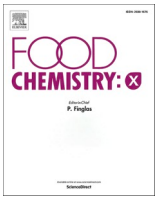
Food Chemistry: X 18 (2023) 100678

Contents lists available at ScienceDirect

Food Chemistry: X

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



Authenticating the geographical origin of the Chinese yam (Tiegun) with stable isotopes and multiple elements

Feng Xiong [a,b,1], Chaogeng Lyu [a,b,1], Chuanzhi Kang [a,b], Xiufu Wan [a,b], Jiahui Sun [a,b], Tielin Wang [a,b], Sheng Wang [a,b], Haiyan Li [c], Jian Yang [a,b,*], Lanping Guo [a,b,*]

[a] State Key Laboratory Breeding Base of Dao-di Herbs, National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, PR China
[b] Key Laboratory of Biology and Cultivation of Herb Medicine, Ministry of Agriculture and Rural Affairs, Beijing 100700, PR China
[c] Institute of Information on Traditional Chinese Medicine, China Academy of Chinese Medical Sciences, Beijing 100700, PR China

ARTICLE INFO

Keywords:
Geographical authentication
Chinese yam (Tiegun)
Stable isotopes
Multielements
Environmental factors

ABSTRACT

The Chinese yam, an important orphan crop with both high nutrient and health promoting value, is mainly produced in the Yellow-Huai-Hai plain near the river basins in China. The protected designation of origin (PDO)-labeled Chinese yam differs greatly from others in market acceptance and price, which has led to fakes and the need for reliable authentication methods. Hence, stable isotope ratios of $\delta^{13}C$, $\delta^{15}N$, $\delta^{2}D$, and $\delta^{18}O$ and 44 multielemental contents were used to explore the authentication of geographical origins and the effect of environmental factors. Twenty-two elements and $\delta^{15}N$ were selected as the key variables to authenticate Chinese yams from three river basins as well as to authenticate them among traditional PDOs and others in the Yellow River basin. Moreover, six environmental factors, including the moisture index, maximum temperature, photosynthetically active radiation, soil organic carbon, total nitrogen and pH, were found to be highly related to these variances.

Introduction

Authenticating the geographical origin of foods is of great importance for guaranteeing food security and quality, especially for high economic value agro-foods with trade markers such as the protected designation of origin (PDO) label (Dias & Mendes, 2018). Many yams of *Dioscorea* spp. are treated as orphan crops (Epping & Laibach, 2020) and contain various nutrients, especially the cultivational variant *Dioscorea opposite* Thunb. cv. Tiegun (*D. opposite* Thunb. cv. Tiegun), usually called Chinese yam (Tiegun). In addition to being a stable food, it is also a functional food and commonly used traditional medicine material in China with many health-promoting properties, such as stimulating the spleen and stomach, promoting body fluid production, and benefiting the lungs (Lyu et al., 2021). In the long history of cultivating the Chinese yam in China, the Chinese yam (Tiegun) was mainly produced in central to eastern China, including Henan, Hebei, Shandong and southern Shanxi Provinces. This was because the widespread river basins provided sandy soil near the rivers, which was beneficial for Chinese yams to grow deep and strait roots (Zhang et al., 2011). Thus, the Yellow-Huai-Hai Plain provides a suitable soil environment as well as climate conditions for the cultivation of the Chinese yam because there are three main rivers, the Yellow River (YeR), the Huai River (HuR), and the Hai River (HaR).

Moreover, the ancient Huaiqingfu region located along the YeR basin has long been treated as a traditional production area of high-quality Chinese yams (Tiegun) and is also labeled as the PDO, which mainly includes two counties, Wuzhi and Wenxian. In the Yellow River basin, the prices of Tiegun yams from traditional areas were much higher than others, and some Tiegun yams with no PDO labels from nontraditional areas were considered fake PDO Chinese yams (Tiegun) through transport to traditional areas for packaging, which led to the demand for a reliable authentication method to recognize the fake yams. However, how to authenticate them from different river basins remains unclear and poorly understood.

Stable isotope ratios (SIRs) and multielement fingerprints are the recently commonly used tracing indices for the geographical origins of

* Corresponding authors at: State Key Laboratory Breeding Base of Dao-di Herbs, National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, PR China.
E-mail addresses: yangchem2012@163.com (J. Yang), glp01@126.com (L. Guo).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.fochx.2023.100678
Received 31 October 2022; Received in revised form 19 March 2023; Accepted 9 April 2023
Available online 17 April 2023

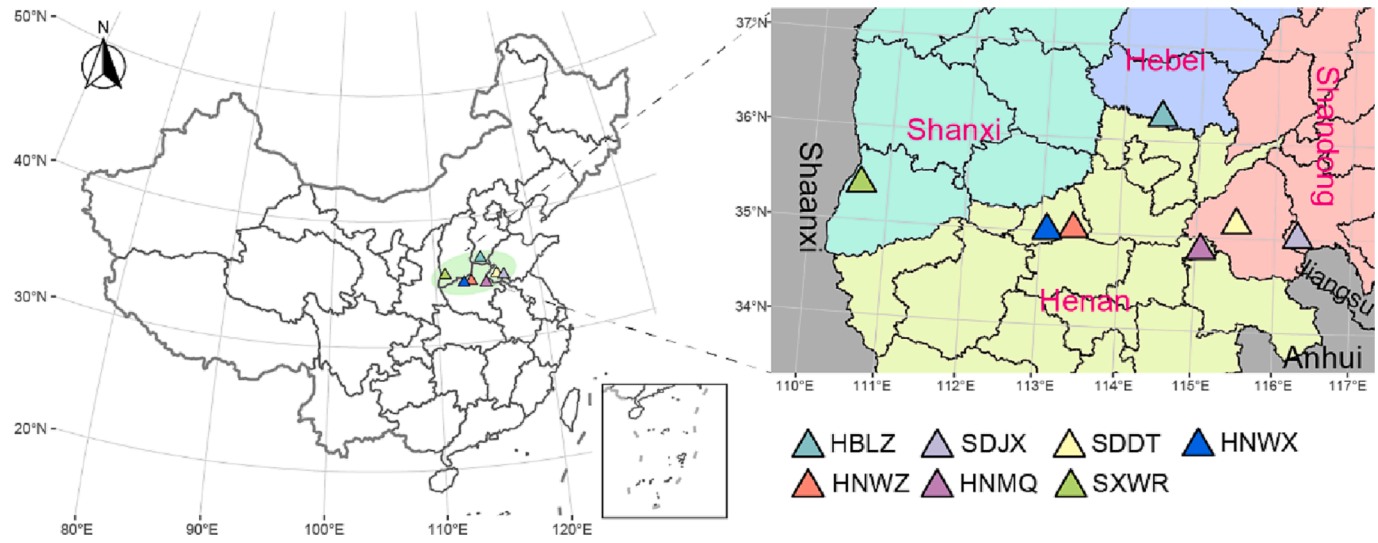


Fig. 1. Geographical locations of the sampling sites for the Chinese yam.

