71st ICoMST
GIRONA - CATALONIA - SPAIN
3rd-8th August - 2025

GC-IMS COMBINES MACHINE LEARNING ALGORITHMS TO EXPLORE KEY VOLATILE COMPOUNDS IN LAMB FOR IDENTIFICATION OF GEOGRAPHICAL INDICATION PRODUCTS

Qi Yang, Chongxin Liu, Muxuan Xu, Li Chen, Xiaochun Zheng, Shaobo Li*, Dequan Zhang*

Institute of Food Science and Technology, Chinese Academy of Agriculture Sciences, Key Laboratory of Agro-Products Quality and Safety Control in Storage and Transport Process, Ministry of Agriculture and Rural Affairs, Beijing 100193, China.

*Corresponding author email: lishaobomeat@126.com, dequan_zhang0118@126.com

I. INTRODUCTION

In recent years, the global meat market has shifted its focus from price competition to quality traceability and brand development. Consumer demand for Geographical Indication (GI) products has grown significantly[1]. The geographical origin, serving as the "geographical fingerprint" of meat quality, imparts unique flavor characteristics to meat through environmental factors such as climate conditions and soil properties, thereby significantly influencing consumer acceptance. However, the relationship between the selection of flavor markers and geographical origin determination remains insufficiently explored. This study employed gas chromatography-ion mobility spectrometry (GC-IMS) to analyze the flavor components of Tan lamb from three different regions and utilized the Recursive Feature Elimination (RFE) method combined with a random forest algorithm to identify the characteristic flavors of Yanchi Tan lamb, a GI-certified product. Furthermore, by comparing the predictive performance of various machine learning models, this study achieved precise identification of Yanchi Tan lamb. The findings provide a solid theoretical foundation for the selection of meat flavor markers and the scientific determination of GI products.

II. MATERIALS AND METHODS

Lamb samples were collected from the Ningxia Hui Autonomous Region, Gansu Province, and the Inner Mongolia Autonomous Region of China. Six individual lambs were selected from each region, with samples taken from the *M. longissimus thoracis et lumborum* and knuckle meat. All samples were obtained from lambs (6-month-old male lambs) raised and slaughtered under the same conditions within each region and stored in a cold room at 4° C for 24 hours post-slaughter. A GC-IMS (FlavourSpec®, G.A.S., Dortmund, Germany) was used as described by Chen et al. (2024) with slight modifications.Flavor marker selection was conducted using the random forest model in MetaboAnalyst 6.0. To minimize randomness, each model was constructed 10 times, with the average out-of-bag (OOB) error rate used to evaluate model performance. Factors were ranked based on their average variable importance, and the lowest 20% of factors were iteratively removed during the selection process. The discrimination model for Yanchi lamb was constructed and analyzed using MATLAB.

III. RESULTS AND DISCUSSION

A total of 33 volatile compounds were detected in Tan lamb from the three regions (Figure A), the GC-IMS fingerprints showed that the content of volatile compounds in Tan lamb from different regions varied significantly.

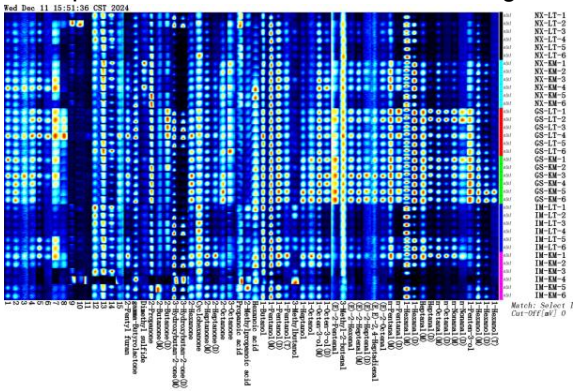


Figure 1. GC-IMS fingerprints of Tan lamb from different regions

Although fingerprint spectra can visually demonstrate the flavor differences of lamb from different regions, the key volatile compounds of Yanchi Tan lamb require further investigation. Principal Component Analysis was conducted

for a preliminary examination of sample similarities. Samples from different regions exhibited clear regional clustering (Figure 2A), indicating the presence of potential key volatile compounds that can effectively distinguish lamb origins. Notably, the IM-KM-4 sample displayed outlier behavior and was excluded from subsequent analyses.
Key volatile compounds distinguishing Yanchi Tan lamb were identified using the RFE method combined with the random forest algorithm. As the number of variables decreased, the classification error of the model showed a declining trend (Figure 2B). When the number of variables was reduced to five, the classification error of the random forest model dropped to zero (Figure 2C). The five selected volatile compounds were identified as key markers for distinguishing Yanchi Tan lamb. Among them, Heptanal (M), 1-Hexanal (D), and n-Nonanal (M) were successfully identified, in addition to two unidentified compounds (Figure 2D).

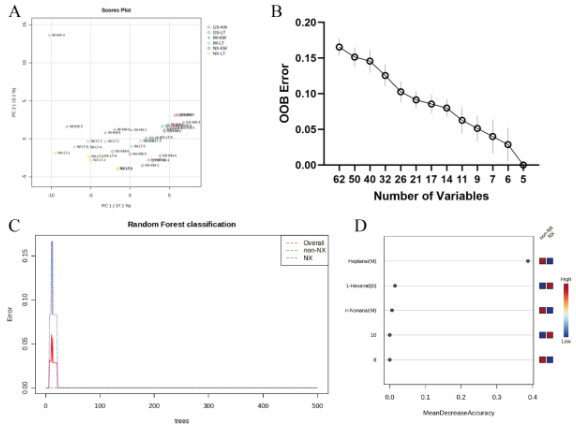


Figure 2. Screening of key volatile compounds in Yanchi Tan lamb

The key volatile compounds screened were further used to construct the discrimination models of Yanchi Tan lamb, and the prediction accuracies of the five discrimination models ranged from 82.14% to 100% in the training set and from 71.43% to 100% in the test set (Table 1), which demonstrated better discrimination performance. The feasibility of combining volatile compounds with machine learning models for GI product identification and random forest RFE for feature marker screening were confirmed.

Table 1 – Performance comparison of machine learning classification models.

Models	BP	GA-BP	PSO-BP	ELM	SVM
Training set discrimination accuracy, %	100.00	100.00	96.43	100.00	82.14
Test set discrimination accuracy, %	85.71	100.00	100.00	71.43	85.71

IV. CONCLUSION

The results of the research showed that geographical origin serves as a critical determinant in shaping the distinctive flavor profiles of lamb. Through the application of RFE, we effectively identified key characteristic markers and subsequently developed machine learning discrimination model that achieved accurate authentication of Yanchi Tan lamb. This research not only provides a scientific basis for the quality certification of Yanchi Tan lamb, but also develops an innovative methodological framework based on the screening of characteristic markers for food traceability technology, but the intrinsic connection between origin and the formation of characteristic flavor still needs to be further explored.

ACKNOWLEDGEMENTS

This work was supported by the National Key R&D Program of China (2022YFD2100500) and National Key R&D Program of Ningxia Hui Autonomous Region, China (2024BBF02022).

REFERENCES

1. Fu, R.; Li, C.; Wang, L.; Gao, Z. (2024) Chinese consumer preference for beef with geographical indications and other attributes. Meat Science, 212, 109475.
2. Chen, J.; Wang, W.; Jin, J.; Li, H.; Chen, F.; Fei, Y.; Wang, Y. (2024) Characterization of the flavor profile and dynamic changes in Chinese traditional fish sauce (Yu-lu) based on electronic nose, SPME-GC-MS and HS-GC-IMS. Food Research International, 192, 114772.