WILEY | Hindawi

Research Article

Identification of Geographical Origin of Milk by Amino Acid Profile Coupled with Chemometric Analysis

Min Kang ⓘ, **Quanhong Yue** ⓘ, **Shikuan Jia** ⓘ, **Jiarong Wang** ⓘ, **Mengdan Zheng** ⓘ, **and Ran Suo** ⓘ

College of Food Science and Technology, Hebei Agricultural University, Baoding 071000, China

Correspondence should be addressed to Ran Suo; ransuo@qq.com

This study aimed to establish a method to identify the geographical origin of milk based on its amino acid profile. High-performance liquid chromatography (HPLC) was carried out to measure amino acid contents. The significant differences of amino acid profiles of milk samples from four regions in China (Hebei, Ningxia, Heilongjiang, and Inner Mongolia) were analyzed by ANOVA. Furthermore, the principal component analysis (PCA) demonstrated the feasibility of geographical origin identification using an amino acid profile, which the first 2 principal components account for 65.62% of total variance. The predictive model for the geographical origin of milk samples was established by orthogonal partial least squares-discriminant analysis (OPLS-DA) with a classification accuracy of 100% and the performance parameters of R^2X 0.98, R^2Y 0.82, and Q^2 0.75. The excellent predictive ability of the model was validated using the validation data set. The analysis of variable importance in projection (VIP) showed that seven amino acids played a key role in the geographical origin identification. This method is a reliable strategy to identify the geographical origin of milk for protecting consumers against mislabeling fraud.

1. Introduction

Milk and its products have become an indispensable part of people's life which provides about 20% of the total protein consumed across the world. Almost all the land on earth has pasture which keeps cows and sheep. The best pastures in the world are concentrated on the temperate grassland at about 40–50 degrees north and south latitudes, internationally recognized as the "golden belt of milk source," where the climate and environment are in favor of the growth of forages and cows. The output, quality, and nutritional composition of milk are closely concerned with the forage quality and cows' body health. Cows with a comfortable and healthy lifestyle can provide high-quality milk. The golden belt of milk sources in China is mainly located in the grasslands of the Northeast, Northwest, and North China, such as Hebei, Ningxia, Inner Mongolia, Heilongjiang, and Xinjiang. These milk sources provide more than 60% of the raw milk in China.

The high quality of milk is generally processed into high-value products, such as infant milk powder produced from the milk sources of New Zealand. The upscale market share of dairy products in the world is almost completely occupied by the milk sourced from these belts. The nutritional and economic value of milk and dairy products is often associated with their geographical origin. Just as the price of wine, coffee, and tea varies enormously depending on where they come from, customers are willing to pay more for the products from some specific geographical regions with favorable acceptance as reliable quality criteria. An example of a preeminent dairy product is the "Emmentaler Switzerland" cheese from the alpine regions, which has the status of a "Protected Designation of Origin (PDO)" and a considerable premium price over Emmental cheese produced elsewhere [1].

There is an increasing demand of robust analytical techniques for the geographical origin traceability of dairy products that can be utilized by regulatory authorities to ensure fair competition and protect consumers against fraud

due to mislabeling. So far, many methods have been developed for identifying the geographical origin of foodstuffs [2, 3]. Isotope ratio mass spectrometry (IRMS) coupled with chemometric methods is the most promising techniques, which has been widely used to determine the authenticity and geographical origin of dairy products [4–7]. The values of the stable isotope ratio of milk vary as the function of environmental factors and animal feeding regimes, which provide a proof correlated with the geographical origin of milk. A pilot study was conducted to evaluate the suitability of multielement isotope ratio analysis for determining the origin of cows' milk from seven dairying regions in Australia and New Zealand. Each milk sample displayed a distinct fingerprint of isotopic ratios of $\delta^{13}C$, $\delta^{15}N$, $\delta^{18}O$, $\delta^{34}S$, and $\delta^{87}Sr$. The potential of IRMS has been verified for determining the geographical origin of dairy products produced within Australasia [8]. The stable isotope ratios of $\delta^{13}C$ and $\delta^{15}N$ for the milk samples were from different Italian origins, and their fractions (fat, casein, and whey) were used to develop a new analytical approach that can simultaneously discriminate milk samples according to their geographical origin and type of processing [9]. By using $\delta^{13}C$ and $\delta^{15}N$ values of extracted proteins and $\delta^{2}H$ and $\delta^{18}O$ values of milk water, IRMS was applied to identify the geographical origin of pure milk from Australia and New Zealand, Germany and France, the USA, and China [10]. Using $\delta^{13}C$, $\delta^{15}N$, $\delta^{2}H$, and $\delta^{18}O$ values to specifically assign geographical origin, Zhao et al. [11] studied the traceability accuracy of cow milk samples from various provinces in China. It was found that different geographical locations with a separation distance greater than 0.7 km can be distinguished using multi-element (C, N, H, and O) stable isotope ratio analysis. In addition, stable isotopic ratios analysis in combination with other techniques, such as inductively coupled plasma atomic emission spectroscopy (ICP-AES), inductively coupled plasma mass spectrometry (ICP-MS), and nuclear magnetic resonance (NMR), was a hopeful way for the geographical origin determination of milk [12, 13]. Regarding dairy products, the determination of the geographical origin of cheeses [14–16], butter [17], and milk powder [18] was also successfully carried out by IRMS coupled with chemometric analysis.

