HNUE JOURNAL OF SCIENCE
Natural Sciences 2024, Volume 69, Issue 3, pp. 154-162
This paper is available online at http://hnuejs.edu.vn/ns
DOI: 10.18173/2354-1059.2024-0045

APPLICATION OF SPECTROSCOPIC METHODS FOR THE AUTHENTICATION OF AGRICULTURAL PRODUCTS OF VIETNAM: A STUDY OF ST25 RICE

Nguyen Quynh Hoa[1], Bui Thu Thuy[2] and To Thi Mai Huong[3]
[1]*Department of General Education, University of Science and Technology of Hanoi, Vietnam Academy of Science and Technology, Hanoi city, Vietnam*
[2]*Department of Chemistry, College of Natural Sciences, Hanyang University, Seoul city, South Korea*
[3]*Department of Life Sciences, University of Science and Technology of Hanoi, Vietnam Academy of Science and Technology, Hanoi city, Vietnam*
[*]Corresponding author: Nguyen Quynh Hoa, e-mail: nguyen-quynh.hoa@usth.edu.vn

Received August 13, 2024. Revised October 4, 2024. Accepted October 31, 2024.

Abstract. Food fraud, particularly in staple commodities like rice, poses significant risks to trademark protection as well as economic challenges. In this study, we explored the efficacy of Raman spectroscopy in differentiating authentic ST25 rice—a premium variety of Vietnam from other rice types, amidst concerns of adulteration. By utilizing both backscattering and transmission Raman spectroscopy, a total of 125 rice samples consisting of both commercial rice and Vietnamese landrace rice varieties were analyzed. All samples were categorized into two main groups namely ST25 and Non-ST25 for the construction of the classification model. Through principal component analysis (PCA) and k-nearest neighbors (kNN) classification method, we achieved classification accuracies of 81.58% for backscattering data and up to 97.37% for transmission data at elevated temperatures. Our findings highlight the efficacy of Raman spectroscopy for rice authenticity verification, nevertheless, modification in spectral measurement schemes is necessary to obtain better method reproducibility as well as to maintain high discriminatory accuracy of the classification model.

Keywords*:* ST25 rice, Raman transmission, Raman backscattering, principal component analysis, k-nearest neighbors.

1. Introduction

Food fraud has been a pressing global issue. The term food fraud or food crime refers to the intentional deception of food for profit [1]. Food fraud occurs frequently in commodities consumed in large quantities [2], in which rice is a vulnerable item for adulteration since it is a staple food for over half of the world's population [3]. Common forms of rice fraud include mislabelling using cheaper, lower-quality products instead of the high-quality products as indicated on the label [4], adding artificial flavors to mask the use of tasteless rice, and mixing high quality rice with low quality or genetically modified rice products [5]. These fraudulent practices can threaten public health and negatively impact the nation's economy of food [6].

Therefore, the use of rapid and reliable analytical methods for the verification of product quality and the assessment of food safety is essential [7]. The authentication of rice and rice products has long been conducted using modern analytical techniques combined with multivariate analysis. Besides popular methods such as isotope ratio mass spectrometry, inductively coupled plasma mass spectrometry, DNA analysis, and omics analysis [8]-[9], spectroscopic methods are also widely used for food authenticity owing to various advantages such as non-destructive analysis, rapid and reliable results with high reproducibility, cost, and labor efficiency [10].

Among spectroscopic techniques commonly used for rice authentication such as near-infrared spectroscopy, laser-induced breakdown spectroscopy, and nuclear magnetic resonance, Raman spectroscopy has been developed as a rapid measurement method for food quality assurance and safety monitoring [11]-[13]. Rice is a product with a complex sample matrix, primarily consisting of a mixture of components such as carbohydrates, lipids, proteins, and other biological components. Raman spectra are capable of simultaneous identification of numerous molecules with fingerprint recognition [14], enabling quick component analysis or online monitoring [15]-[16]. Meanwhile, the rapid development of Raman reading devices and smart data processing algorithms has propelled the widespread application of Raman spectroscopy in food monitoring. Factors such as post-harvest processing conditions, geographical origin, or variety may considerably affect the chemical compositions of food, and in turn, this is reflected in different levels of radiation absorption at specific wavelengths. Such information in combination with data analysis allows the differentiation of various food types, thereby expanding the scope of applications and the number of publications in this field. By combining data processing and chemical analysis, Raman spectroscopy provides rapid identification of rice species, varieties, production regions, and market monitoring [17]-[18].

