Acta Scientific NUTRITIONAL HEALTH (ISSN:2582-1423)

Volume 9 Issue 1 January 2025

Research Article

Predictive Model to Authenticate the Origin of Pakchoi in Northern Vietnam through Metal Content from the Soil

Mac The Vinh[1,2], Truong Ngoc Minh[3] and Nguyen Thi Minh Tu[1,4]*

[1]*School of Chemistry and Life Sciences, Hanoi University of Science and Technology, Vietnam*

[2]*Hanoi University of Industry, Vietnam*

[3]*Center for High Technology Research and Development - Vietnam Academy of Science and Technology, Vietnam*

[4]*New Industry Creation Hatchery Center (NICHe), Tohoku University, Sendai, Miyagi, Japan*

***Corresponding Author:** Nguyen Thi Minh Tu New Industry Creation Hatchery Center (NICHe), Tohoku University, Sendai, Miyagi, Japan.

Received: December 16, 2024

Published: December 30, 2024

Abstract

Pakchoi (*Brassica rapa* subsp. Chinensis) is widely grown in Vietnam and many other Southeast Asian countries. Brassica rapa has long been known for its ability to accumulate many types of metals. Along with the metal functional type in the soil, the main source of the metal functional type of Brassica rapa, the metal composition becomes a key to support the determination of the geographical origin of the plant. Our study has built an analytical model and predicted the geographical location of the Brassica rapa growing area based on the metal composition in the soil of the crop. Twenty-seven metals in the soil of 9 regular cultivation sites serving the Hanoi market from September 2021 to May 2024 were determined by ICP-MS, and then the data were analyzed and a predictive model was built using the LDA analysis method. The results of this study provide a target origin model for improvement, as well as a way to identify the origin of other vegetable crops in the food market in Northern Vietnam.

Keywords: Hanoi; Soil; LDA; ICP-MS; Pakchoi; Predictive Model

Abbreviations

ICP-MS: Inductively Coupled Plasma Mass Spectrometry; LDA: Linear Discriminant Analysis

Introduction

Pakchoi (*Brassica rapa* subsp. Chinensis), is a plant widely grown in Vietnam and Southeast Asian countries such as China, Malaysia, etc. Pakchoi is heat-resistant, moisture-resistant, and relatively resistant to soil-borne pests. Used as a familiar daily food, Pakchoi is selected because of its yield of 10-20 tons/ha for small varieties and 20-30 tons/ha for large varieties [1]. Pakchoi is a vegetable that contains many flavonoids - a group of bioactive compounds that are beneficial to health with 633-982 µg/g carotenoids, of which lutein accounts for 40-43%, Violaxanthin 17-28%, Neoxanthinmade 13%, β-carotene 19% - 27%. Vitamin C or ascorbic acid accounts for about 9.2% (dry season) to 13.8% (rainy season). Pakchoi also has many vitamins A, B, C, minerals: Calcium, Manganese, Potassium, Zinc, Iron, Sodium, Mg, Selenium, Phosphorus. In which, calcium content ranges from 21.8 mg/g to 36.9 mg/g, iron from 534 µg/g to 644 µg/g, zinc from 49 µg/g to 61 µg/g, glucosinolate content from 12.2 - 21.0 µmol/g [2]. In Vietnam, the Pakchoi variety grown and used is mainly the Green Fortune variety. Other Pakchoi varieties include Tai sai, White Pakchoi, Green Fortune [3], Dwarf Carton White, Joi choi, Meiquin... are rarely planted or not planted at all, partly because their disease resistance and heat tolerance are not as good as Green Fortune. Among them, the Pakchoi variety grown and used in Vietnam is mainly the Green Fortune variety due to its good disease resistance and heat tolerance.

Although Pakchoi is resistant to pests and heat, it is not as easy to grow as some other mustards. This leads to the fact that Pakchoi in Vietnam is rarely grown spontaneously but mainly grown in some specialized cooperatives, then harvested and brought to the market for sale. Thus, the location of Pakchoi is often quite fixed.

