from Guizhou were significantly lower than in YN and SC tobacco ($p < 0.05$).

3.2. Geographical profiling of isotopes and elements at a municipal scale

Similarly, 84 tobacco samples collected from 5 cities in YN province at a municipal scale were analyzed for stable isotopes (Fig. 2). Mean δ^{13}C values for each city ranged from −28.5‰ to −27.1‰, consistent with the range observed in C3 plants (−34‰ ~ −24‰) (Krueger & Reesman, 1982). We found no differences across most cities, with the exception of YNYX. According to the weather conditions in Yunnan Province in 2020, the temperature of YNYX was higher than that of other regions, resulting in a higher δ^{13}C value. Geographical δ^{13}C trends were in alignment with a large-scale investigation of δ^{13}C in leaves across China (MX et al., 2017). Additionally, YNYX tobacco leaves showed higher δ^{15}N values compared to the other four cities. For different cities, a decreasing δ^{2}H trend of YNQJ (−99.3 ± 1.4‰) > YNWS (−106.6 ± 4.5‰) > YNPE (−118.4 ± 4.8‰) > YNYX (−123.0 ± 3.6‰) > YNCX (−130.2 ± 9.6‰) was observed. This is primarily because the precipitation in YNCX and YNYX during the tobacco leaf growth period (May to July) is relatively low compared to other regions. In addition, major YNQJ samples were collected in high-altitude areas compared to YNWS and YNPE. Overall, differences in stable isotopes are chiefly due to latitudinal effects, precipitation, light irradiance conditions, and photosynthetic reactions (Sanchez-Bragado, Serret, Marimon, Bort, & Araus, 2019).

The full results of the elemental analysis are presented in Table S2. YNQJ samples were richer in Zn, Rb, Cd, $^{207/206}$Pb, and $^{208/206}$Pb compared to other tobacco samples. Na, Mg, Al, K, Ga, Sn, and Pb contents were significantly higher in YNYX tobacco than in tobacco grown in other regions, while the $^{208/206}$Pb ratio was lower. YNWS tobacco leaves were richer in rare earth elements (Y, La, Pr, Nd, Sm, Dy, Er, and Tl) and contained lower levels of $^{207/206}$Pb than tobacco from the other four cities. YNCX tobacco had the highest Sr concentration, with mean values of 896.8 mg/kg.

Minerals are acquired from the soil, fertilizers, and industrial pollution by growing tobacco plants. Lead isotopes ($^{208/206}$Pb) from soil mineralization reflect the soil origin and geological age, increasing geographic resolution as lithology varies by region (Zhang, Sun, Zhou, Fan, Zhai, & Yin, 2002). In our study, $^{87/86}$Sr values of tobacco from YNQJ (1.8 ± 0.1%) and YNWS (1.9 ± 0.7%) were more positive than findings from YNYX (1.0 ± 0.03%) and YNCX (0.9 ± 0.1%). This is likely due to the content of ^{87}Sr produced by the radioactive decay of ^{87}Rb in the bedrock of each area being different, resulting in $^{87/86}$Sr having regional characteristics (Durante et al., 2021).

A heat map generated by hierarchical Pearson's clustering was constructed to study the main sources of variability at the municipal scale based on their stable isotopes and elemental profile (Fig. S3). HCA, an unsupervised pattern recognition technique, can more comprehensively and intuitively reveal correlations between variables (stable isotopes and elements) and geographic origin. Two primary tobacco clusters were evaluated based on the mean of the 35 variables. The first group included YNYX and YNCX tobacco, while the second group included all other regions. YNPE and YNQJ tobacco leaves formed a subgroup in the second group, as well as WS. These clusters are strongly correlated with geographic location (Fig. 1). YNYX and YNCX tobacco aligned with longitude and adjacent, lower Y, La, Pr, Nd, Sm, Dy, Er, Tl contents, and $^{87/86}$Sr ratio, and had higher Na, Mg, K, Cr, and Cu values. However, YNPE, YNQJ, and YNWS tobacco had lower Sn, Pb, δ^{13}C, Al, Se, Sr, Cr, Na, Mg, and K contents and an increased $^{87/86}$Sr ratio. HCA results provide a preliminary indication of tobacco origins, with cluster classification aligned to geographic grouping.