agroproducts due to their high accuracy (Dias & Mendes, 2018; Liu et al., 2022). A previous study also proved that the SIRs and multielement variables could discriminate between organically and conventionally cultivated Chinese yams (Lyu et al., 2021). However, the characteristics of SIRs and multielement fingerprints in Chinese yams (Tiegun) from different production areas remain to be discovered, and thus the important variables to separate Chinese yams (Tiegun) from different origins need to be determined. It is unclear whether SIRs and multielement fingerprints provide powerful effects in geographical origin traceability for Chinese yams (Tiegun) and what the most important indices are.

In addition, how the SIRs and elements in Chinese yams (Tiegun) react to different environmental factors has also been investigated. SIRs such as $\delta^{13}C$, $\delta^{15}N$, $\delta^{2}H$, and $\delta^{18}O$ in plants are largely correlated with climate and soil factors at large geographical scales (Xiong et al., 2022; Yoneyama, Fujiwara, & Wilson, 1998), while multielement fingerprints are more correlated with the soil background, which is thought to be heterogeneous (Hao et al., 2021; Zhao & Yang, 2022). Consequently, the cause of the variation in SIRs and elemental characteristics in Chinese yams (Tiegun) from different regions was investigated. It is unclear whether the important variables in the above context varied according to environmental factors or geographical origins and what the key environmental factors that affect those variables are.

Consequently, the purposes of this study were as follows: (1) to draw the features of multielement fingerprints and stable isotope ratios of Chinese yams (Tiegun) from different production areas and reveal the differences between the basins of the YeR, HuR, and HaR systems and between the PDO Tiegun yams (produced in the ancient Huaiqingfu area) and others; (2) to explore environmental factors that affect the variance of multielements and stable isotopes of Chinese yams (Tiegun) from different origins; and (3) to develop a geographical origin authentication method based on a machine learning algorithm to authenticate geo-authentic Chinese yams by multielement analysis and stable isotope ratios.

Materials and methods

Sample collection and data origin

A total of 239 batches of samples were collected in the seven main Chinese yam-producing areas (Fig. 1 and Table S1), including Linzhang County (HBLZ) in Hebei Province, Jinxiang County (SDJX) and Dingtao County (SDDT) in Shandong Province, Wenxian County (HNWX), Wuzhi County (HNWZ), and Mingquan County (HNMQ) in Henan Province, and Wanrong County (SXWQ) in Shanxi Province (Table S1). Samples from the seven counties cover all three main production areas of Chinese yams (Tiegun) and include the two PDOs (the ancient Huaiqingfu region including HNWZ and HNWX). In each production area, fresh yam samples were collected from both the different merchants in local markets and directly bought from farmers in the field. Samples were identified as *D. opposite* Thunb. cv. Tiegun by Prof. Yan Jin (National Resource Center for Chinese Materia Medica), and voucher specimens were deposited in the National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences (Beijing). Finally, the fresh samples were washed, peeled, and dried in a drying oven at 50 °C for 36 h, and the dried samples were ground into fine powder and stored in a desiccator before further analyses.

The GPS information and the 18 collected environmental variables of the 7 production areas are listed in Table S1. The bioclimate factors were obtained from the Science Data Bank and extracted by the GPS location (Wei Lin-Feng, Hu Xiao-Fei, Cheng Qi, Wu Xing-Qi, & Ni Jian, 2022), and the soil backgrounds were extracted from the Soil Grids (https://soilgrids.org). In total, the 18 environmental factors included altitude (m), MAP (mean annual precipitation, mm), MAT (mean annual temperature, °C), PAR (photosynthetically active radiation, Wm-2), SH (sunlight hours, h), DI (annual drought index), GDD0 (accumulated temperature of growing degree days above 0 °C, °C), GDD5 (accumulated temperature of growing degree days above 5 °C, °C), GP (growing season precipitation, mm), MI (annual moisture index), MTCO (mean temperature of the coldest month, °C), MTWA (mean temperature of the warmest month, °C), Tmax (absolute maximum temperature, °C), Tmin (absolute minimum temperature, °C), soil pH, SOC (soil organic carbon content, dg kg^{-1}), TN (soil total nitrogen content, cg kg^{-1}), and CEC (cation exchange capacity, mmol(c) kg^{-1}).

Chemicals and reagents

To obtain purified water, a Mill-Q system (Millipore, USA) was used. Multielement standard solutions were purchased from Spex (Metuchen, NJ, USA). Analytical grade H_2O_2 was obtained from the Damao Chemical Reagent Factory (Tianjin, China), and ultra-pure HNO_3 was obtained from Merck (Darmstadt, Germany). HCl was purchased from CNW (Shanghai, China).

Stable isotope analysis

Methods for analyzing $\delta^{13}C$, $\delta^{15}N$, $\delta^{2}H$, and $\delta^{18}O$ followed our previous work (Lyu et al., 2021). Approximately 1 mg of powder sample

were weighed in tin capsules in duplicate and then packed into the sample tray of an isotope cube elemental analyzer (Elementar, Germany) for $\delta^{13}C$ and $\delta^{15}N$ analysis. GC columns were used to separate CO_2 and N_2 generated by combusted and reduced samples (99.999% helium).

Samples were combusted, carried by helium (99.999%), and passed through a GC column to separate CO2 and N2. Then, an Isoprime 100 mass spectrometer (Isoprime, Cheadle, UK) was used to detect the gases. The reference materials for δ13C and δ15N were IAEA-600 (−27.77 ‰ ± 0.04 ‰ vs. V-PDB) and IAEA-CH-6 (−10.45 ‰ ± 0.03 ‰ vs. V-PDB) and USGS40 (−4.5 ‰ ± 0.1 ‰ vs. air N_2) and IAEA-N-2 (+20.41 ‰ ± 0.12 ‰ vs. air N_2), respectively. The long-term standard deviations for $\delta^{13}C$ and $\delta^{15}N$ were 0.05‰ and 0.1‰, respectively.