It has been demonstrated that isotope ratio analysis is a powerful tool for the traceability and identification of milk and dairy products. However, this technique has some limitations, such as the high cost of sample analysis and the high price of the instrument [5, 19]. The objective of this work was to provide a new low-cost method for identifying the geographical origin of milk and dairy products. This method was based on the characteristics of the amino acid profile of raw milk, as same as the isotope ratio value which was correlated with the living environment and feeding regimes of cows. This procedure was carried out by amino acid analysis using high-performance liquid chromatography (HPLC) coupled with chemometric analysis to develop a classification model for the geographical origin of milk. The method could be used to prevent from the mislabeling fraud of the geographical origin of milk.

2. Materials and Methods

2.1. Sample Collection.
Cow's milk was sampled from dairy farms located at four Chinese provinces (Hebei, Ningxia, Heilongjiang, and Inner Mongolia) in August 2018. A total of 178 fresh milk samples were collected to ensure a representative data set (Appendix 1). The collected samples were transported to the laboratory by cold chain logistics. Milk samples were kept frozen at −20°C until preparation. In addition, three samples were purchased from markets and had production origin labels. That is, A, B, and C were labeled as Inner Mongolia, Heilongjiang, and Ningxia, respectively.

2.2. Sample Preparation.
An approximate 50–100 g (fresh weight) homogeneous milk sample was placed on a glass plate and lyophilized for 24 h to dry powder. After being ground, 0.2 g of freeze-dried milk powder was weighed into a 12 mL glass vial, added 10 mL of 6 mol/L HCL solution (containing 1 g/L of phenol) and screwed the cap tightly. Subsequently, the hydrolysis was carried out in an air-blowing thermostatic oven at 110°C for 24 h.

1 mL of hydrolysate was transferred into a 50 mL eggplant-shaped flask and evaporated in vacuo to dryness at 70°C. The residue obtained was redissolved with 2 mL of 0.1 mol/L HCL solution by vortex mixing and passed through a 0.45 μm inorganic filter membrane. 100 μL of hydrolysate filtrate, 200 μL of buffer solution (pH 9.0), and 100 μL of derivative agent (300 mg/mL of 2,4-dinitrochlorobenzene) were successively transferred into a 1.5 mL glass vial and screwed the cap tightly. After vortex mixing, the derivatization was carried out in a thermostatic oven at 90°C for 90 min.

After derivatization, the solution obtained was adjusted to pH 7 with 50 μL of 10% (V/V) acetic acid and diluted with 600 μL of ultrapure water. Finally, the derivative solution was filtered through a 0.45 μm organic membrane for subsequent analysis by HPLC.

2.3. Analysis of Amino Acids.
The milk sample set was analyzed by HPLC (2695, Waters Ltd., America) equipped with a C18 column (4.6 mm * 250 mm * 5 μm, Kromat Universil, America) and a PDA detector. The chromatography was carried out with the following conditions: column temperature 40°C; detection wavelength 360 nm; flow rate 1 mL/min; injection volume 10 μL; mobile phase A was acetonitrile, mobile phase B was buffer solution (0.03 mol/L sodium acetate solution, 0.15% triethylamine, pH 5.25 ± 0.05), gradient elution procedure (0 ~ 10 min, 18% A; 10 ~ 15 min, 18% ~ 20% A; 15 ~ 30 min, 20% ~ 34% A; 30 ~ 35 min, 34% ~ 45% A; 35 ~ 38 min, 45% ~ 55% A; 38 ~ 42 min, 55% ~ 60% A; 42 ~ 45 min, 60% ~ 18% A).

Seventeen amino acids were selected for determination. The 17 analytes selected by short name were Asp, Thr, Ser, Asn, Glu, Gln, Pro, Gly, Ala, Cit, Abu, Val, Met, Ile, Leu, Tyr, Phe, Lys, His, and Arg (in elution order).

2.4. Statistical Analysis. The data acquired by HPLC were exported to Microsoft Excel (Microsoft Corp., USA), The amino acid ratio, i.e., proportion of each amino acid content in total 17 amino acid content, was used for multivariate data analysis. Statistical analysis of amino acid was performed using SPSS 22.0, and a post hoc Duncan's test of analysis of variance (ANOVA) was performed to determine significant differences between samples from different origins. The preprocessed data was subjected to principal component analysis (PCA) by the Origin software package (Northampton, MA, USA). For modeling, the samples of each class were divided into calibration sets (125 samples) and validation sets (53 samples) by applying the Kennard–Stone (KS) uniform sampling algorithm. A calibration set was used to develop a model by using the supervised technique OPLS-DA from SIMCA 14.2 software (Umetrics, Umeå, Sweden).

The permutation test ($n = 200$) was used to evaluate whether the data was overfitted or not. Furthermore, 7-fold cross-validation was run, and its validation metrics were Q^2 and the lowest root mean square error of cross-validation (RMSECV). The external validation was performed according to a previously reported procedure [20, 21]. The external validation data set was imported into the developed model under the "specify toolbar" of SIMCA. Its validation metrics were the correct discrimination rate and receiver operating curve (ROC). The area under the curve (AUC) of the ROC illustrates the method performance; the closer to 1 the value is, the better the performance is.