ST25 rice is a premium agricultural product of Vietnam that was developed by Mr. Ho Quang Cua for high yields and resilience against various climatic conditions. This rice variety has won the "World's Best Rice" awards twice in 2019 and 2023, making it a valuable fraudulent target for unscrupulous vendors for economic benefits. Therefore, in this study, we aimed to investigate the potential of Raman spectroscopy in distinguishing authentic ST25 rice from other rice types. This can serve as preliminary research toward developing rapid detection methods for the authenticity of commonly traded rice varieties in the market.

2. Content

2.1. Methods

2.1.1. Sample collection

A total of 125 samples were used in this analysis including 46 commercial ST25 rice samples, 33 commercial non-ST25 rice samples, and 46 rice samples from the Vietnamese rice collection. Commercial ST25 rice samples were purchased in 2022 from different brands with proper labels providing adequate product information such as origin, production date, and expiration date. Non-ST25 commercial rice samples were acquired similarly. Examples of commercial ST25 and non-ST25 rice samples are shown in Figure 1. Samples from the Vietnamese landrace rice varieties were provided by the Plant Resources Center in An Khanh, Hanoi, and were cultivated at the Biological and Experimental Station of the Vietnam Academy of Science and Technology in 2023. The commercial ST25 rice samples were categorized into the ST25 group, while the other rice samples were categorized into the Non-ST25 group for the classification model.

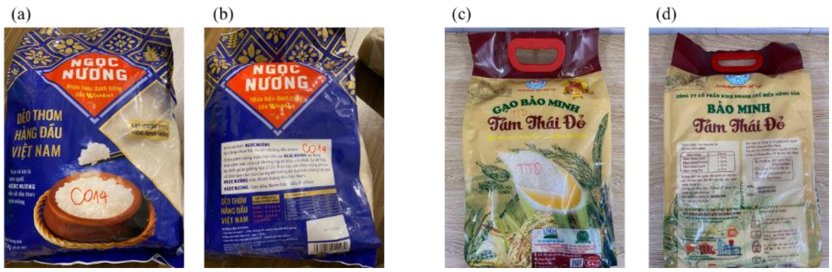

Figure 1. Examples of a commercial ST25 rice sample (a: front view and b: back view of a package) and a commercial non-ST25 rice sample (c: front view and d: back view of a package)

2.1.2. Raman measurement

Raman spectral signals were collected using two different methods: transmission and backscattering. For transmission Raman measurement, a laser source (785 nm, Invictus, Kaiser Optical Inc., USA) with a laser illumination diameter of 1 mm was irradiated on the side of the vial (9.8 mm diameter x 20 mm depth) containing rice grains, and sample-interacting photons were collected at the opposite side using a PhAT probe connected to a Raman spectrometer (RamanRxn1 unit, Kaiser Optical Inc.). Meanwhile, for backscattering Raman measurement, each sample of rice grains was placed in a rotating holder with dimensions of 35mm in diameter and 10 mm in depth, and the laser with a 6-mm laser illumination diameter was used to irradiate the sample. The scattered radiation was collected by a detector placed on the same side as the laser source [19]. The backscattering Raman signal was measured at 20 °C, while the transmission Raman signal was conducted at four different temperatures: 20, 30, 40, and 50 °C to assess the feasibility of different measurement methods. Triplicate Raman spectra of each sample were collected and the average spectrum of these triplicate measurements was employed for subsequent analysis.

2.1.3. Data processing

The baseline of each raw spectrum was corrected by asymmetric least squares smoothing (AsLS), and then the AsLS-processed spectrum was normalized by dividing the intensities of each peak by the peak area calculated over the 1700-800 cm^{-1} range. For 125 rice samples belonging to two groups, ST25 and Non-ST25, principal component analysis (PCA) was used to evaluate the overall separation between the groups. Subsequently, the k-nearest neighbors (kNN) method was used to differentiate between the two groups. Two-level cross-validation (CV) [18] was used to evaluate the classification accuracy of the kNN model with three-fold cross-validation and 100 repetitions. Baseline correction, spectral normalization, PCA, kNN analyses, and graphical representations were conducted using MATLAB R2020b.