To avoid extra disturbance from environmental water, samples were stored in a 2 mL tube within a desiccator until analysis. To analyze δ^2H and $\delta^{18}O$, 0.5 mg of sample powder was weighed, packed in silver capsules and analyzed in a PYRO cube autosampler (Elementar, Germany). After balancing 72 h in the laboratory to fully expose to the local atmospheric conditions as a equilibration procedure (Wassenaar & Hobson, 2003), H and O in the samples were converted into H_2 and CO at a temperature of 1450 °C with carbon. Then, a GC column was used to separate the gases, and an Isoprime 100 mass spectrometer (Isoprime, UK) was used to detect the gases. The data was corrected by a three-point normalization line. To calibrate δ^2H and $\delta^{18}O$, the reference materials USGS54 ($\delta^2H = -150.4$ ‰ ± 1.1 ‰ and $\delta^{18}O = +17.79$ ‰ ± 0.15 ‰ vs. V-SMOW), USGS55 ($\delta^2H = -28.2$ ‰ ± 1.7 ‰ and $\delta^{18}O = +19.12$ ‰ ± 0.07 ‰ vs. V-SMOW) and USGS56 ($\delta^2H = -44.0$ ‰ ± 1.8 ‰ and $\delta^{18}O = +27.23$ ‰ ± 0.03 ‰ vs. V-SMOW) were used, and the long-term standard deviations were 5‰ and 0.4‰ for δ^2H and $\delta^{18}O$, respectively.

The following equation was used to calculate the SIR (Coplen, 2011; Skrzypek et al., 2022):

$$\delta = R_{sample}/R_{standard} - 1$$

where R_{sample} is the abundance ratio of the heavier isotope versus the lighter isotope in the sample, namely, $^{13}C/^{12}C$, $^{15}N/^{14}N$, $^{18}O/^{16}O$, and $^{2}H/^{1}H$, and $R_{standard}$ is the ratio of the heavier isotope versus the lighter isotope in the reference material. δ values are presented as δ‰. The $\delta^{15}N$ and $\delta^{13}C$ values are relative to atmospheric nitrogen gas (AIR) and Vienna Pee Dee Belemnite (VPDB), respectively. The δ^2H and $\delta^{18}O$ values are relative to the Vienna Standard Mean Ocean Water (VSMOW).

Elemental analysis

First, 100 mg of sample powder was weighed into a Teflon vessel, and the mixed acid (5 mL of 65% HNO_3 and 1 mL of 30% H_2O_2) was added. Then, a Mars 6 microwave digestion system (CEM Corp, USA) was used to digest samples in the vessel. The digestion methods followed our previous study (Lyu et al., 2021). With a 1500 W microwave power, the microwave temperature was raised to 100 °C in 6 min and held for 3 min, then further increased to 160 °C in 7 min and held for 4 min, and again increased to 190 °C in 6 min and held at 190 °C for 30 min. After cooling to room temperature, an electronically heated plate at 180 °C was used to evaporate the solvent to nearly 1 mL, and the remaining solution was diluted to 10 mL with dd H_2O and filtered through a 0.45 μm syringe filter. The filtered solutions were sample solutions prepared for the normal element content tests. For the testing of Na, Mg, K, and Ca, the sample solutions were further diluted to 50 mL with dd H_2O. A standard reference solution was prepared by accurately measuring Multielement Solution I and Multielement Solution II and diluted with 2% HNO_3 until the concentrations of all elements in the reference solution were 10 μg/mL. A Rh solution of 10 μg/mL was used as an internal standard (Varrà, Husáková, Patočka, Ghidini, & Zanardi, 2021).

An ICAP-Q combined with a plasma mass spectrometer (ICP–MS) was used in elemental content testing. Argon gas with flow rates of 0.8 mL/min, 1.05 mL/min, and 14.0 mL/min was used as the auxiliary, carrier, and cold gas, respectively. The atomization chamber was set at 2 °C, and the scan mode was the STD mode with three sampling points and three repeats. The detection limits and recoveries were 0.003–4.826 ng/mL and 85.5%–122.3%, respectively. The contents of elements were expressed as a dry weight basis (mg/kg dw).

Data analysis

R software (version 4.1.0) was used for the data analysis and the production of figures.

First, the unsupervised discriminate method of principal component analysis (PCA) was applied to explore data without supervised separations using the package 'vegan'. PLS-DA was then employed to explore the important variables, and the variables with VIP values larger than 1 were selected for the following study. They were explored by the Welch-test ('onewaytests' package) to evaluate differences between groups. A 'Games-Howell' post hoc test was used to test multiple comparisons among groups, which provides a powerful test when sample size and variance are unequal (Games & John, 1976). Furthermore, Pearson correlation relationships were used to explore correlations between elements.

Then, the environmental factors were collected to explore the relationships between important variables and environmental factors. The mantel test and partial mantel test by the package 'vegan' were used to test whether geographical distance or environmental factors were relevant to the SIRs and elemental characteristics. Based on the GPS information, the geographical distance between production areas was calculated by the package 'geosphere', and the environmental Euclidean distance and the Euclidean distance between samples were also calculated based on their environmental variables and elemental contents by the package 'vegan' (Oksanen et al., 2013). The specific Spearman relationships between Chinese yam SIRs and elemental features with exact environmental factors were calculated and plotted by the package 'LinkET' (Huang, 2021). As a result, the selected environmental factors with high correlation coefficients were used to perform a redundancy analysis (RDA) with the package 'vegan'.