3. Results and Discussion

3.1. Differences in Amino Acid Profile of Milk from Different Geographical Origins. The content of 17 amino acids in milk samples was determined by HPLC. The amino acid ratios, i.e., proportion of each amino acid content to the total 17 amino acid content, were calculated and shown in Table 1. Unlike the amino acid content, the value of amino acid ratio is only related to the protein composition, which is not affected by other components, such as fat, in milk samples. So the amino acid ratio, rather than the amino acid content, was selected to show the amino acid profile of milk. The differences in amino acid profiles of milk samples from four regions were analyzed by post hoc Duncan's test of ANOVA. The ratios of Asp, Cys, and Ala of samples were significantly different between four regions ($P < 0.05$). The ratios of Glu and His in the samples were significantly different between three of the four regions ($P < 0.05$). The amino acid profiles of samples from Hebei and Ningxia were relatively similar. The results of ANOVA indicate the potential feasibility of using amino acid ratios as an indicator of geographical origin traceability. It was reported that the amino acid profile of milk is linked to feed [22]. The differences in pasture conditions, such as forage quality, feeding strategy, and climate, lead to the differences in the amino acid profiles of milk from different geographical origins.

3.2. Potential of Geographical Origin Classification Based on Amino Acid Profiles. PCA was first applied for data visualization, which demonstrated the general potential to differentiate between the geographical origin of milk samples using amino acid profiles. PCA is the most commonly used variable-reduction method, which decomposes the data matrix with n rows (samples) and P columns (variables) into the product of a score matrix [22]. All principal components (PCs) are mutually orthogonal. Each successive PC contains less of the total variability of the initial data set, and the scores show the position of samples in the space of the PCs.

PCA was carried out on the 178-sample data set of 4 geographical origins. The scores are plotted as a multiclass model; i.e., each geographical origin of the milk sample is separately presented as a class (Figure 1). The first two PCs accounted for 65.62% of the total cumulative variance, which interpreted a majority of the variables from the raw data. Samples from Inner Mongolia are well separated from the other three groups of samples. The three groups of samples from Hebei, Ningxia, and Heilongjiang overlapped to a certain extent, which was consistent with the results of the previous ANOVA. The results suggested that the amino acid profiles had the potential for the identification of geographical origin.

3.3. Establishment of the OPLS-DA Model. As an unsupervised chemometric method, PCA just shows the data as they are, which is frequently seen as a practical indicator for the potential of OPLS-DA model [23]. Conversely, OPLS-DA is a supervised chemometric method that can determine features within data and is explicitly oriented to address particular issues, such as food authentication and geographical origin traceability. By constructing the predictive models, OPLS-DA can separate and classify new data points, which allows it to be used as a nontargeted method to analyze whether an unknown sample is accepted or rejected by a predefined class. Through orthogonal signal correction (OSC), by filtering out the useless orthogonal information in the independent variable X matrix which is not related to the dependent variable Y, the correlation between X and Y is strengthened, and the explanatory ability and accuracy of predicting model are improved. [24, 25].

The predicting model of OPLS-DA was established using the calibration data set with unit variance scaling and principal components of "$3 + 8 + 0$." It was shown that there was a clear clustering of milk samples from different regions with obvious separation boundaries (Figure 2(a)). Note that the new variables t1 and t2 summarize the X-variables. Score t1, which is the first component, corresponds to the largest variation of the X space, followed by t2, and so on. Inner Mongolia samples were negatively affected by t1, while samples from other provinces both had positive score values for t1. Ningxia, Hebei, and Heilongjiang samples were separated according to t2, where Heilongjiang samples were negatively affected. The prediction performance of the model was frequently assessed according to the cumulative coefficient of determination (R^2 (cum)) and cross-validated coefficient of determination (Q^2 (cum)). R^2 evaluates the fitting degree and Q^2 indicates the predictability. And R^2 values were evaluated based on their components attributed to the input variables (R^2X (cum)) and

TABLE 1: Amino acid ratios (mean values ± SD, %) of milk samples from different regions.