3.3. Geographical origin discrimination at two geographical scales

The biplots (Scores-Loading plot) for OPLS-DA models for distinguishing YN, GZ, and SC tobacco at the provincial scale are displayed in Fig. 3A. The biplots simultaneously visualize scores centroid and loading plots to interpret the observations with respect to individual variables. 176 tobacco samples were divided into three categories using 35 variables. LV1 and LV2 explained 11.6% and 7.4% of the variance of raw data, respectively. The R^2Y (0.827) and Q^2 (0.78) of the model

Table 1
Elements (including five isotope ratios) of flue-cured tobacco at a provincial scale (Yunnan Guizhou and Sichuan provinces). Mg, Al, K, and Mn concentrations are expressed in mg/g; Ga, Sn, Sb, Cs, Pr, Sm, Dy, Er, and Tl concentrations are expressed in μg/kg; all other elements concentrations are expressed in mg/kg.

Provincial scale	YN (n = 84)	GZ (n = 52)	SC (n = 40)	*P*
Li	2.2 ± 2.6b	4.1 ± 5.2a	3 ± 3.4ab	*
Na	37.7 ± 35.1a	22 ± 11.9b	13.7 ± 35.6b	***
Mg	2 ± 2.6b	4.5 ± 2.2a	3.8 ± 1.6a	***
Al	0.3 ± 0.2a	0.1 ± 0.05b	0.2 ± 0.2a	***
K	7.3 ± 10.2b	24.4 ± 5.3a	25.8 ± 8.3a	***
Cr	0.5 ± 0.3a	0.2 ± 0.1b	0.5 ± 0.3a	***
Mn	0.2 ± 0.1a	0.2 ± 0.1a	0.3 ± 0.3a	ns
Co	0.3 ± 0.2a	0.3 ± 0.2a	0.4 ± 0.4a	ns
Ni	0.8 ± 0.4a	0.5 ± 0.3b	0.9 ± 0.8a	***
Cu	12.2 ± 7.1a	6.1 ± 3.7b	6 ± 4.2b	***
Zn	47.4 ± 13.8a	45.4 ± 15.5ab	37.8 ± 20.2b	**
Ga	73.4 ± 47.7a	31.8 ± 17.2b	91.3 ± 105.1a	***
As	0.3 ± 0.2a	0.1 ± 0.04b	0.1 ± 0.01b	***
Se	0.3 ± 0.6a	0 ± 0.1b	0.5 ± 0.3a	***
Rb	7.3 ± 5.0a	3.6 ± 1.8b	3.4 ± 1.9b	***
Y	0.4 ± 0.4a	0.1 ± 0.1b	0.4 ± 0.7a	***
Cd	5.6 ± 4.8a	3.1 ± 2.2b	2 ± 1.2b	***
Sn	61 ± 59.7a	59.3 ± 30.7a	63.2 ± 44.7a	ns
Sb	41.2 ± 28.8a	34.9 ± 16.4ab	24.5 ± 28b	*
Cs	62.3 ± 28.3a	13.6 ± 8.7c	32.5 ± 24.8b	***
Ba	25.7 ± 16b	23 ± 11.3b	35.9 ± 29.3a	**
La	0.9 ± 0.9ab	0.3 ± 0.3b	1.2 ± 2.8a	*
Pr	114.1 ± 90.3a	40.9 ± 27.4b	149.1 ± 275.5a	**
Nd	0.4 ± 0.3a	0.1 ± 0.1b	0.5 ± 0.8a	***
Sm	69.4 ± 47.2a	20.3 ± 11.5b	69.7 ± 88.4a	***
Dy	48.3 ± 35.6a	12.5 ± 6.8b	42.4 ± 67.5a	***
Er	21.3 ± 15.5a	6.4 ± 3.3b	17.8 ± 27.6a	***
Tl	45.0 ± 23.8a	42.1 ± 25.3a	26.0 ± 19.5b	***
Sr	276.5 ± 327.7a	64.6 ± 33.2b	69.6 ± 39.7b	***
Pb	1.5 ± 1.5a	1.2 ± 0.4a	1 ± 0.9a	ns
$^{84/86}$Sr	0.1 ± 0.01a	0.1 ± 0.01a	0.1 ± 0.01a	ns
$^{87/86}$Sr	1.5 ± 0.5a	1.2 ± 0.3b	1.1 ± 0.2b	**
$^{88/86}$Sr	9.2 ± 0.1b	9.2 ± 0.1b	9.4 ± 0.2a	***
$^{207/206}$Pb	0.8 ± 0.01b	0.8 ± 0.01b	0.9 ± 0.01a	***
$^{208/206}$Pb	2.1 ± 0.05a	2.1 ± 0.03b	2.1 ± 0.02a	***

p-value of the ANOVA; ns: *p*-value > 0.05, *:0.05 > *p*-value > 0.01, **: 0.01 > *p*-value > 0.001, ***: *p*-value ≤ 0.001. Different letters on each variable denote differences between different geographical origins.