In addition, Pakchoi has the characteristics of Brassica species, which can accumulate excessively [4]. This makes it possible to use the metal content and ratio in the crop and surrounding factors such as soil, water, manure, air, etc. as the key to determine the geographical origin of Pakchoi growing areas.

Materials and Methods

Material

Soil samples

Soil was collected from Bac Tu Liem, Chuong My, Dan Phuong, Long Bien, Me Linh, and Thuong Tin districts in Hanoi, Vietnam, in the 5th week of 6 harvests from September 2021 to May 2024 and from Dong Hoi, Van Noi, and Thanh Tri districts in Hanoi, Vietnam, in the 5th week of 6 harvests from January 2022 to May 2024.

Soil was collected at a depth of 5 cm from the surface. The samples were then coarsely ground in a grinder and dried at 70 ^{o}C until the moisture content was approximately 15% before being ground into powder. Finally, the samples were stored in clean, labeled plastic bags and stored in a dry place.

Samples digestion

Weigh 0.1 g of soil samples from the respective plastic sample bags into polytetrafluoroethylene (PTFE) teflon tubes, add 10 ml of 65% HNO_3 and pre-digest for at least 24 h in a 50 ml falcon tube in a laboratory fume hood. The soil samples were then digested in a Mars 6 microwave digester according to a preset program. Accordingly, the samples were heated to 175^{o}C for 5 min 30 s and held for 4 min 30 s before cooling to room temperature. After cooling, the solution was filtered and diluted to 1000 ml in a volumetric flask with ultrapure water. The samples were then collected for metal content analysis by Inductively coupled plasma mass spectrometry (ICP-MS) equipment. It should be noted that the absolute moisture content of the samples was measured before microwave digestion.

Chemicals

Ultrapure deionized water (resistivity of 18.2mΩcm) was obtained from a Milli-Q plus water filtration system (Millipore, Bedford, MA, USA). Nitric acid 65% (HNO_3) was purchased from Merck, USA. Hydrogen peroxide 30% (H_2O_2) was purchased from Fisher Chemical. ICP-MS method was run with a multi-component standard solution including Li, Bo, Mg, Al, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, As, Rb, Sr, Nb, Mo, Ag, Cd, Sb, Cs, Ba, Hg, Tl, Pb, Bi, (TraceCert).

Methods

Microwave digestion

The soil samples, after being harvested and prepared, was put into the Microwave Digestion System – MARS 6. Digestion was performed according to EPA method 3050B. After that, the samples were filtered using filter paper with filter holes size 15-20 µm and titrated to 100ml using a volumetric flask. Samples were finally stored to measure metal content using ICP-MS equipment.

IPC-MS method

The equipment used was an Agilent 7900 ICP-MS instrument (Agilent Technologies, Tokyo, Japan). ICP-MS analysis parameters were RF source at 1550 W, RF coupling at 2.0 V, cell entrance at -40 V, cell exit -60 V, cell energy discrimination at 5.0 V, spray chamber temperature at 2 °C. Carrier gas was used was Argon at a flow of 1.09 L/min. The auxiliary gas was used was helium at 4.3 L/min. Quantitative data were obtained with respect to diverse standards matched to matrices prepared in 1% HNO_3. Instrumental detection limits in this study were calculated using data Raw material from standard and blank sample (using ultrapure 2% nitric acid) and calculated according to the following formula: IDL = 3SDblank x Cx/(Sx-Sblank). Where, SDblank is the standard deviation of the intensity of multiple blank measurements, Cx is the average signal used for the standard, SX is the signal for Cx and Sblank is the signal for the blank sample.

Data analysis method

The data in this study were the average values of three analyses of the metal content in a sample of Pakchoi soil. Data from soil samples grown in different regions was then processed using Linear Discriminant Analysis (LDA) [5]. Data were analyzed using Microsoft Excel in Office 365, upgraded with XLSTAT software version 2024.