Amino acids	Hebei	Inner Mongolia	Ningxia	Heilongjiang
Asp (%)	9.11 ± 0.24^b	7.80 ± 0.76^d	9.58 ± 0.21^a	8.50 ± 0.30^c
Glu (%)	22.78 ± 0.73^b	20.94 ± 1.61^c	23.83 ± 0.46^a	21.04 ± 0.94^c
His (%)	2.60 ± 0.24^a	2.67 ± 0.23^a	2.50 ± 0.19^b	2.17 ± 0.14^c
Ser (%)	5.66 ± 0.47^{bc}	5.82 ± 0.63^{ab}	5.95 ± 0.34^a	5.47 ± 0.52^c
Arg (%)	3.84 ± 0.38^a	3.52 ± 0.33^b	3.79 ± 0.43^a	3.53 ± 0.52^b
Gly (%)	1.96 ± 0.22^a	2.05 ± 0.27^a	1.73 ± 0.23^b	1.74 ± 0.20^b
Thr (%)	4.66 ± 0.32^a	4.66 ± 0.43^a	4.60 ± 0.10^{ab}	4.38 ± 0.42^b
Pro (%)	8.53 ± 0.62^b	10.31 ± 1.20^a	7.78 ± 0.32^c	8.17 ± 0.78^{bc}
Ala (%)	3.59 ± 0.11^b	3.15 ± 0.21^d	3.68 ± 0.09^a	3.47 ± 0.15^c
Val (%)	6.05 ± 0.24^b	6.59 ± 0.44^a	6.02 ± 0.17^{bc}	5.89 ± 0.21^c
Met (%)	2.11 ± 0.18^a	1.91 ± 0.23^b	2.13 ± 0.18^a	2.18 ± 0.14^a
Cys (%)	0.65 ± 0.07^d	0.81 ± 0.13^a	0.70 ± 0.08^c	0.74 ± 0.06^b
Ile (%)	4.75 ± 0.23^b	5.10 ± 0.32^a	4.73 ± 0.15^b	4.78 ± 0.19^b
Leu (%)	8.94 ± 0.24^b	10.12 ± 0.55^a	8.80 ± 0.27^{bc}	8.73 ± 0.25^c
Phe (%)	4.42 ± 0.21^{ab}	4.51 ± 0.31^a	4.46 ± 0.20^a	4.35 ± 0.15^b
Lys (%)	7.61 ± 1.69^b	6.58 ± 3.39^b	7.25 ± 1.37^b	10.32 ± 1.57^a
Tyr (%)	2.73 ± 0.94^{bc}	3.45 ± 2.24^b	2.47 ± 0.72^c	4.54 ± 1.43^a

Mean values in a row with different superscript letters are significantly different by post hoc Duncan's test ($P < 0.05$).
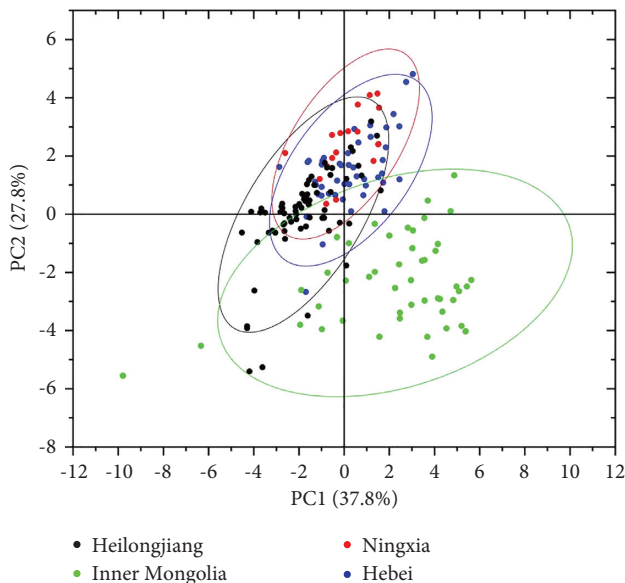

FIGURE 1: PCA score plot for different geographical origins of milk samples.

class response (R^2Y (cum)). The model was considered stable and robust when the values of R^2 and Q^2 were greater than 0.5, and the closer to 1 these values were, the better the model was [26]. The model fitting parameters of R^2X (cum) and R^2Y (cum) were 0.98 and 0.82, respectively, and the prediction parameter of Q^2 (cum) was 0.75. In addition, the lowest root mean square error of cross-validation (RMSECV) for the proposed model was 0.18. Note that the closer to 0 the value is, the better the model is. The above results indicated that the model had a good fitness and a strong ability of prediction. In the score plot (Figure 2(a)), it seems that the data point distribution is relatively close between samples from Hebei and Ningxia. However, these two groups are clearly separated from each other in the corresponding 3D model plot (Appendix 2). The data points of the Inner Mongolia samples were separated

very successfully from those of the other three regions, probably because their large differences in latitude and longitude led to differences in amino acid profiles.

The permutation test ($n = 200$) was performed to assess whether the OPLS-DA model overfitted the data or not (Figure 2(b)). The results showed that the intercept value of Q^2Y was below 0, and the values of R^2 and Q^2 on the left were all lower than the original points on the right, which indicated that the model was valid and did not exhibit overfitting [27]. Moreover, the statistical significance of the OPLS-DA model was also validated by $P_{\text{CV-ANOVA}}$ value, which was 0. The result of the ROC curve can also represent the ability of the model to classify samples correctly. Of those, an ACU value of 1 in Figure 2(c) revealed an excellent performance of this model.

The analysis of variable importance in projection (VIP) was performed to evaluate the contribution of independent variables to the model classification. The potential key markers for differentiation between classes were determined according to the criteria of both the variable importance in projection (VIP) value ≥1 and $P < 0.05$ [28]. As can be seen from Figure 2(d) and Appendix 3, Asp, Glu, Leu, Cys, Ala, Pro, and Val gave a major contribution to the model classification and were proposed as potential markers between four different geographical origins of milk samples.