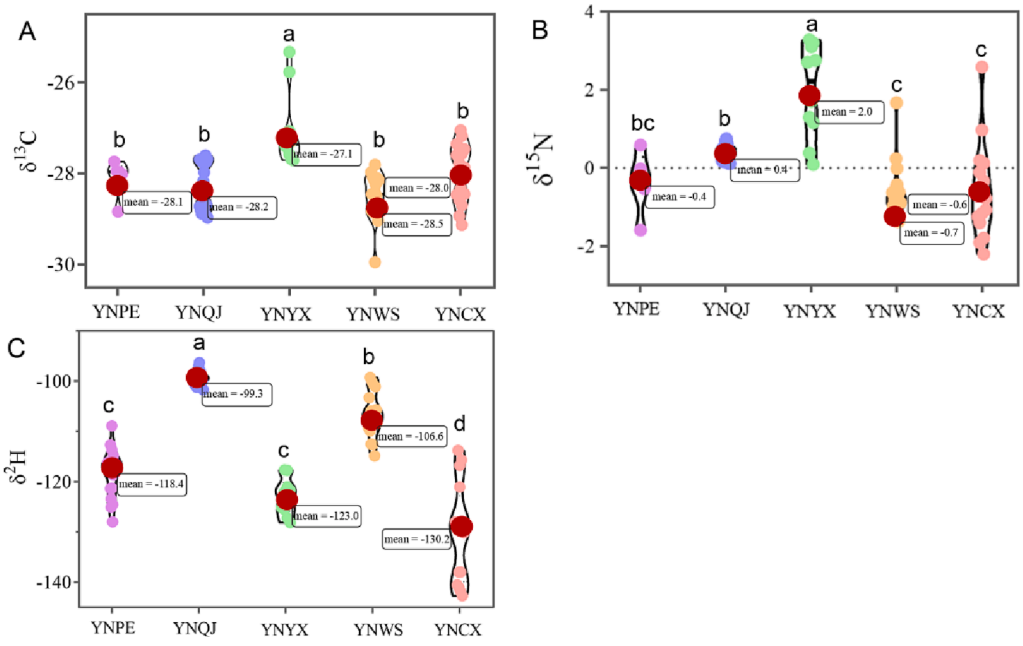


Fig. 2. Stable isotopes of (A) $\delta^{13}C$, (B) $\delta^{15}N$, and (C) $\delta^{2}H$ of different geographical origins at a municipal scale.

