Acta Manilana

Chemometric differentiation of aromatic rice cultivars using gas-phase Fourier transform infrared spectroscopy

Michael E. Serafico[1,2] and Fortunato B. Sevilla III[1,3*]

[1]*The Graduate School, University of Santo Tomas, España Blvd., Manila, Philippines 1008*
[2]*Department of Science and Technology-Food and Nutrition Research Institute, DOST Compound, General Santos Avenue, Bicutan, Taguig City, Philippines*
[3]*Research Center for the Natural and Applied Sciences, University of Santo Tomas, España, Manila, Philippines*

Graphical Abstract

Aromatic rice sample　　**Headspace sample**　　**FTIR spectra**　　**PCA biplot**
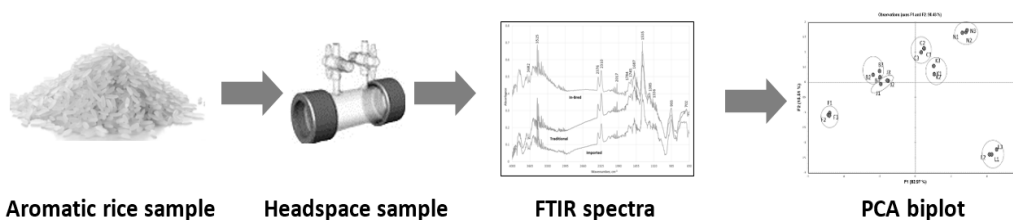
Abstract

Infrared spectroscopy is a standard method in analytical chemistry that transforms the vibration of atoms in a compound into images. In the mid-infrared (MIR) spectral region extending from 4,000-400 cm^{-1}, the chemical, structural, and compositional information of a molecule, regarded as its fingerprint region, in the solid, liquid, and gas phases can be deduced. Fourier transform infrared (FTIR) spectrometers are predominantly used over conventional dispersive instruments because modern computer technology allows real-time, fast Fourier transform algorithms. This study explored the utility of gas-phase Fourier transform infrared spectroscopy for the chemometric differentiation of aromatic rice. The seven cultivars Basmati, *Dinorado*, Jasmine, *Milagrosa*, NSICs Rc148, Rc342, and Rc344 were characterized by different infrared patterns. Agglomerative Hierarchical Clustering (AHC) generated clusters that define the rice's similar characteristics. Correspondingly, Principal Component Analysis (PCA) was able to precisely classify the samples into different varieties (≈99%). The results demonstrated that chemometric analysis of FTIR spectrograms can be a reliable technique and has a high potential to discriminate aromatic rice samples based on their infrared fingerprints. The study provided a possible inexpensive and non-destructive alternative that has not been explored before to assess the authenticity of rice varieties through an existing analytical platform.

Introduction

In the fields of agriculture and food science, the demand for quality assurance and precise differentiation of agricultural products has grown exponentially. Among these products, rice, considered as a staple food for nearly half of the global population, is often a target of adulteration [1]. The market for specialty rice, particularly aromatic cultivars has expanded due to their distinctive flavors, fragrances, and cultural significance. As consumer preferences continue to evolve, the need for rapid and accurate methods to differentiate these rice varieties becomes increasingly crucial.

Traditional methods for distinguishing rice cultivars involve time-consuming and often subjective procedures, such as visual inspection and sensory analysis [2,3,4]. However, advancements in analytical methods have paved the way for more modern techniques. One such technique, gas-phase Fourier Transform Infrared (FTIR) spectroscopy, has emerged as a powerful tool for qualitative and quantitative analysis of organic compounds [5]. FTIR spectroscopy exploits the characteristic vibrational modes of molecular bonds to produce unique spectral fingerprints that can differentiate substances.

Using gas-phase FTIR spectroscopy to discriminate aromatic rice cultivars offers several advantages over conventional methods. This non-destructive technique requires minimal sample preparation and provides real-time results, enabling rapid and high-throughput analysis. By targeting specific chemical functional groups within the rice samples, such as carbohydrates, lipids, and proteins, FTIR spectroscopy can capture intricate variations in molecular compositions among different cultivars.