The amino acids required by dairy cows include essential and nonessential amino acids. Essential amino acids are those that cannot be synthesized by the cow herself and need to be absorbed directly from feed or metabolites by the microbiome of the rumen, including Arg, His, Ile, Leu, Lys, Met, Phe, Thr, and Val [29, 30]. Nonessential amino acids are synthesized through the cow's metabolism using feed and amino acid profiles are controlled by genes. The amino acids in milk are fractionated from the amino acids in the blood through mammary metabolism. Therefore, the reasons for amino acid differences between cow milk origins may include: metabolism differences of essential amino acids between cow milk origins are mainly influenced by feeding
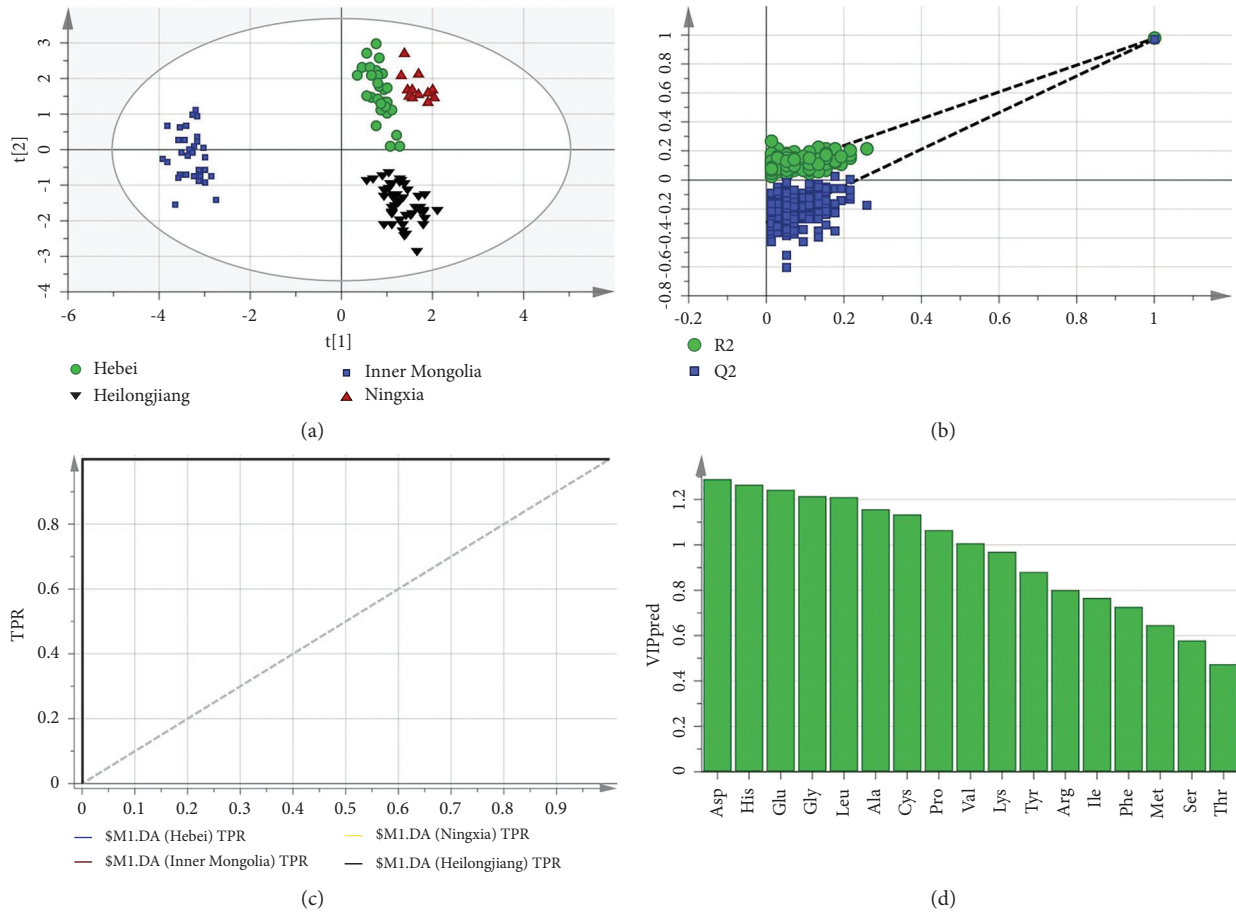
(a)

(b)

(c)

(d)

Figure 2: OPLS-DA analysis based on amino acid ratios. (a) Score scatter plot of the OPLS-DA model for different geographical origins of milk samples. (b) Result of permutation test for OPLS-DA model (intercepts: $R^2 = (0.0, 0.05)$; $Q^2 = (0.0, -0.29)$). (c) ROC curves and (d) VIP values of OPLS-DA.

differences, as different topography and environment result in different feed population. Metabolism differences of nonessential amino acid between cow milk origins are mainly influenced by genetic differences because different regions may raise different breeds of cows and different environments can also cause genetic mutations for cows on adapting to the environment. In addition, the microbiome of the rumen may also vary depending on the environment, and it has been reported that some amino acids in cows are derived from these microbial metabolites [29]. In conclusion, amino acid differences between cow milk origins are influenced by a combination of many factors that are representative of the origin, including the regional climate conditions (rainfall, temperature, possibility to graze) and the breeding of different breeds.