Result and Discussion

To develop a model to distinguish groups of data using the LDA, we prepare data on metal content in soil in 3 years: A model-building group and a prediction group. In which the model building group is selected as the results of metal composition of soil in 6 regions of Bac Tu Liem, Chuong My, Dan Phuong, Long Bien, Me Linh, and Thuong Tin from January 2022 to May 2024 (3 locations in each region) and 3 regions of Dong Hoi, Van Noi, and Thanh Tri from February 2022 to May 2024. The selected prediction group is the result of the metal composition of soil in 6 regions of Bac Tu Liem, Chuong My, Dan Phuong, Long Bien, Me Linh, and Thuong Tin from May 2021 to December 2021 (3 locations in each region) and 3 regions of Dong Hoi, Van Noi, and Thanh Tri from January 2023 to December 2023 (duplicate samples were prepared separately, not identical to the original samples already in the construction group).

Wilks Lambda test allows testing whether the vectors of mean values for different groups are equal or not. In table 1,2, we can see that the difference between the vectors of mean values of the groups is significant. The Eigenvalues table in table 3 shows the eigenvalues and the corresponding % variance. Accordingly, 40.79% of the variance is represented by the first factor, and 34.74% of the variance is represented by the second factor. These two factors were able to represent 75% of the variance.

The variables chart in figure 1 shows the correlation of the original variables with two factors, F1 and F2. Factor F1 is correlated

	Location	Location	
Me Linh	21°10'19.7"N 105°45'38.3"E 21°10'27.6"N 105°45'31.7"E 21°10'36.4"N 105°45'39.9"E	Chuong My	20°51'45.6"N 105°39'40.0"E 20°51'55.3"N 105°39'25.9"E 20°52'02.0"N 105°39'40.0"E
Bac Tu Liem	21°03'56.1"N 105°45'16.2"E 21°03'58.9"N 105°45'04.8"E 21°04'04.7"N 105°45'17.2"E	Dong Hoi	21°04'20.8"N 105°51'59.1"E 21°04'23.7"N 105°52'02.0"E 21°04'14.2"N 105°52'16.5"E
Thuong Tin	20°46'42.6"N 105°54'11.4"E 20°46'39.5"N 105°53'57.5"E 20°46'42.7"N 105°54'06.7"E	Van Noi	21°09'18.0"N 105°49'12.1"E 21°09'15.5"N 105°49'19.7"E 21°08'05.5"N 105°48'59.0"E
Long Bien	21°04'08.6"N, 105°51'16.4"E 21°04'06.5"N, 105°51'17.4"E 21°04'07.9"N, 105°51'19.6"E	Thanh Tri	20°55'56.0"N 105°51'15.1"E 20°56'23.3"N 105°50'10.4"E 20°56'18.2"N 105°50'10.0"E
Dan Phuong	21°05'06.0"N 105°42'22.3"E 21°05'08.9"N 105°42'25.4"E 21°05'11.8"N 105°42'24.2"E		

Table 1: Pakchoi's planting locations.

Lambda	0.000
F (Observed value)	224.366
F (Critical value)	1.171
DF1	216
DF2	2793.771
p-value (Two-tailed)	<0.0001
alpha	0.05

Table 2: Wilks' Lambda test of Pakchoi's soil form 9 regions in LDA

Wilks' Lambda test (Rao's approximation):

*Signification codes: 0 < "***" < 0.001 < "**" < 0.01 < "*" < 0.05 < "." < 0.1 < " " < 1*

Test interpretation: H0: The means vectors of the 9 classes are equal.

Ha: At least one of the means vector is different from another.