To harness the potential of FTIR spectroscopy for aromatic rice differentiation, chemometric approaches play a pivotal role. Chemometrics, the applications of statistical and mathematical techniques to chemical data, enables the extraction of meaningful information from complex spectroscopic datasets [6,7]. Multivariate statistical analyses, such as Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA), allow the visualization of patterns and relationships within the data, aiding in classifying rice cultivars based on their chemical profiles.

In this context, this study explores the feasibility of using gas-phase FTIR spectroscopy coupled with chemometric techniques to differentiate aromatic rice cultivars. By analyzing the unique spectral signatures generated by FTIR, combined with advanced data processing methods, it becomes possible to achieve accurate and rapid discrimination among different varieties of aromatic rice. Such an approach holds significant potential for enhancing the quality control of rice products and supporting agricultural practices by enabling better cultivar management and traceability.

Materials and Methods

Rice Samples. Three in-bred cultivars of the Mabango series (NSIC Rc148, NSIC Rc342, and NSIC Rc344) obtained as hulls from the Philippine Rice Research Institute and International Rice Research Institute, and traditional Philippine aromatic rice varieties *Dinorado* and *Milagrosa*, and imported varieties, Basmati and Jasmine, purchased from the market, were studied. Three packs of every rice variety were mixed and homogenously divided into three laboratory-scale samples (1 kg each) and stored in air-tight containers.

Static Headspace Generation. Generation of headspace volatiles was conducted one hour before analysis. In brief, a 50-mL glass vessel with a two-hole cover containing 10 g of rice was purged with nitrogen gas for two minutes. A 50-mL syringe filled with nitrogen was attached to one hole of the cover while the other hole was sealed. The system was allowed to equilibrate for an hour at room temperature. The generated headspace was collected using another syringe through the other hole.

Gas-Phase Fourier Transform Infrared Spectroscopy. Analysis was carried out on an FTIR spectrometer (Shimadzu IR Prestige-21) equipped with a deuterated L-alanine-doped triglycemic sulfate detector. The generated headspace collected was transferred into a 100-mm short path EC gas cell (Pike Technologies) with a 25-mm x 4mm KBr window using a syringe. The FT-IR spectra were recorded continuously from $400 - 4000$ cm^{-1} with a resolution of 2.0 cm^{-1}. These spectra were automatically baseline-corrected. Moreover, a spectral atmospheric correction was applied to remove the background signals due to water and carbon dioxide. Each sample was measured thrice and the averaged, baseline, and background corrected spectrum was utilized for chemometric analysis.

Chemometric Analysis. Clustering is a technique of arranging and classifying components based on their similar characteristics in space. This results in high homogeneity among related groups and high heterogeneity among distinct groups. It assumes that the samples having similar numerical values will move closer to each other in space and thus would lie in areas with high proximity. This similarity is measured based on the distances of different data points with zero distance being perfectly related and represented by clusters. The most common type of clustering, AHC, was used in this study. This technique has been applied in discriminating varieties based on multi-element analysis of rice [7,8].

The large spectral dataset was transformed into small uncorrelated variables using PCA. Each variable, known as the principal component, is a linear combination regarded as a compressed representation of the original spectral data obtained by reducing the number of dimensions without much loss of information based on similarities and differences. Accordingly, when four or more variables are measured, data visualization becomes possible through various plotting systems [9]. While the whole spectral range may establish groupings, it is better to process the data and remove principal components that are highly correlated to avoid redundancy.

Both AHC and PCA analyses were performed using the XLSTAT software package (Addinsoft, USA). The PCA was processed using the wavenumber and peak area as variables and intensities, respectively.

Results and Discussion

Headspace generation. Static headspace sampling is a simple method carried out manually with a syringe that aims to establish equilibrium between the volatiles that a food material contains and those in the vapor phase above it. For this purpose, the food sample must be allowed to reach an equilibrium within a given time. In this study, rice samples in sealed glass containers were allowed to equilibrate at room temperature for one hour before the headspace was collected and transferred to the FTIR sample cell using a syringe. The headspace sample was kept small to avoid changes in sample characteristics due to sampling. In fact, Aganda et al. [10] observed a decrease in the total absorbance in the spectra when the sample weight was increased. Because of this, the method is unlikely to be successful for determining concentrations lower than or equal to parts per million range. Thus, the data was more useful in untargeted metabolomics or fingerprinting employing chemometric techniques.