3.4. Validation of the OPLS-DA Model.

The predicting model of OPLS-DA established was validated using the validation data set. The OPLS-DA model is a binary classification method that can only assign imaginary value of 1 and 0. For example, if the imaginary values of 1 (deviation <0.5) is assigned to the class predefined as Hebei, the value of 0 (deviation <0.5) will be assigned to the other three classes

(Inner Mongolia, Ningxia, and Heilongjiang). The predicted values of classification were separately calculated by the OPLS-DA model for each of four classes, and the results of four binary classification predictions are shown in Figure 3. The classification accuracy of the predicting model was 100% for the four classes (Table 2). All the samples were correctly classified, which suggested the powerful predicting ability of the model.

3.5. Practical Application of the OPLS-DA Model.

Chemometric analysis, particularly using the partial least squares-related methods, can intuitively visualize the classification results, and is widely used in the field of geographical origin traceability. Xie et al. [13] established an OPLS-DA model for milk traceability in a small-scale region. Chen et al. [31] used PCA and LDA to trace the geographical origin of *Thelephora ganbajun*. However, these studies did not perform the practical application to the real samples. It is crucial to provide a practical process for identifying the geographical origin of commercial milk using the proposed model. The OPLS-DA model established has been applied to the geographical origin identification of commercial milk in this study. Firstly, amino acid ratios in the unknown milk
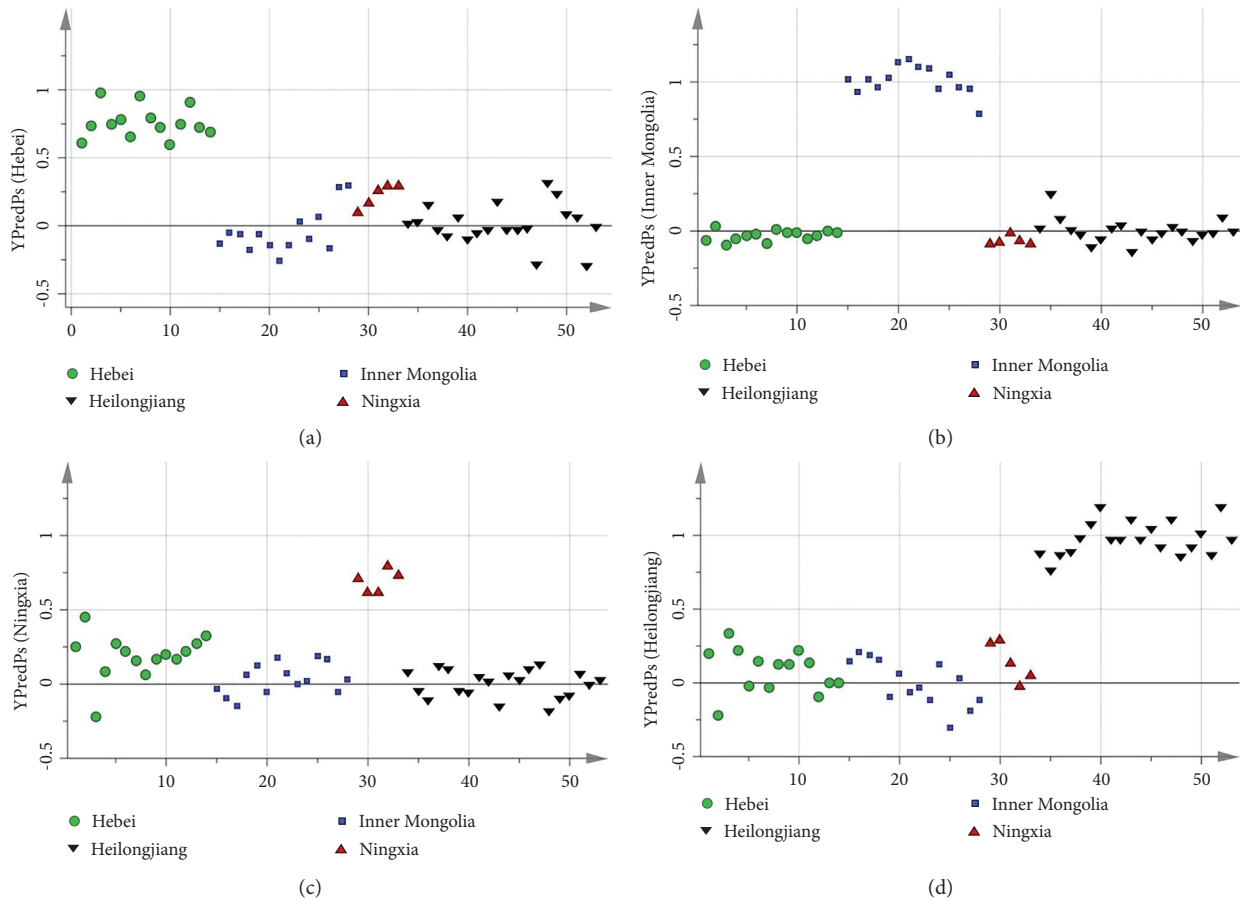
(a)

(b)

(c)

(d)

Figure 3: The results of the four binary classification predictions. The values of 1 are assigned to (a) Hebei, (b) Inner Mongolia, (c) Ningxia, and (d) Heilongjiang, respectively.

Table 2: Discrimination accuracy (%) of the model and validation.

Class	Model		Validation	
	Sample size	Correct (%)	Sample size	Correct (%)
Hebei	33	100	14	100
Inner Mongolia	33	100	14	100
Ningxia	12	100	5	100
Heilongjiang	53	100	14	100

Table 3: The predicted values of the hold out samples in the 4 binary categories.