2.2. Results and discussion

2.2.1. Raman spectra data

The backscattering Raman spectra of the rice samples displayed similar spectral patterns (Figure 2). The main differences appeared in the wavelength ranging from 1550 to 1700 cm^{-1}, which may be related to the protein content of the samples [20]. The transmission Raman signal of Vietnamese rice collection samples at different temperatures showed significant variation across the wavelength range, whereas the transmission signal of the commercial rice samples was more stable (Figure 3). This could be due to the heterogeneous inner structure of the rice grains [21]. The rice grains consist of three main parts: the bran layer, the endosperm, and the embryo, each of which is composed of different chemical components. Consequently, the Raman signal at each position varies accordingly. The collection of transmission Raman data from the rice samples was also affected by these components, where the outermost layer of samples from the Vietnamese rice collection often still has bran because they have not been milled thoroughly, while the outermost layer of other commercial rice samples has been milled. In addition, the transmission Raman measurement scheme may provide less variation in spectral intensities of the commercial rice samples [20]. However, examination of a small amount of rice grains may induce spectral variation between the Vietnamese landrace rice samples and the commercial samples. A larger sample coverage area or the increment in measurement from different random positioning of samples may help reduce such among-sample variation as well as increase the reproducibility of the measurement method [22].

In contrast, the backscattering Raman signal of the samples from the Vietnamese rice collection was relatively stable and had a pattern similar to that of the commercial ST25 rice samples and the commercial non-ST25 rice samples. The signal intensities of the commercial samples, regardless of ST25 or non-ST25 rice were relatively similar, with only a few different regions between the two groups, such as the wavelength range from 1480 to 1700 cm^{-1}, corresponding to the protein content, or the wavelength range from 1300 to 1450 cm^{-1}, corresponding to amylose/amylopectin content [20].

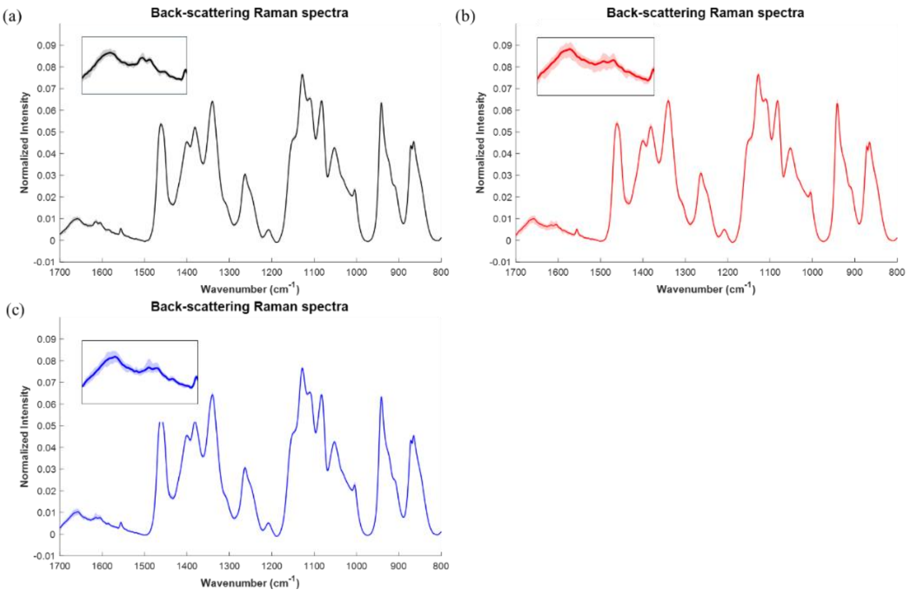

Figure 2. Backscattering Raman spectral data presented as mean (solid line) ± standard deviation (shaded area) of (a) rice samples from the Vietnamese landrace rice varieties, (b) commercial ST25 rice samples, and (c) commercial non-ST25 rice samples. Magnified regions from 1550 to 1700 cm^{-1} are shown in the insets