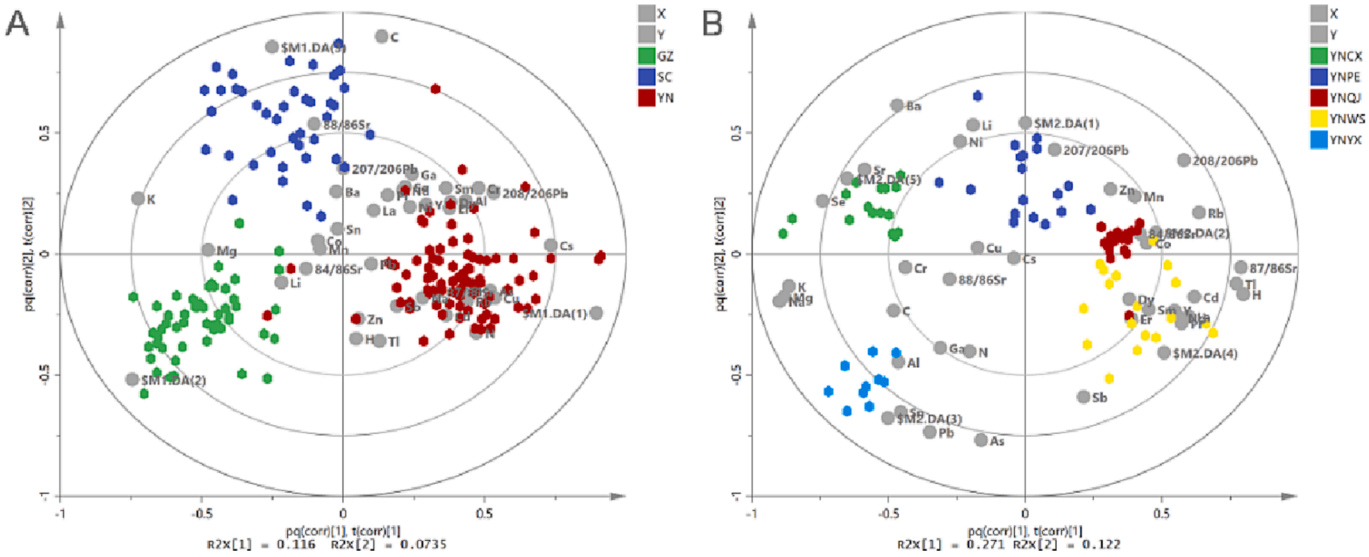


Fig. 3. OPLS-DA modeling of stable isotope and multi-element signatures at (A) provincial scale and (B) municipal scale.

indicated sufficient predictability, where R^2Y explained a proportion of cumulative component variance in the model data, and Q^2 represented the predictability of the model. The loading distribution of variables indicated that Cs, $^{208/206}Pb$, Cu, and Cr were representative variables for YN tobacco, while $\delta^{13}C$, K, and $^{88/86}Sr$ were more representative of SC tobacco. GZ tobacco was correlated with Mg. Cross-validation (Table 2) of this model resulted in a total discriminant accuracy of 98.3%. The cross-validation discriminant accuracies were 96.4% for the YN province and 100% for both the GZ and SC provinces (Table 2). Similar to prior reports on potatoes (Opatić et al., 2018) and olive oil (Benincasa, Lewis, Perri, Sindona, & Tagarelli, 2007), we observed that the "fingerprint" of rare earth elements played a significant role in the geographical authentication of tobacco.

At the municipal scale, stable isotope and multi-element authentication of five sites in the YN province are displayed in Fig. 3B. Tobacco samples were separated into five groups according to their geographical

Table 2
Discriminant accuracies of tobacco leaves at different spatial scales using the OPLS-DA model.

Scales	Date	Discriminant attribution	Discriminant accuracy
Provincial scale	YN	81/84	96.4%
	GZ	52/52	100%
	SC	40/40	100%
	Total	173/176	98.3%
Municipal scale	YNPE	18/18	100%
	YNQJ	21/22	95.5%
	YNYX	10/10	100%
	YNWS	17/18	94.4%
	YNCX	16/16	100%
	Total	82/84	97.6%

origins. Plots of samples and variables projected from the first two latent variables (LV1 and LV2) generally distributed samples into five clusters with a slight overlap between the YNQJ and YNWS regions, where LV1 and LV2 explained 27.1% and 12.2% variance of raw data, respectively. YNCX tobacco was correlated with Sr, Se, and Ba. YNPE tobacco was correlated with the $^{208/206}$Pb and $^{207/206}$Pb ratios. YNQJ tobacco was characterized by Rb and $^{208/206}$Pb. YNWS tobacco was characterized by Cd, Sb Tl, $^{88/86}$Sr, and δ^2H variables. YNYX samples were linked to Pb, As, and Sn. Cross-validation (Table 2) of this model resulted in a total discriminant accuracy of 97.6%. The prediction accuracy was 95.5% for tobacco from YNQJ and 94.4% for YNWS samples, while the remaining three regions could be accurately identified (100%). This can be explained, by YNQJ and YNCX being geographically adjacent and having similar climates in the middle and later stages of flue-cured tobacco growth compared to other regions. Overall, stable isotopes and elemental analysis of tobacco combined with chemometrics provide a promising approach for origin traceability in small growing zones.