Samples	YPred (Hebei)	YPred (Inner Mongolia)	YPred (Ningxia)	YPred (Heilongjiang)
A	0.29	0.91	−0.09	−0.11
B	0.04	−0.06	−0.00	1.03
C	0.26	−0.03	0.59	0.18

samples were measured. Subsequently, these data were imported into the developed model under the "specify toolbar" of SIMCA. The geographical origin class of the sample was judged according to the predicted value. If the predicted value of an unknown sample locates at the range from 0.5 to 1.5 in one of the four binary classifications, it will be identified as the corresponding class of geographical origin with an assigned value of 1. According to the above procedure, the predicted values of three samples were obtained, which are detailed in Table 3. The predicted results show that the samples of A, B, and C are from Inner Mongolia, Heilongjiang, and Ningxia, respectively, which is consistent with the labeled information of geographical origin.

4. Conclusions

This study provided a method to identify the geographical origin of milk based on its amino acid profile. The significant differences of amino acid profiles of milk samples from four regions in China (Hebei, Ningxia, Heilongjiang, and Inner Mongolia) were analyzed by ANOVA and PCA. The results suggested that the amino acid profiles had the potential for the identification of geographical origin. The predictive model for the geographical origin of milk samples was established by OPLS-DA with the correct classification accuracy of 100%. The excellent predictive ability of the model was validated using the validation data set. The VIP analysis showed that the amino acids of Asp, Glu, Leu, Cys, Ala, Pro, and Val gave a major contribution to the model classification. This method is a reliable strategy to identify the geographical origin of milk for protecting consumers against mislabeling fraud.

Data Availability

The data used to support the study are included within the article. Raw data can be acquired from the corresponding author upon reasonable request.

Conflicts of Interest

All authors declare that there are no conflicts of interest.

Acknowledgments

Supplementary Materials

Appendix 1. Sample information. Appendix 2. 3D plot of the OPLS-DA model for different origins of milk samples. Appendix 3. Amino acids that gave a major contribution to the model classification. (*Supplementary Materials*)

References

[1] G. Downey, *Advances in Food Authenticity Testing*, Woodhead Publishing, Cambridge, UK, 2016.

[2] B. Peres, N. Barlet, G. Loiseau, and D. Montet, "Review of the current methods of analytical traceability allowing determination of the origin of foodstuffs," *Food Control*, vol. 18, no. 3, pp. 228–235, 2007.

[3] L. M. Reid, C. P. O'Donnell, and G. Downey, "Recent technological advances for the determination of food authenticity," *Trends in Food Science & Technology*, vol. 17, no. 7, pp. 344–353, 2006.

[4] F. Camin, M. Boner, L. Bontempo et al., "Stable isotope techniques for verifying the declared geographical origin of food in legal cases," *Trends in Food Science & Technology*, vol. 61, pp. 176–187, 2017.

[5] M. Kamal and R. Karoui, "Analytical methods coupled with chemometric tools for determining the authenticity and detecting the adulteration of dairy products: a review," *Trends in Food Science & Technology*, vol. 46, no. 1, pp. 27–48, 2015.

[6] S. A. Drivelos and C. A. Georgiou, "Multi-element and multi-isotope-ratio analysis to determine the geographical origin of foods in the European Union," *TrAC Trends in Analytical Chemistry*, vol. 40, pp. 38–51, 2012.

[7] Y. Zhao, B. Zhang, G. Chen, A. Chen, S. Yang, and Z. Ye, "Recent developments in application of stable isotope analysis on agro-product authenticity and traceability," *Food Chemistry*, vol. 145, pp. 300–305, 2014.

[8] R. G. Crittenden, A. S. Andrew, M. LeFournour, M. D. Young, H. Middleton, and R. Stockmann, "Determining the geographic origin of milk in Australasia using multi-element stable isotope ratio analysis," *International Dairy Journal*, vol. 17, no. 5, pp. 421–428, 2007.

[9] M. Scampicchio, T. Mimmo, C. Capici et al., "Identification of milk origin and process-induced changes in milk by stable isotope ratio mass spectrometry," *Journal of Agricultural and Food Chemistry*, vol. 60, no. 45, pp. 11268–11273, 2012.

[10] D. Luo, H. Dong, H. Luo, Y. Xian, X. Guo, and Y. Wu, "Multi-element (C, N, H, O) stable isotope ratio analysis for determining the geographical origin of pure milk from different regions," *Food Analytical Methods*, vol. 9, no. 2, pp. 437–442, 2016.

[11] S. Zhao, Y. Zhao, K. M. Rogers, G. Chen, A. Chen, and S. Yang, "Application of multi-element (C, N, H, O) stable isotope ratio analysis for the traceability of milk samples from China," *Food Chemistry*, vol. 310, 2020.

[12] D. Sacco, M. A. Brescia, A. Sgaramella et al., "Discrimination between southern Italy and foreign milk samples using spectroscopic and analytical data," *Food Chemistry*, vol. 114, no. 4, pp. 1559–1563, 2009.

[13] L. Xie, S. Zhao, K. M. Rogers et al., "A case of milk traceability in small-scale districts-Inner Mongolia of China by nutritional and geographical parameters," *Food Chemistry*, vol. 316, 2020.