Since the data was intended for untargeted metabolomics, varying the temperature for headspace generation was not investigated. In previous studies on headspace generation using an electronic nose, it was observed that volatile components of rice naturally stem out of the grains in a few minutes without heat application [11,12]. These volatile compounds have low molecular weight and are in gas form at room temperature. Indeed, despite the minimal contents of volatile compounds in rice, 2-acetyl pyrroline (2AP) and hexanal, can be easily perceived by the human olfactory system whose scents are indicative of whether the consumer accepts or rejects the samples.

Spectral Features of Aromatic Rice. A characteristic signature of chemical substances that are present in the rice samples featuring molecular vibrations such as stretching and bending in particular regions is provided by FTIR. Figure 1 shows the normalized FTIR spectra of the different rice varieties. Each spectrum contains a multitude of bands that exemplify the characteristic fingerprint of aromatic rice samples.
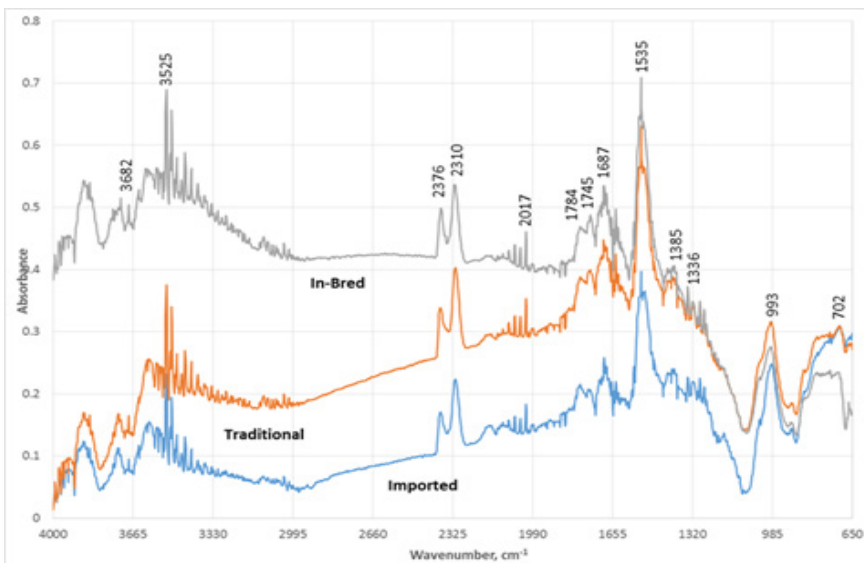

Figure 1. FTIR spectra of in-bred, traditional & imported rices in the mid-infrared region (4000-650 cm^{-1}).
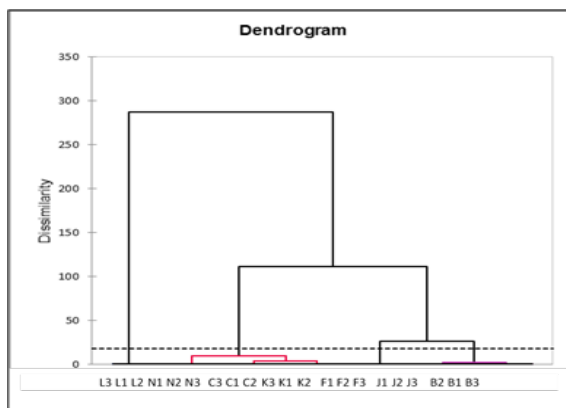
Figure 2. Dendogram of cluster analysis (Cluster I – *Mabango* 2 (L), Cluster II – *Dinorado* (N), *Mabango* 4 (C), and *Mabango* 5 (K), Cluster III – *Milagrosa* (F), Cluster IV – Jasmine (J) and Basmati (B))

The region from 3600-3000 cm^{-1} shows a broad vibrational band relating to –OH groups from carbohydrates, amide, and alcoholic structures and from moisture in the samples. Signals corresponding to the primary components of the atmosphere such as water and carbon dioxide were shown at 3525 cm^{-1} and double-peak at 2300 cm^{-1}, respectively. The sharp signal range around 2017 cm^{-1} arises from the stretching vibrations of –NCS bonds from isothiocyanite while the signal range from 1784-1687 cm^{-1} corresponds to the bending of –CH bonds arising from overtones of aromatic compounds. These results support our earlier study wherein an electronic nose sensor for sulfur compounds was instrumental in differentiating the aromatic rice samples [13].