VIP plots of stable isotope and multi-element at provincial and municipal scales were obtained from the OPLS-DA model above. The rankings of the importance of given variables (VIP > 1) were different between the two scales (Fig. 4). For the provincial scale, the most important variable was δ^{13}C, followed by K, Cs, $^{208/206}$Pb, $^{88/86}$Sr, Cu, Cr, δ^{15}N, As, Al, Rb, Mg, Sm, Sr, and Cd. For the municipal scale, the most discriminant variable was Sr, closely followed by Se, Pb, Sb, Rb, Tl, Mg, Cu, As, Cd, $^{208/206}$Pb, Sn, Ba, Na, K, δ^2H, δ^{15}N, and $^{84/86}$Sr. While K, Sr, $^{208/206}$Pb, Cu, As, Rb, Mg, Cd, and δ^{15}N were key variables across both provincial and municipal scales, their contributions were different. Moreover, the number of discriminant variables for municipal-scale authentication was more pronounced than at the provincial scale. The elements Se, Pb, Sb, Sn, Ba, Na, δ^2H, and $^{88/86}$Sr were important for municipal scale authentication, but not necessarily for the provincial scale, while δ^{13}C was relevant for provincial scale authentication, which is similar to observations in *D. officinale* (Xiong et al., 2022). Additionally, the contour lines of δ^{13}C and K calculated using spatial interpolation also indicate significant differences across the three provinces. The distribution of Sr and Se is significantly different for the municipal scale (Fig. S4). This result indicates that stable isotopes and elemental indicators have different distribution patterns in flue-cured tobacco at different spatial scales.

4. Conclusion

Our study indicates that stable isotope and element fingerprinting in combination with statistical analysis (e.g., OPLS-DA) can effectively differentiate tobacco obtained from different regions. Isotope and elemental indicators can provide information regarding the geographical origin of flue-cured tobacco at two spatial scales. At the provincial scale, YN had higher values of Na, Cs, and Sr and lower values of K concentration than GZ and SC. At the municipal scale (considering YNPE, YNQJ, YNYX, YNWS, and YNCX) the key variables for tobacco authentication were Sr, Se, and Pb. Furthermore, the discriminant accuracy at the provincial scale was higher (max 98.3%) than at the municipal scale (max 97.6%). The importance of ranking stable isotopes and elements differs with the spatial scale being evaluated.

Stable isotopes and elemental concentrations may be impacted by the varietal and inter-annual differences, while our research is limited to data from one year. This work established a foundation for small region traceability of tobacco, but discrimination accuracy based on samples from a single year may not be comprehensive. More sampling sites from across different provinces would also make the conclusion more dependable at the provincial scale. Further investigation is therefore needed to assess the influence of regions, years, seasons, and climate change on stable isotopes and elemental contents.

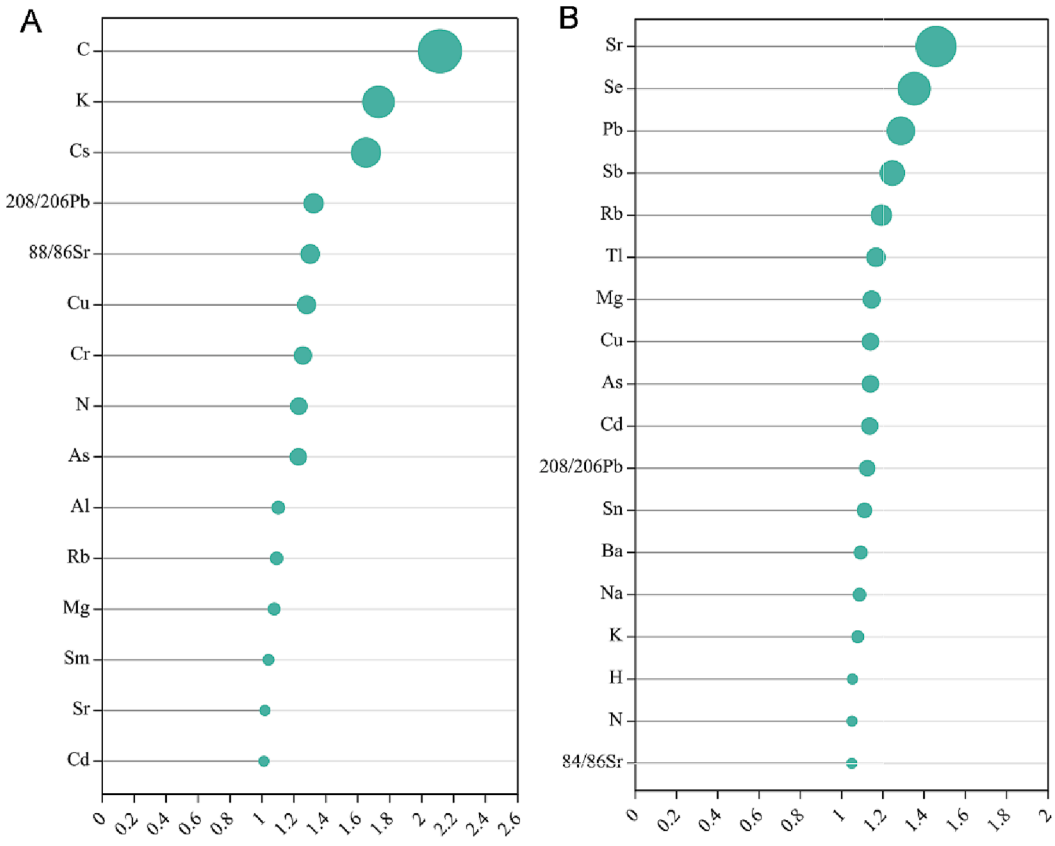


Fig. 4. VIP plots based on stable isotopes and elemental contents at (A) provincial scale and (B) municipal scale.