As the computed p-value is lower than the significance level alpha=0.05, one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

	F1	F2	F3	F4	F5	F6	F7	F8
Eigenvalue	132.520	112.854	26.485	23.649	21.161	6.431	1.500	0.244
Discrimination (%)	40.795	34.741	8.153	7.280	6.514	1.980	0.462	0.075
Cumulative % (%)	40.795	75.536	83.689	90.969	97.483	99.463	99.925	100.000

Table 3: Eigenvalue table of Pakchoi's soil with LDA.

with the metal content of Li, Co, Fe, Ba, Hg, Mg, Nb, Ti, Ag, Sr, and Cd, while Factor F2 is correlated with the metal content of Pb, Cr, Ni, Cu, Rb, Mn, Mo, V, Al, As, Sb, Bi, and Cs. The observation chart in figure 2-4 shows us the distribution of soil samples on the chart built from two factors, F1 and F2. Soil samples from different areas are in separate clusters with different colors. The values of predicted samples are also shown on the chart with different color compared to the standard samples in the same area. An ellipse drawn based on the centroids of the samples at the 95% confidence interval shows the range of likelihood of the soil samples occurring.

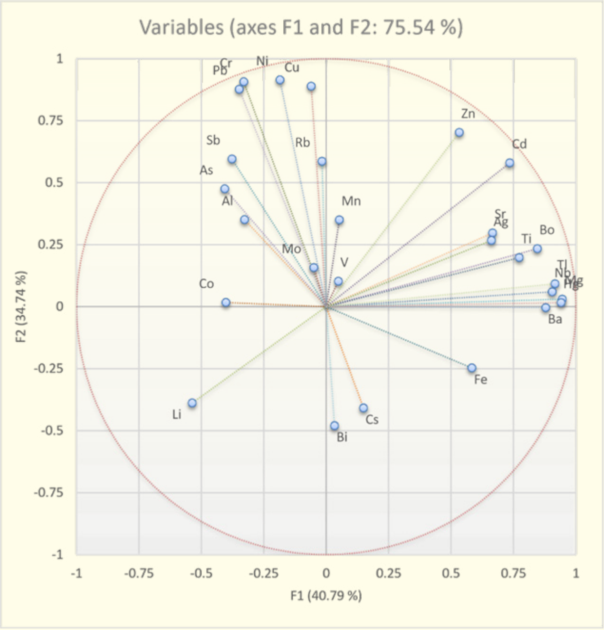

Figure 1: Variables chart from the LDA method.

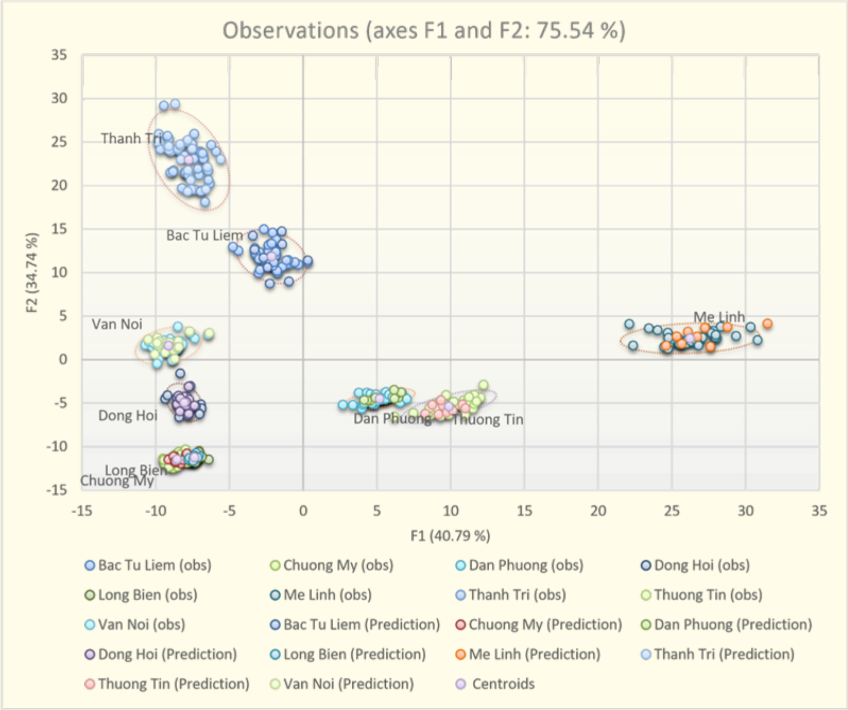

Figure 2: Observations chart obtained from the LDA method.