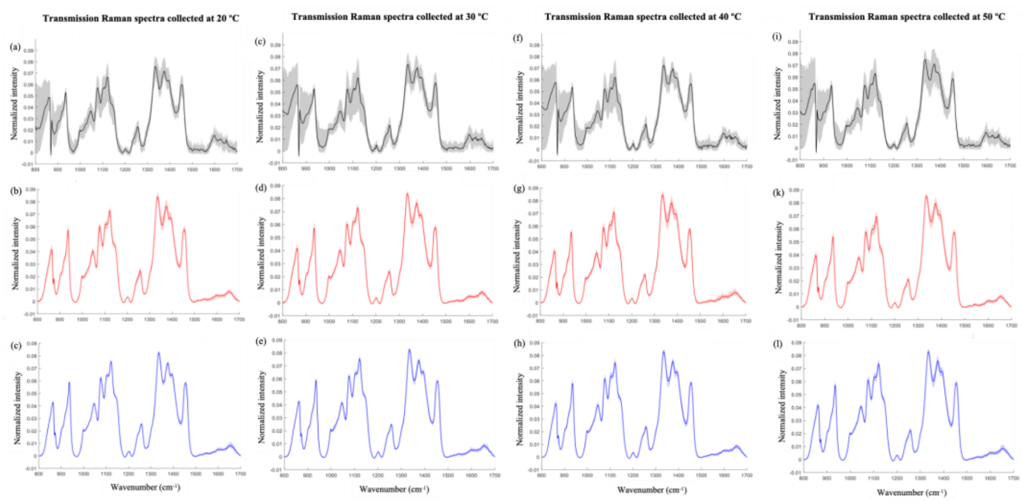

Figure 3. Transmission Raman spectral data measured at 20 to 50 ºC of rice samples from the Vietnamese rice collection (a, c, f, i), commercial ST25 rice samples (b, d, g, k), and commercial non-ST25 rice samples (c, e, h, l)

2.2.2. Discrimination between ST25 and Non-ST25

The PCA results indicate that the backscattering Raman signals of the rice samples were quite similar (Figure 4). Despite some differences in the spectral signals at specific wavelengths, the samples primarily overlapped in signal intensity. The first two principal components of the PCA model explained 51.11% of the total variance, meanwhile, the supervised classification model kNN applied across the entire range of wavelengths of the backscattering Raman dataset achieved an accuracy of 81.58%. This could be due to the small amount of rice grains examined in this study which may not fully reflect all the characteristics of the rice samples although the backscattering Raman signals of the rice samples were relatively stable. Further analysis is needed to examine backscattering variations from a larger sample coverage area, as well as to increase the number of measurements at different random positions of rice grains.

Regarding transmission Raman spectra, PCA results at all four different temperature levels showed distinct clustering between the two rice groups, ST25 and Non-ST25 (Figure 5). The first two principal components explained 87.63%, 93.27%, 93.79%, and 92.61% of the total variance of the dataset at temperatures ranging from 20 to 50°C, respectively. Overall, the samples from the two groups exhibited clear separation in signals with no overlap in any data points. The total variance explained by the first two principal components was highest at 50 °C and lowest at 20 and 30 °C. The supervised clustering model kNN applied across the entire wavelengths for the transmission of Raman data at temperatures ranging from 20 °C to 50 °C achieved accuracies of 89.47%, 89.47%, 92.10%, and 97.37%, respectively.

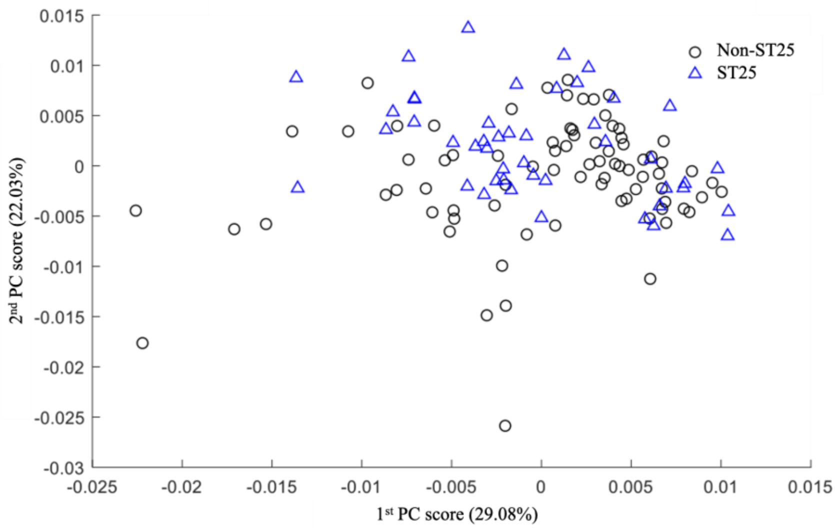

Figure 4. PCA results for the discrimination between ST25 and Non-ST25 groups using the backscattering mode. The first and second principal components explained 29.08% and 22.03% of the total variance, respectively