Finally, linear discriminant analysis (LDA), partial least-squares discriminant analysis (PLS-DA), and machine learning methods such as random forest (RF), support vector machine (SVM) and K-nearest neighbors (KNN) were used as discriminant classifiers (Deng et al., 2020; Liu et al., 2022). Furthermore, 20% and 80% of the data were set as testing and training data, respectively. In total, ten calculations were applied, and the mean value was considered the prediction accuracy of the model. A process of hyperparameter tuning (the 'tuneParams' function in the 'mlr' package) was used to find proper hyperparameters of RF, KNN, and SVM, as displayed in Table S2. The initial prior ranges for hyperparameters ntree, mtry, nodesize, and maxnodes for the RF model were set as 300–800, 2–29, 1–10, and 10–35, respectively. For the SVM, the hyperparameters kernels include 'linear,' 'polynomial,' 'radial,' and 'sigmoid.' The degree, cost, and gamma ranges were set as 1–3, 0.1–10, and 0.1–10, respectively. The k value of KNN was set as 1–10. Packages 'mlr' combined with 'randomForest,' 'tidyverse,' and 'SwarmSVM' were used to analyze the LDA, RF, KNN and SVM, respectively.

In addition, the packages 'maptools', 'sf', 'ggspatial', 'tidyverse', 'ggplot2', and 'cowplot' were used to produce the figure of the sampling site location, while the packages 'raster' and 'rgdal' were used to extract the environmental factors of the sites from raster data.

Results and discussion

Overall characteristics of SIRs and multiple elements in Chinese yams

The elemental contents of Chinese yams from different geographical origins are shown in Table S3, and most of the elements were

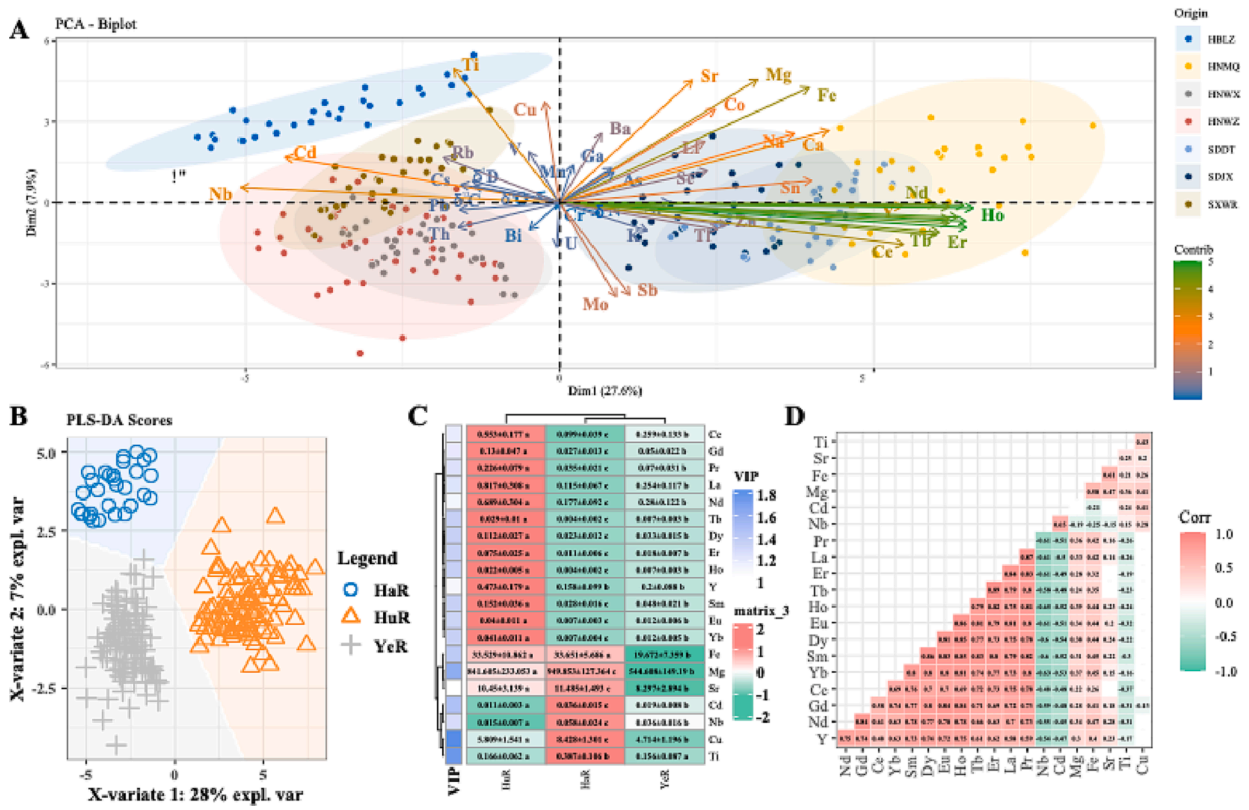


Fig. 2. (A) PCA results for the SIRs and elements in Chinese yams (Tiegun); (B) PLS-DA score plot to separate Chinese yams (Tiegun) from different river basins; (C) Importance and heatmap plot of the variables in the PLS-DA result with a VIP value > 1; and (D) Pearson correlations among these importance variables.

significantly different among different origins, except for V and Ga. Multivariance analysis (Fig. 2A) revealed that Chinese yams from 7 production areas could be clustered into three classes. These class characteristics fit the geographical locations of the 7 counties. Specifically, site HBLZ belongs to the Hai River basin (HaR), sites HNMQ, SDDT, and SDJX belong to the Huai River basin (HuR), and sites HNWX, HNWZ, and SXWR belong to the Yellow River basin (YeR). This result revealed that the SIRs and multielement characteristics differed among Chinese yams from three river basins, even though they were close in geological location. In addition, in the YeR region, it was obvious that HNWZ and HNWX were much closer to each other than to SXWR (Fig. 2A and Fig. S1A), and the former two counties belong to the ancient Huaiqingfu area, where the Chinese yam has been known as a famous geo-authentic production (named Chinese Huai yam) since the Ming dynasty (Wang, Dai, Yuan, & Guo, 2018).