Figure 3

The confusion matrix summarizes the reclassification of observations and allows a quick look at the % of observations that are classified correctly, which is the ratio of the number of observations that are classified correctly to the total number of observations. Here, we observe 396 random soil metal samples from different vegetable-growing areas in the entire dataset. The Confusion matrix for the cross-validation results in table 4 shows that there are no incorrect observations here.

From\to	Bac Tu Liem	Chuong My	Dan Phuong	Dong Hoi	Long Bien	Me Linh	Thanh Tri	Thuong Tin	Van Noi	Total	% Correct
Bac Tu Liem	45	0	0	0	0	0	0	0	0	45	100.00%
Chuong My	0	45	0	0	0	0	0	0	0	45	100.00%
Dan Phuong	0	0	45	0	0	0	0	0	0	45	100.00%
Dong Hoi	0	0	0	42	0	0	0	0	0	42	100.00%
Long Bien	0	0	0	0	45	0	0	0	0	45	100.00%
Me Linh	0	0	0	0	0	45	0	0	0	45	100.00%
Thanh Tri	0	0	0	0	0	0	42	0	0	42	100.00%
Thuong Tin	0	0	0	0	0	0	0	45	0	45	100.00%
Van Noi	0	0	0	0	0	0	0	0	42	42	100.00%
Total	45	45	45	42	45	45	42	45	42	396	100.00%

Table 4: Confusion matrix for the cross-validation results.

Using the predicted sample to test the pattern recognition ability of the newly built model, we get the result that the model correctly recognizes 100% of the prepared samples. Thus, the LDA model promises high ability in distinguishing soil samples for growing Pakchoi.

Conclusion

Twenty-seven metals in soil at a total of 9 Pakchoi regular cultivation sites serving the Hanoi market from September 2021 to May 2024 were analyzed by ICP-MS. Using the linear method of discriminating the same composition and ratio of metals in the soil of growing pakchoi, the research team built an analytical model to predict the geographical location of the growing area of pakchoi. The model gave good discrimination results up to 100% and completely recognized the predicted samples. This research result provides a model to help identify the geographical origin of Pakchoi through the metal composition of the soil, as well as a reference for determining the origin of other vegetable crops in the food market.

Acknowledgements

This research was funded by the Vietnam Academy of Science and Technology under grant number NCXS.

Bibliography

1. Li Chia-wen. "Chinese Cabbages of China. Beijing, China: Agricultural Publishing Company (1984).

2. Hanson P., *et al*. "Contents of Carotenoids, Ascorbic Acid, Minerals and Total Glucosinolates in Leafy Brassica Pakchoi (Brassica Rapal.Chinensis) as Affected by Season and Variety". *Journal of the Science of Food and Agriculture* 89.5 (2009): 906-914.

3. Kalisz Andrzej., *et al*. "Yield and Nutritional Quality of Several Non-Heading Chinese Cabbage (Brassica Rapa Var. Chinensis) Cultivars with Different Growing Period and Its Modelling". *Korean Journal of Horticultural Science and Technology* 30.6 (2012): 650-656.

4. Malecka A., *et al*. "Insight into the Phytoremediation Capability of *Brassica Juncea* (V. Malopolska): Metal Accumulation and Antioxidant Enzyme Activity". *International Journal of Molecular Sciences* 20.18 (2019).

5. Esteki M., *et al*. "Use of Spectroscopic Methods in Combination with Linear Discriminant Analysis for Authentication of Food Products". *Food Control* 91 (2018): 100-112.