Author contributions

Lili Cui: conceptualization, data curation and writing-original draft; Huan Chen: data curation; Yuwei Yuan and Jing Nie: formal analysis and writing-review; Fengpeng Zhu, Shulei Han, and Ya'ning Fu: methodology; Hongwei Hou: investigation and supervision; Qingyuan Hu: funding acquisition and project administration; Zengping Chen: software and validation. The final version was approved by all authors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

This work was supported by the Provincial and Ministerial Major Project of China under grants 110202001006 (XX-02).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fochx.2023.100716.

References

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics, 2*(4), 433–459. https://doi.org/10.1002/wics.101

Benincasa, C., Lewis, J., Perri, E., Sindona, G., & Tagarelli, A. (2007). Determination of trace element in Italian virgin olive oils and their characterization according to geographical origin by statistical analysis. *Analytica Chimica Acta, 585*(2), 366–370. https://doi.org/10.1016/j.aca.2006.12.040

Binette, M.-J., Lafontaine, P., Vanier, M., & Ng, L.-K. (2009). Characterization of Canadian cigarettes using multi-stable isotope analysis by gas chromatography–Isotope ratio mass spectrometry. *Journal of Agricultural and Food Chemistry, 57*(4), 1151–1155. https://doi.org/10.1021/jf802642d

Boccard, J., & Rutledge, D. N. (2013). A consensus orthogonal partial least squares discriminant analysis (OPLS-DA) strategy for multiblock Omics data fusion. *Analytica Chimica Acta, 769*, 30–39. https://doi.org/10.1016/j.aca.2013.01.022Get

Bontempo, L., Bertoldi, D., Franceschi, P., Rossi, F., & Larcher, R. (2021). Elemental and isotopic characterization of tobacco from Umbria. *Metabolites, 11*(186), 1–12. https://doi.org/10.3390/metabo11030186

Bontempo, L., Camin, F., Manzocco, L., Nicolini, G., Wehrens, R., Ziller, L., & Larcher, R. (2011). Traceability along the production chain of Italian tomato products on the basis of stable isotopes and mineral composition. *Rapid Communications in Mass Spectrometry, 25*(7), 899–909. https://doi.org/10.1002/rcm.4935

Buchmann, N., Brooks, J. R., Rapp, K. D., & Ehleringer, J. R. (1996). Carbon isotope composition of C4 grasses is influenced by light and water supply. *Plant, Cell & Environment, 19*(4), 392–402. https://doi.org/10.1111/j.1365-3040.1996.tb00331.x

Dansgaard, W. (1964). Stable isotopes in precipitation. *Tellus, 16*(4), 436–468. https://doi.org/10.1111/j.2153-3490.1964.tb00181.x

Durante, C., Lancellotti, L., Manzini, D., Rossi, M. C., Sighinolfi, S., Marchetti, A., & Tassi, L. (2021). 87Sr/86Sr ratio as traceability marker for Modena's balsamic vinegars. *LWT, 147*, 111571. https://doi.org/10.1016/j.lwt.2021.111571

Evans, J. R., & Von Caemmerer, S. (2013). Temperature response of carbon isotope discrimination and mesophyll conductance in tobacco. *Plant, Cell & Environment, 36*(4), 745–756. https://doi.org/10.1111/j.1365-3040.2012.02591.x

Hu, Y., Cheng, H., & Tao, S. (2016). The challenges and solutions for cadmium-contaminated rice in China: A critical review. *Environment International, 92–93*, 515–532. https://doi.org/10.1016/j.envint.2016.04.042

Hua, D. A., Bt, B., Sc, B., Cx, B., Xl, B., Lei, Z. B., … Zc, A. (2020). Combination of stable isotopes and multi-elements analysis with chemometric for determining the geographical origins of Rhizoma Coptidis - ScienceDirect. *Microchemical Journal, 152*(104427). https://doi.org/10.1016/j.microc.2019.104427