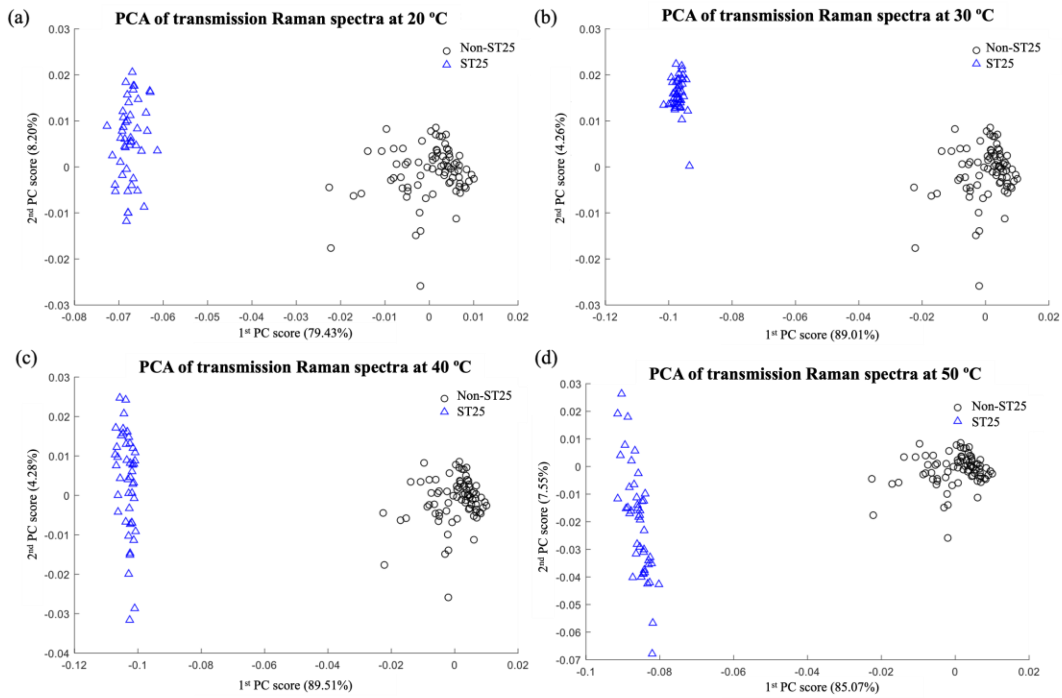

Figure 5. PCA results for distinguishing between the two rice groups: ST25 and Non-ST25 using transmission Raman data at different temperature levels

Even though high discrimination accuracy was obtained at higher temperatures, Raman spectra of high variation could lead to reduced reproducibility, which eventually severely affects the discrimination model. Further research examining a larger sample coverage area and a higher number of spectra measurements should be conducted to better develop a measurement scheme that could combine the advantages observed in both backscattering and transmission schemes.

3. Conclusions

Overall, this study has presented the feasibility of implementing Raman spectroscopy in authenticating rice varieties, with ST25 rice as a study species. By successfully distinguishing ST25 rice from non-ST25 counterparts with high accuracy, regardless of commercial products or samples collected from the Vietnamese landrace rice varieties, this research contributes to the development of rapid detection methods that can safeguard consumer health and maintain market integrity. Further analysis should be conducted for the optimization of the spectral measurement scheme for better reproducibility and high accuracy of the discrimination method. Given the complexities involved in rice adulteration and the importance of reliable food safety measures, future investigations should focus on refining the measurement techniques to capitalize on the advantages of both backscattering and transmission spectroscopy. As food fraud continues to be a pressing issue globally, implementing robust analytical methods like Raman spectroscopy will be essential in protecting consumers and ensuring the authenticity of food products.

Acknowledgments. Nguyen Quynh Hoa was funded by the Postdoctoral Scholarship Programme of Vingroup Innovation Foundation (VINIF), code VINIF.2023.STS.41.

REFERENCES

[1] Spink J & Moyer DC, (2011). Defining the public health threat of food fraud. *Journal of Food Science*, 76(9), R157-R163.

[2] Kowalska A, Soon JM & Manning L, (2018). A study on adulteration in cereals and bakery products from Poland including a review of definitions. *Food Control*, 92, 348-356.

[3] Nguyen VN & Ferrero A, (2006). Meeting the challenges of global rice production. *Paddy and Water Environment*, 4, 1-9.

[4] Nader WF, Brendel T & Schubbert R, (2016). Advances in DNA fingerprinting for food authenticity testing: *Advances in food authenticity testing*, p. 7-33, Woodhead Publishing.

[5] Food Standard Agency, (2017). *Final report: Application of multispectral imaging (MSI) to food and feed sampling and analysis*. https://www.food.gov.uk/sites/default/files/media/document/fs301017finrep.pdf. (Accessed on October 7, 2024).