A further supervised PLS-DA was thus applied according to the river systems, including the YeR, the HaR, and the HuR. As a result, the predictive ability (Q2Y(cum)) and cumulative variation (R2Y(cum)) were 0.863 and 0.886, respectively, which revealed that the model was good in both predictive and explanatory ability. As shown in Fig. 2B, these three classes were totally separated according to PLS-DA, and the first two axes explained 28% and 7% of the variance, respectively. The elements selected by VIP > 1 are plotted in Fig. 2C, which included 20 elements ranked as Cu, Ti, Mg, Fe, Dy, Tb, Ho, Sm, Er, Yb, Eu, Nb, Pr, La, Gd, Y and Sr. The important variables did not include the SIRs, this was because SIRs among different river basins demonstrated less variability. The $\delta^{13}C$ and $\delta^{18}O$ was only higher in the YeR than the HuR ($p = 0.006$ and $p = 0.003$, respectively), and the $\delta^{15}N$ was only higher in the YeR than the HaR ($p = 0.03$) (Fig. S2). This may be because the spatial scale in this study was small, while the SIRs usually react to environmental differences on a larger spatial scale. For example, $\delta^{13}C$ in plants are regulated by the process of photosynthesis and water use efficiency, so it is affected by temperature, light, CO_2 deviation pressure in air, environmental water conditions, and plant photosynthetic pathway type (Yoneyama et al., 1998). $\delta^{15}N$ in plants are affected by soil $\delta^{15}N$, and thus are more correlated with the fertilizer measurements and soil microbial communities (Lyu et al., 2021). Especially in the agricultural ecosystems, fertilizer is the key factors, because $\delta^{15}N$ in synthetic fertilizers are much lower than the nature abundance in soil, which will decrease the $\delta^{15}N$ (Choi et al., 2017). $\delta^{2}H$ and $\delta^{18}O$ in plant are mainly influenced by the habitat water stable isotope abundance, and the water related process in plant, hence, distance to the shoreline, latitude, precipitation, soil water contents and temperature could have influence on plant $\delta^{2}H$ and $\delta^{18}O$(Michener & Lajtha, 2008).

Interestingly, most of the important elements are rare earth (RE) elements, including 12 lanthanides (Tb, Dy, Er, Ho, Sm, Yb, Eu, Pr, La, Gd, Ce, and Nd) and yttrium (Y). The rest of the important elements were alkaline-earth elements (Sr and Mg) and transition elements (Cu, Ti, Fe, Nb, and Cd). Similarly, PLS-DA was conducted to select important variables in discriminating PDO-labeled Chinese yams (Tiegun) from others in the YeR. There were 17 variables selected (Fig. S1C), and 14 of them were the same as the above elements (including 9 lanthanides Tb, Dy, Sm, Eu, Pr, La, Gd, Ce, and Nd, 2 alkaline-earth elements Sr and Mg, and 3 transition elements Ti, Fe, and Nb), while U, Na, and $\delta^{15}N$ were different.

The characteristics of important elements in different river systems are shown in Fig. 2C. The elements were divided into two groups according to the complete linkage hierarchical cluster. The first group includes the 13 rare earth elements, and they accumulated the most in Chinese yams (Tiegun) from the Huai River system and the least in Chinese yams (Tiegun) from the Yellow River system. The other group includes the alkaline-earth elements and transition elements, and they accumulated the most in Chinese yams (Tiegun) from the Hai River system but the least in Chinese yams (Tiegun) from the Yellow River system. For Huaiqingfu and others in the YeR (Fig. S1C), the rare earth elements and U were higher in the Huaiqingfu area, while the alkaline

A

Correlation between elements dissimilarity (Euclidian) and:	r	p
Environmental distance	0.4481	0.001
Geographical distance	0.4399	0.001
Controlling for:		
Geographical distance	0.2071	0.001
Environmental distance	0.1849	0.001

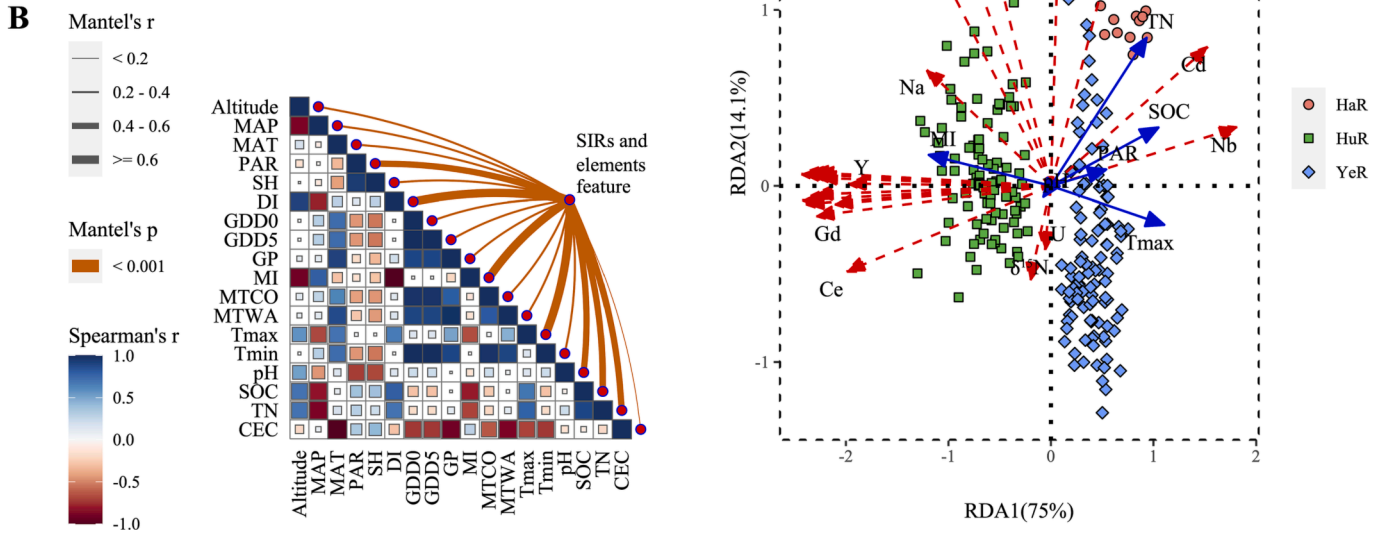


Fig. 3. Correlation of important elements with environmental factors.

earth and transition elements were lower in the Huaiqingfu area. Considering the Chinese yams (Tiegun) cultivated in different regions may have experienced different fertilization measurement according to the local soil backgrounds, the $\delta^{15}N$ was higher in the Huaiqingfu area than in the SXWR area was largely due to more organic fertilizer were used in the Huaiqingfu area than in SXWR (Lyu et al., 2021). Compare with chemical fertilizer, organic fertilizer could largely increase $\delta^{15}N$ in foods (Li et al., 2021). And this might also be one of the reasons to explain better quality of Chinese yams (Tiegun) in the Huaiqingfu region. Although the $\delta^{2}H$ and $\delta^{18}O$ were usually important geographic tracers, they revealed less difference in this study. This was because the majority of $\delta^{2}H$ and $\delta^{18}O$ in plant came from precipitation and the local water sources (Michener & Lajtha, 2008), while the study scale was small (<2° in latitude and 3° in longitude, Table S1), which lead to similar local $\delta^{2}H$ and $\delta^{18}O$ in the environment. Moreover, the SXWR and the Huaiqingfu region are both belong to the YeR, which further increased similarity of $\delta^{2}H$ and $\delta^{18}O$ in environments among them. The correlations between these important elements are also explored in Fig. 2D. Similar to the cluster results, the rare earth elements were tightly positively correlated, with correlation coefficients of 0.48–0.87. Meanwhile, Nb, Cd and Ti were negatively correlated with most elements, and Cu was insignificantly correlated with most elements.