Jamin, E., Naulet, N., & Martin, G. (1997). Multi-element and multi-site isotopic analysis of nicotine from tobacco leaves. *Plant, Cell & Environment, 20*(5), 589–599. https://doi.org/10.1111/j.1365-3040.1997.00099.x

Krueger, H. W., & Reesman, R. H. (1982). Carbon isotope analyses in food technology. *Mass Spectrometry Reviews, 1*(3), 205–236. https://doi.org/10.1002/mas.1280010302

Li, Y., Ren, K., Zou, C., Xie, J., He, X., Chen, Y., … Chen, J. (2019). Effects of ferrous iron toxicity on agronomic, physiological, and quality indices of flue-cured tobacco. *Agronomy Journal, 111*(5), 2193–2206. https://doi.org/10.2134/agronj2018.12.0786

Liu, G., & Liu, J. (2010). *Practical guide to leaf tobacco production techniques in China* (pp. 194–200). Beijing: China National Leaf Tobacco Corporation.

Liu, X., Liu, Z., Qian, Q., Song, W., & Yuan, Y. (2020). Isotope chemometrics determines farming methods and geographical origin of vegetables from Yangtze River Delta Region, China. *Food Chemistry, 342*(128379), 1–9. https://doi.org/10.1016/j.foodchem.2020.128379

Liu, Z., Zhang, W., Zhang, Y., Chen, T., Shao, S., Zhou, L., … Rogers, K. M. (2019). Assuring food safety and traceability of polished rice from different production regions in China and Southeast Asia using chemometric models. *Food Control, 99*, 1–10. https://doi.org/10.1016/j.foodcont.2018.12.011

Luo, R., Jiang, T., Chen, X., Zheng, C., Liu, H., & Yang, J. (2019). Determination of geographic origin of Chinese mitten crab (Eriocheir sinensis) using integrated stable isotope and multi-element analyses. *Food Chemistry, 274*, 1–7. https://doi.org/10.1016/j.foodchem.2018.08.104

Matthews, A. M., Fu, R., Dana, T., & Chou, R. (2016). Intranasal or transdermal nicotine for the treatment of postoperative pain. *Cochrane Database of Systematic Reviews, 1*. https://doi.org/10.1002/14651858.CD009634.pub2

Muoz-Redondo, J. M., Bertoldi, D., Tonon, A., Ziller, L., & Moreno-Rojas, J. M. (2021). Tracing the geographical origin of Spanish mango (Mangifera indica L.) using stable isotopes ratios and multi-element profiles. *Food Control, 125*(2–3), 107961. doi: j.foodcont.2021.107961.

Peng, M. X., Wang, C. H., Yang, Y. Z., & Zhu, Q. A. (2017). Spatial patterns of leaf C-13 and its relationship with plant functional groups and environmental factors in China. *J. Geophys. Res.-Biogeo., 122*(7), 1564–1575. https://doi.org/10.1002/2016JG003529

Nguyen, H. T. T., Giang, L. T., & Pham, T. N. (2020). Empirical analysis on the illicit trade of cigarettes in Vietnam. *Tobacco Control, 29*(4), 281–286. https://doi.org/10.1136/tobaccocontrol-2019-055598

Nie, J., Shao, S., Zhang, Y., Li, C., & Yuan, Y. (2021). Discriminating protected geographical indication Chinese Jinxiang garlic from other origins using stable isotopes and chemometrics. *Journal of Food Composition and Analysis, 99*(2), 103856. https://doi.org/10.1016/j.jfca.2021.103856

Nie, J., Weng, R., Li, C., Liu, X., Wang, F., Rogers, K. M., … Yuan, Y. (2022). Chemometric origin classification of Chinese garlic using sulfur-containing compounds, assisted by stable isotopes and bioelements. *Food Chemistry, 394*(133557), 1–11. https://doi.org/10.1016/j.foodchem.2022.133557

Opatić, A. M., Nečemer, M., Budič, B., & Lojen, S. (2018). Stable isotope analysis of major bioelements, multi-element profiling, and discriminant analysis for geographical origins of organically grown potato. *Journal of Food Composition and Analysis, 71*, 17–24. https://doi.org/10.1016/j.jfca.2018.04.005