[6] Food and Agriculture Organization of the United Nations, (2017). *Rice Market Monitor (RMM).* http://www.fao.org/economic/est/publications/rice-publications/rice-mar ket-monitor-rmm/en/. (Accessed on October 7, 2024).

[7] Lohumi S, Lee S, Lee H & Cho BK, (2015). A review of vibrational spectroscopic techniques for the detection of food authenticity and adulteration. *Trends in Food Science & Technology*, 46(1), 85-98.

[8] Teye E, Amuah CL, McGrath T & Elliott C, (2019). Innovative and rapid analysis for rice authenticity using hand-held NIR spectrometry and chemometrics. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 217, 147-154.

[9] Xu Y, Zhong P, Jiang A, Shen X, Li X, Xu Z, Shen Y, Sun Y & Lei H, (2020). Raman spectroscopy coupled with chemometrics for food authentication: A review. *TrAC Trends in Analytical Chemistry*, 131, 116017.

[10] Wadood SA, Chunlin L, Nie J, Rogers KM, Mei H, Zhang Y, Shah I, Qamar A & Yuan Y, (2024). Stable isotopic fingerprinting of authentic basmati rice from Pakistan. *Food Control*, 157, 110166.

[11] Lim WY, Goh CH, Thevarajah TM, Goh BT & Khor SM, (2020). Using SERS-based microfluidic paper-based device (μPAD) for calibration-free quantitative measurement of AMI cardiac biomarkers. *Biosensors and Bioelectronics*, 147, 111792.

[12] McVey C, Elliott CT, Cannavan A, Kelly SD, Petchkongkaew A & Haughey SA, (2021). Portable spectroscopy for high throughput food authenticity screening: Advancements in technology and integration into digital traceability systems. *Trends in Food Science & Technology*, 118, 777-790.

[13] Wang K, Li Z, Li J & Lin H, (2021). Raman spectroscopic techniques for nondestructive analysis of agri-foods: A state-of-the-art review. *Trends in Food Science & Technology*, 118, 490-504.

[14] Zhang Y, Ren L, Wang Q, Wen Z, Liu C & Ding Y, (2022). Raman spectroscopy: A potential diagnostic tool for oral diseases. *Frontiers in Cellular and Infection Microbiology*, 12, 775236.

[15] Hirsch E, Pataki H, Domján J, Farkas A, Vass P, Fehér C, Barta Z, Nagy ZK, Marosi GJ & Csontos I, (2019). Inline noninvasive Raman monitoring and feedback control of glucose concentration during ethanol fermentation. *Biotechnology Progress*, 35(5), e2848.

[16] Liu Y, Ren X, Yu H, Cheng Y, Guo Y, Yao W & Xie Y, (2020). Non-destructive and online egg freshness assessment from the eggshell based on Raman spectroscopy. *Food Control*, 118, 107426.

[17] Feng X, Zhang Q, Cong P & Zhu Z, (2013). Preliminary study on classification of rice and detection of paraffin in the adulterated samples by Raman spectroscopy combined with multivariate analysis. *Talanta*, 115, 548-555.

[18] Guo Y, Cai W, Tu K, Tu S, Wang S, Zhu X & Zhang W, (2013). Infrared and Raman spectroscopic characterization of structural changes in albumin, globulin, glutelin, and prolamin during rice aging. *Journal of Agricultural and Food Chemistry*, 61(1), 185-192.

[19] Kim M, Chung H, Woo Y & Kemper M, (2006). New reliable Raman collection system using the wide area illumination (WAI) scheme combined with the synchronous intensity correction standard for the analysis of pharmaceutical tablets. *Analytica Chimica Acta*, 579(2), 209-216.

[20] Bui TT, Jeong S, Jeong H, Le GT, Nguyen HQ & Chung H, (2023). Authentication of ST25 rice using temperature-perturbed Raman measurement with variable selection by Incremental Association Markov Blanket. *Food Chemistry*, 429, 136985.

[21] Hwang J, Kang S, Lee K & Chung H, (2012). Enhanced Raman spectroscopic discrimination of the geographical origins of rice samples via transmission spectral collection through packed grains. *Talanta*, 101, 488-494.

[22] Jeong H, Cho J, Han J, Yoon YS, Kim HG, Kim J & Chung H, (2025). Large-area coverage–transmission near-infrared measurement of dried laver to determine crude protein content. *Food Control*, 168, 110934.