Environmental effects on the elemental characteristics of the Chinese yam (Tiegun)

To assess the environmental effects on the elemental characteristics of the Chinese yam (Tiegun), a mantel test was first conducted to investigate the overall relationships between the selected elements in the above context with environmental factors and geographical distance. The 23 variables, including SIRs and elements selected in the former section, were employed as the dependent variable. Spearman's correlation coefficients for the element dissimilarity with geographical and environmental distances were 0.44 and 0.45, while the partial correlation coefficients were 0.18 and 0.21 (Fig. 3A), respectively. These results indicate that both environmental factors and geographical distance have significant effects on the elements and SIR characteristics of the Chinese yam (Tiegun), and their effects were similar. This result supports that elements and SIR features in plants could reveal their response to environmental factors (Baxter & Dilkes, 2012). Similarly, it has been suggested that plant element stoichiometry could reflect the environment more than genotype (Ågren & Weih, 2012). A control experiment on peanuts also revealed that the effect of the subregion on elemental contents was larger than that of variety (Zhao & Yang, 2022).

Based on this, the relationships between elemental characteristics and environmental factors were explored in detail. As shown in Fig. 3B, the elemental characteristics were significantly correlated with all of the environmental factors ($p < 0.001$), among which the correlation coefficient ranged from 0.17 to 0.72, and the most correlated environmental factors were the moisture index (MI, r = 0.72), dry index (DI, r = 0.71), and maximum temperature (Tmax, r = 0.62), followed by the soil organic carbon (SOC, r = 0.60), total nitrogen (TN, r = 0.47), pH (r = 0.44), and photosynthetically active radiation (PAR, r = 0.41). Although the parent materials for soil in the Yellow-Huai-Hai Plain are all acidic alluvial materials (S. Liu et al., 2010), soil features such as the SOC, TN and pH are important for plant element selection (Tian et al., 2019). Climate factors such as the MI, DI, Tmax, and PAR could affect strategies to adapt soil nutrients for plants (Baxter & Dilkes, 2012).

The highly correlated environmental factors were selected for further RDA to reduce the dimension, among which the DI and MI were extremely correlated with a Pearson's r of −1; hence, the MI, Tmax, SOC, TN, pH, and PAR were selected. As a result, the six environmental factors constrained 61.75% of the variance (the adjusted R^2), and the model was significant ($p < 0.001$) according to a permutation test by 999 permutations, with the six environmental factors as the independent variables and the 23 elements (except for U, for which the p value was 0.011) as the dependent variables. As shown in Fig. 3C, the first two axes accounted for 75.0% and 14.1% of the total variance in all environmental–element relationships, respectively. The rare earth elements were positively correlated with the MI and pH but negatively correlated with the Tmax, SOC, TN, and PAR. The rest of the elements were differently affected by the environmental factors. For example, Na, Fe, Mg, and Sr were positively correlated with the MI and negatively correlated with the Tmax but were less correlated with the TN, SOC, and

Table 1
Accuracy of different chemometric models.

Dataset		Accuracy (%)				
		LDA	PLS-DA	RF	SVM	KNN
Three river basins	Train (n = 191)	99.48	70.16	100.00	100.00	81.68
	Test (n = 48)	99.17	67.92	98.96	98.96	63.54
Huai yam or not in the YeR	Train (n = 88)	98.86	81.82	100.00	100.00	88.64
	Test (n = 23)	96.52	86.96	98.26	96.09	85.65

pH. However, Cu and Ti were positively correlated with the TN, SOC, and PAR but had less correlation with other factors. U and δ^{15}N were exactly opposite, and they were negatively correlated with the TN, SOC, and PAR. In addition, in contrast to rare earth elements, Cd and Nb were positively correlated with the TN, SOC, Tmax, and DI but negatively correlated with the pH and MI.

Geo-origin trace of Chinese yams (Tiegun) by a chemometric algorithm

In total, five chemometric models were established for the discrimination of Chinese yams (Tiegun) from different river basins (Table 1). As a result, the LDA, RF and SVM models revealed good performance, in which the mean accuracies for the test set (10 rounds) were 99.17%, 98.96%, and 98.96%, respectively. The other two models (PLS-DA and KNN) only revealed accuracies of 67.92% and 63.54%, respectively. Based on this, the confusion matrices of the LDA, RF, and SVM models were plotted (Fig. 4). For the LDA model, incorrect discrimination mostly occurred between the Huai River and Yellow River, while the SVM model usually incorrectly discriminated between the Hai River and Yellow River. The RF model predicted the Yellow River samples as Huai River and Hai River samples or predicted the Hai River samples as Yellow River samples.

Apart from the discrimination of three river basins, the discrimination models further contributed to discriminating samples in the Yellow River basin between the Huaiqingfu region and non-Huaiqingfu regions, as Huaiqingfu has historically been the geo-authentic region for the Chinese yam (Tiegun). Similarly, the LDA, RF, and SVM models revealed good performance, with accuracies in the test sets of 96.52%, 98.26%, and 96.09%, respectively. The confusion matrix (Fig. 4D-F) for these three models revealed that the LDA and SVM models both performed better in the non-Huaiqingfu group, while the RF model performed better in the Huaiqingfu region. For the RF model, only 1.11 ± 2.35% of the samples in the test set from the Huaiqingfu region were falsely discriminated as the *N*-Huai group in the total of 10 rounds of tests.