Petrini, R., Sansone, L., Slejko, F., Buccianti, A., Marcuzzo, P., & Tomasi, D. (2015). The 87Sr/86Sr strontium isotopic systematics applied to Glera vineyards: A tracer for the geographical origin of the Prosecco. *Food Chemistry, 170*, 138–144. https://doi.org/10.1016/j.foodchem.2014.08.051

Sanchez-Bragado, R., Serret, M. D., Marimon, R. M., Bort, J., & Araus, J. L. (2019). The hydrogen isotope composition δ2H reflects plant performance. *Plant Physiology, 180*(2), 793–812. https://doi.org/10.1104/pp.19.00238

Simon-Delso, N., Amaral-Rogers, V., Belzunces, L. P., Bonmatin, J.-M., Chagnon, M., Downs, C., … Girolami, V. (2015). Systemic insecticides (neonicotinoids and fipronil): Trends, uses, mode of action and metabolites. *Environmental Science and Pollution Research, 22*(1), 5–34. https://doi.org/10.1007/s11356-014-3470-y

Tang, Z., Chen, L., Chen, Z., Fu, Y., Sun, X., Wang, B., & Xia, T. (2020). Climatic factors determine the yield and quality of Honghe flue-cured tobacco. *Scientific Reports, 10*(1), 1–12. https://doi.org/10.1038/s41598-020-76919-0

Wadood, S. A., Nie, J., Li, C., Rogers, K. M., Zhang, Y., & Yuan, Y. (2022). Geographical origin classification of peanuts and processed fractions using stable isotopes. *Food Chemistry: X, 16*, 100456. https://doi.org/10.1016/j.fochx.2022.100456

Wang, C., He, W. Z., Zhao, D. Y., Liu, Z., Fan, Y. Y., Tian, W. N., … Rogers, K. M. (2020). Modeling of stable isotope and multi-element compositions of jujube (Ziziphus jujuba Mill.) for origin traceability of protected geographical indication (PGI) products in Xinjiang, China. *Journal of Food Composition and Analysis, 92*(103577), 1–12. https://doi.org/10.1016/j.jfca.2020.103577

Wang, J., Chen, T., Zhang, W., Zhao, Y., Yang, S., & Chen, A. (2020). Tracing the geographical origin of rice by stable isotopic analyses combined with chemometrics. *Food Chemistry, 313*, 126093. https://doi.org/10.1016/j.foodchem.2019.126093

Wang, Y., Kang, L., Zhao, Y., Xiong, F., Yuan, Y., Nie, J., … Yang, J. (2022). Stable isotope and multi-element profiling of Cassiae Semen tea combined with chemometrics for geographical discrimination. *Journal of Food Composition and Analysis, 107*, 104359. https://doi.org/10.1016/j.jfca.2021.104359

Wu, W., Tang, X.-P., Yang, C., Liu, H.-B., & Guo, N.-J. (2013). Investigation of ecological factors controlling quality of flue-cured tobacco (Nicotiana tabacum L.) using classification methods. *Ecological Informatics, 16*, 53–61. https://doi.org/10.1016/j.ecoinf.2013.04.008

Xia, W., Li, C., Nie, J., Shao, S., Rogers, K. M., Zhang, Y., … Yuan, Y. (2022). Stable isotope and photosynthetic response of tea grown under different temperature and light conditions. *Food Chemistry, 368*, 130771. https://doi.org/10.1016/j.foodchem.2021.130771

Xia, W., Li, Z., Yu, C., Liu, Z., Nie, J., Li, C., … Yuan, Y. (2021). Understanding processing, maturity and harvest period effects to authenticate early-spring Longjing

tea using stable isotopes and chemometric analyses. *Food Control, 124*, 107907. https://doi.org/10.1016/j.foodcont.2021.107907

Xiong, F., Yuan, Y., Li, C., Lyu, C., Wan, X., Nie, J., … Guo, L. (2022). Stable isotopic and elemental characteristics with chemometrics for the geographical origin authentication of Dendrobium officinale at two spatial scales. *LWT, 167*, 113871. https://doi.org/10.1016/j.lwt.2022.113871

Zhang, H.-F., Sun, M., Zhou, X.-H., Fan, W.-M., Zhai, M.-G., & Yin, J.-F. (2002). Mesozoic lithosphere destruction beneath the North China Craton: Evidence from major-, trace-element and Sr–Nd–Pb isotope studies of Fangcheng basalts. *Contributions to Mineralogy and Petrology, 144*(2), 241–254. https://doi.org/10.1007/s00410-002-0395-0