[14] F. Camin, R. Wehrens, D. Bertoldi et al., "H, C, N and S stable isotopes and mineral profiles to objectively guarantee the authenticity of grated hard cheeses," *Analytica Chimica Acta*, vol. 711, pp. 54–59, 2012.

[15] A. V. Silva, J. F. Hélie, F. d. A. Caxito, H. Monardes, A. F. Mustafa, and R. Stevenson, "Multi-stable isotope analysis as a tool for assessing the geographic provenance of dairy products: a case study using buffalo's milk and cheese samples from the Amazon basin, Brazil," *International Dairy Journal*, vol. 35, no. 2, pp. 107–110, 2014.

[16] C. Capici, T. Mimmo, L. Kerschbaumer, S. Cesco, and M. Scampicchio, "Determination of cheese authenticity by carbon and nitrogen isotope analysis: stelvio cheese as a case study," *Food Analytical Methods*, vol. 8, no. 8, pp. 2157–2162, 2015.

[17] A. Rossmann, G. Haberhauer, S. Hölzl, P. Horn, F. Pichlmayer, and S. Voerkelius, "The potential of multielement stable isotope analysis for regional origin assignment of butter," *European Food Research and Technology*, vol. 211, no. 1, pp. 32–40, 2000.

[18] L. Liang, J. Chen, J. Hou, and L. Zhang, "Traces of infant formula milk powders by elemental analysis-stable isotope ratio mass spectrometry," *Journal of Chinese Mass Spectrometry Society*, vol. 31, no. 1, pp. 66–71, 2015.

[19] R. Karoui and J. Debaerdemaeker, "A review of the analytical methods coupled with chemometric tools for the determination of the quality and identity of dairy products," *Food Chemistry*, vol. 102, no. 3, pp. 621–640, 2007.

[20] J. Riedl, S. Esslinger, and C. Fauhl-Hassek, "Review of validation and reporting of non-targeted fingerprinting

approaches for food authentication," *Analytica Chimica Acta*, vol. 885, pp. 17–32, 2015.

[21] P. Galvin-King, S. A. Haughey, and C. T. Elliott, "Garlic adulteration detection using NIR and FTIR spectroscopy and chemometrics," *Journal of Food Composition and Analysis*, vol. 96, 2021.

[22] B. D. Nkosi, J. S. M. Phenya, I. M. M. Malebana, M. C. Muya, and M. D. Motiang, "Nutrient evaluation and ruminal degradation of dry matter and protein from amarula (Sclerocarya birrea), macadamia (integrifolia) and baobab (Adansonia digitata L.) oilcakes as dietary supplements for ruminants," *Tropical Animal Health and Production*, vol. 51, no. 7, pp. 1981–1988, 2019.

[23] B. Worley and R. Powers, "PCA as a practical indicator of OPLS-DA model reliability," *Current Metabolomics*, vol. 4, no. 2, pp. 97–103, 2016.

[24] H. Sadeghi-Bazargani, S. I. Bangdiwala, K. Mohammad, H. Maghsoudi, and R. Mohammadi, "Compared application of the new OPLS-DA statistical model versus partial least squares regression to manage large numbers of variables in an injury case-control study," *Scientific Research and Essays*, vol. 6, no. 20, pp. 4369–4377, 2011.

[25] T. Shi, M. Zhu, Y. Chen et al., "1H NMR combined with chemometrics for the rapid detection of adulteration in camellia oils," *Food Chemistry*, vol. 242, pp. 308–315, 2018.

[26] S. Mi, W. Yu, J. Li, M. Liu, Y. Sang, and X. Wang, "Characterization and discrimination of chilli peppers based on multi-element and non-targeted metabolomics analysis," *Lwt-Food Science and Technology*, vol. 131, 2020.

[27] F. He, J. Duan, J. Zhao et al., "Different distillation stages Baijiu classification by temperature-programmed headspace-gas chromatography-ion mobility spectrometry and gas chromatography-olfactometry-mass spectrometry combined with chemometric strategies," *Food Chemistry*, vol. 365, Article ID 130430, 2021.

[28] J.-L. Wu, C.-X. Zhou, P.-J. Wu et al., "Brain metabolomic profiling of eastern honey bee (*Apis cerana*) infested with the mite Varroa destructor," *PLoS One*, vol. 12, no. 4, Article ID e0175573, 2017.

[29] S. I. Arriola Apelo, A. L. Bell, K. Estes, J. Ropelewski, M. J. de Veth, and M. D. Hanigan, "Effects of reduced dietary protein and supplemental rumen-protected essential amino acids on the nitrogen efficiency of dairy cows," *Journal of Dairy Science*, vol. 97, no. 9, pp. 5688–5699, 2014.

[30] C. G. Schwab, L. D. Satter, and A. B. Clay, "Response of lactating dairy cows to abomasal infusion of amino acids," *Journal of Dairy Science*, vol. 59, no. 7, pp. 1254–1270, 1976.

[31] L.-p. Chen, H.-y. Zhu, Y.-f. Li et al., "Combining multielement analysis and chemometrics to trace the geographical origin of *Thelephora ganbajun*," *Journal of Food Composition and Analysis*, vol. 96, 2021.