Conclusion

In this study, 4 SIRs and 44 elements were determined to explore the features of SIRs and multiple elements of Chinese yams (Tiegun) among different main production areas in the Yellow-Huai-Hai River plain and among the PDOs (ancient Huaiqingfu region) and others in the Yellow River basin. The results in this study revealed that SIRs and multielement features in Chinese yams (Tiegun) could indicate geographical origins. According to the PLS-DA, 20 and 17 SIRs and elemental variables were selected as the key variables in discriminating Chinese yams (Tiegun) from different river basins and PDOs, respectively. A total of 23 variables, including δ^{15}N, Tb, Dy, Er, Ho, Sm, Yb, Eu, Pr, La, Gd, Ce, Nd, Y, Sr, Mg, Cu, Ti, Fe, Nb, Cd, U, and Na, were used for further exploration. The Mantel test revealed a significant effect of both the environment and geographical distance on these variable features of Chinese yams (Tiegun). The moisture index (MI, r = 0.72), dry index (DI, r = 0.71), maximum temperature (Tmax, r = 0.62), soil organic carbon (SOC, r = 0.60), total nitrogen (TN, r = 0.47), soil pH (r = 0.44), and photosynthetically active radiation (PAR, r = 0.41) were the most effective factors, and the MI, Tmax, PAR, SOC, TN, and pH expressed 61.75% of the variance in SIRs and multielement features of Chinese yams (Tiegun).

In addition, the 23 variables were employed in 5 discriminate models

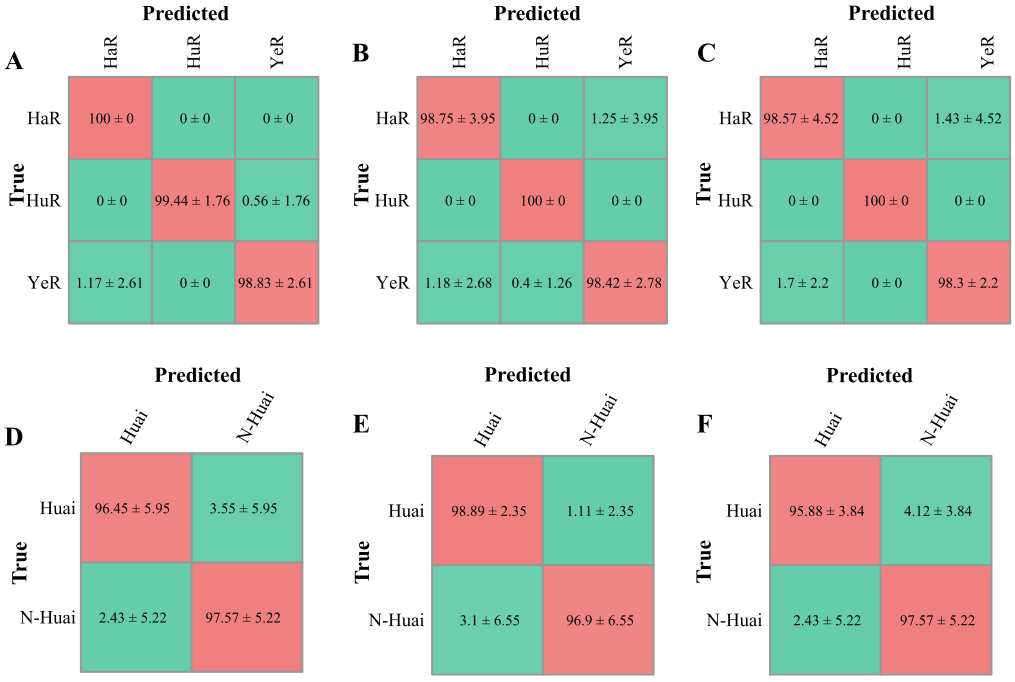


Fig. 4. The confusion matrix for the LDA (A), RF (B), and SVM (C) models in discriminating Chinese yams (Tiegun) from three river basins and for the LDA (D), RF (E), and SVM (F) models in discriminating Huai yams and *N*-Huai yams in the YeR basin. The data are expressed as the mean ± standard deviation.

to authenticate Chinese yams (Tiegun) from the 3 river basins and from the Huaiqingfu or non-Huaiqingfu regions. The LDA, RF and SVM models provided good accuracy (>96% in the test sets). For authenticating Chinese yams (Tiegun) from 3 river basins, the LDA model performed best with accuracies of 100%, 99.44% and 98.83% for the Hai River, Huai River, and Yellow River basins, respectively. The RF model revealed the highest accuracy in authenticating Chinese yams (Tiegun) with PDO labels, with accuracies of 98.89% and 96.90% for the PDO and non-PDO, respectively. The high accuracy of the models in testing sets indicates these models are possibly used in the real market for an unknown sample. Further, considering climate and soil factors have significant effects in the shaping of Chinese yam (Tiegun) SIRs and multielement features, climate change in the future may affect these features and thus the model accuracy. Hence, more data from different time points should be considered in future studies to improve the model's adaptability.

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.

Acknowledgment

This work was financially supported by the Scientific and technological innovation project of China Academy of Chinese Medical Sciences (Grant No. CI2021A04011, CI2021A04005), National Key R&D Program of China (Grant No. 2020YFC1712700), Fundamental Research Funds for the Central public welfare research institutes (ZZ15-YQ-059), the Shandong Provincial Key Research and Development Program (Major Technological Innovation Project) (Grant No. 2021CXGC010508), and the Key Project at Central Government Level: the ability establishment of sustainable use for valuable Chinese medicine resources (Grant No. 2060302).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fochx.2023.100678.

References

Ågren, G. I., & Weih, M. (2012). Plant stoichiometry at different scales: Element concentration patterns reflect environment more than genotype. *New Phytologist, 194* (4), 944–952. https://doi.org/10.1111/j.1469-8137.2012.04114.x

Baxter, I., & Dilkes, B. P. (2012). Elemental Profiles Reflect Plant Adaptations to the Environment. *Science, 336*(6089), 1661–1663. https://doi.org/10.1126/science.1219992

Choi, W.-J., Kwak, J.-H., Lim, S.-S., Park, H.-J., Chang, S. X., Lee, S.-M., … Kim, H.-Y. (2017). Synthetic fertilizer and livestock manure differently affect δ15N in the agricultural landscape: A review. *Agriculture, Ecosystems & Environment, 237*, 1–15. https://doi.org/10.1016/j.agee.2016.12.020

Coplen, T. B. (2011). Guidelines and recommended terms for expression of stable-isotope-ratio and gas-ratio measurement results. *Rapid Communications in Mass Spectrometry, 25*(17), 2538–2560. https://doi.org/10.1002/rcm.5129

Deng, X., Liu, Z., Zhan, Y., Ni, K., Zhang, Y., Ma, W., … Rogers, K. M. (2020). Predictive geographical authentication of green tea with protected designation of origin using a random forest model. *Food Control, 107*, Article 106807. https://doi.org/10.1016/j.foodcont.2019.106807

Dias, C., & Mendes, L. (2018). Protected Designation of Origin (PDO), Protected Geographical Indication (PGI) and Traditional Speciality Guaranteed (TSG): A bibiliometric analysis. *Food Research International, 103*, 492–508. https://doi.org/10.1016/j.foodres.2017.09.059

Epping, J., & Laibach, N. (2020). An underutilized orphan tuber crop—Chinese yam: A review. *Planta, 252*(4), 58. https://doi.org/10.1007/s00425-020-03458-3

Games, P. A., & Howell, J. F. (1976). Pairwise Multiple Comparison Procedures with Unequal N's and/or Variances: A Monte Carlo Study. *J Educ. Stat., 1*(2), 113–125. https://doi.org/10.3102/10769986001002113

Hao, X., Gao, F., Wu, H., Song, Y., Zhang, L., Li, H., & Wang, H. (2021). From Soil to Grape and Wine: Geographical Variations in Elemental Profiles in Different Chinese Regions. *Foods, 10*(12), 3108. https://doi.org/10.3390/foods10123108

Huang, H. (2021). *LinkET: Everything is Linkable. R Package Version 0.0. 4*.

Li, C., Wang, Q., Shao, S., Chen, Z., Nie, J., Liu, Z., … Yuan, Y. (2021). Stable Isotope Effects of Biogas Slurry Applied as an Organic Fertilizer to Rice, Straw, and Soil. *Journal of Agricultural and Food Chemistry, 69*(29), 8090–8097. https://doi.org/10.1021/acs.jafc.1c01740

Liu, S., Mo, X., Lin, Z., Xu, Y., Ji, J., Wen, G., & Richey, J. (2010). Crop yield responses to climate change in the Huang-Huai-Hai Plain of China. *Agricultural Water Management, 97*(8), 1195–1209. https://doi.org/10.1016/j.agwat.2010.03.001

Liu, X., Mu, J., Tan, D., Mao, K., Zhang, J., Ahmed Sadiq, F., … Zhang, A. (2022). Application of stable isotopic and mineral elemental fingerprints in identifying the geographical origin of concentrated apple juice in China. *Food Chemistry, 391*, Article 133269. https://doi.org/10.1016/j.foodchem.2022.133269

Lyu, C., Yang, J., Wang, T., Kang, C., Wang, S., Wang, H., … Guo, L. (2021). A field trials-based authentication study of conventionally and organically grown Chinese yams using light stable isotopes and multi-elemental analysis combined with machine learning algorithms. *Food Chem, 343*, Article 128506. https://doi.org/10.1016/j.foodchem.2020.128506

Michener, R., & Lajtha, K. (2008). *Stable isotopes in ecology and environmental science*. John Wiley & Sons.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R. B., … Wagner, H. (2013). Package 'vegan'. *Community Ecology Package, Version, 2*(9), 1–295.

Skrzypek, G., Allison, C. E., Böhlke, J. K., Bontempo, L., Brewer, P., Camin, F., … Dunn, P. J. H. (2022). Minimum requirements for publishing hydrogen, carbon, nitrogen, oxygen and sulfur stable-isotope delta results (IUPAC Technical Report). *Pure and Applied Chemistry, 94*(11–12), 1249–1255. https://doi.org/10.1515/pac-2021-1108

Tian, D., Reich, P. B., Chen, H. Y. H., Xiang, Y., Luo, Y., Shen, Y., … Niu, S. (2019). Global changes alter plant multi-element stoichiometric coupling. *New Phytol, 221* (2), 807–817. https://doi.org/10.1111/nph.15428

Varrà, M. O., Husáková, L., Patočka, J., Ghidini, S., & Zanardi, E. (2021). Multi-element signature of cuttlefish and its potential for the discrimination of different geographical provenances and traceability. *Food Chemistry, 356*, Article 129687. https://doi.org/10.1016/j.foodchem.2021.129687

Wang, N., Dai, Y., Yuan, Y., Guo, Z., Kong, 李婷, Jia, X., … Shan, X. (2018). Analysis on Historical Origin of Dioscoreae Rhizoma and Preliminary Construction of Its Standard System. *Chinese Journal of Experimental Traditional Medical Formulae 24* (04), 222–228. https://doi.org/10.13422/j.cnki.syfjx.2018040222.

Wassenaar, L. I., & Hobson, K. A. (2003). Comparative equilibration and online technique for determination of non-exchangeable hydrogen of keratins for use in animal migration studies. *Isotopes in Environmental and Health Studies, 39*(3), 211–217. https://doi.org/10.1080/1025601031000096781

Wei Lin-Feng, W.-L.-F., Hu Xiao-Fei, H.-X.-F., Cheng Qi, C. Q., Wu Xing-Qi, W.-X.-Q., & Ni Jian, N. J. (2022). A dataset of spatial distribution of bioclimatic variables in China at 1km resolution. *Science Data Bank Datasets*. https://doi.org/10.11922/sciencedb.01434

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-Food Science and Technology, 167*, Article 113871. https://doi.org/10.1016/j.lwt.2022.113871

Yoneyama, T., Fujiwara, H., & Wilson, J. M. (1998). Variations in fractionation of carbon and nitrogen isotopes in higher plants: N metabolism and partitioning in phloem and xylem. *Stable Isotopes*. Garland Science.

Zhang, L., Bai, B., Liu, X., Wang, Y., Li, M., & Zhao, D. (2011). α-Glucosidase inhibitors from Chinese Yam (Dioscorea opposita Thunb.). *Food Chemistry, 126*(1), 203–206. https://doi.org/10.1016/j.foodchem.2010.10.100

Zhao, H., & Yang, Q. (2022). Study on influence factors and sources of mineral elements in peanut kernels for authenticity. *Food Chemistry, 382*, Article 132385. https://doi.org/10.1016/j.foodchem.2022.132385