The spectral range encompassing peak 1535 cm^{-1} relates to the secondary structure of proteins [14,15]. A medium signal at 1385 cm^{-1} signifies the bending of –CH bonds from aldehydes. Peaks below 1000 cm^{-1} could be from the bending of mono- and di-substituted alkenes or structures mostly related to carbohydrates [16].

The spectra illustrate bands associated with the main groups of organic molecules present in the samples. Table 1 presents the major peak ranges of rice samples captured from the spectra with their corresponding functional group vibrations and their probable sources.

Table 1. Major bands of rice samples, functional groups, and probable sources.

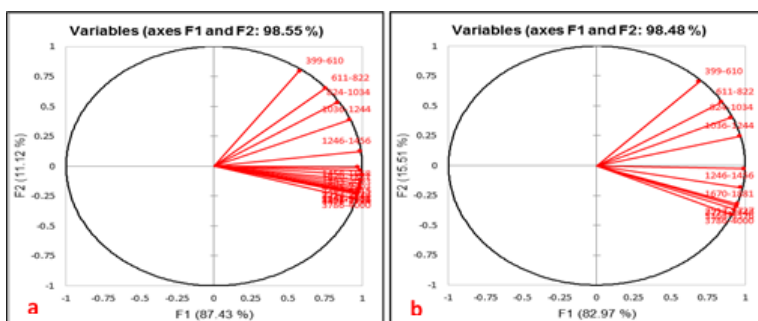| Frequency (cm^{-1})	Functional group vibration	Probable source
3700 – 3584	O-H stretching	Alcohol
2140 – 1990	N=C=S stretching	Thiocyanate
2000 – 1650	C-H bending	Aromatic compound
1745	C=O	Ketone
1550 – 1500	N-O stretching	Nitro compound
1390 – 1380	C-H bending	Aldehyde
1420 – 1330	O-H bending	Alcohol
995 – 985	C=C bending	Alkene
730 – 665		

Figure 3. Discrimination rate from PCA using a) seventeen and b) ten wavelength ranges

Based on the spectra, major organic volatiles believed to be present in the rice samples were alcohols, ketones, aldehydes, sulfur, and aromatic compounds. Differentiations of rice using electronic noses have suggested that these major compounds are present in rice [10,17] while gas chromatographic studies conducted on rice volatiles substantiate the presence of these compounds in most rice varieties [18,19,20].

Although the spectral curves in Figure 1 showed the same patterns, the absorbance of the inbred varieties was relatively higher than the others. This could be due to the fresh nature of the in-bred samples being kept as rough rice before the analysis. However, overlapping points manifested in the fingerprint region (1500-650 cm^{-1}). No observable differences were seen that can indicate that the rice samples can be identified directly from the figure. To resolve this, spectral differentiation was conducted.

Agglomerative Hierarchical Clustering Analysis and Rice Classification. AHC was done to identify the natural groupings of the varieties based on their Euclidian distance. The resulting dendrogram, presented in Figure 2, illustrates four clusters. The first cluster consists of an in-bred variety (*Mabango* 2). The second cluster consists of 2 inbred varieties (*Mabango* 4 and 5) and *Dinorado*. Cluster three consists of *Milagrosa* while the fourth cluster includes the imported varieties Basmati and Jasmine. It shows further that the dissimilarity between the two in-bred rice varieties in Cluster 2 was very small while those between the imported rice varieties was almost zero.

Principal Component Analysis and Dimension Reduction. PCA was performed for exploratory data analysis to identify natural clustering according to similarities and differences of the aromatic rice samples from the three groups that could influence the subsequent multivariate classification.

The variable plot of the PCA utilizing seventeen wavenumber ranges discriminated 98.55% of the samples (Figure 3a). Reducing this range to ten resulted to a non-significant decreased rate of discrimination (Figure 3b).
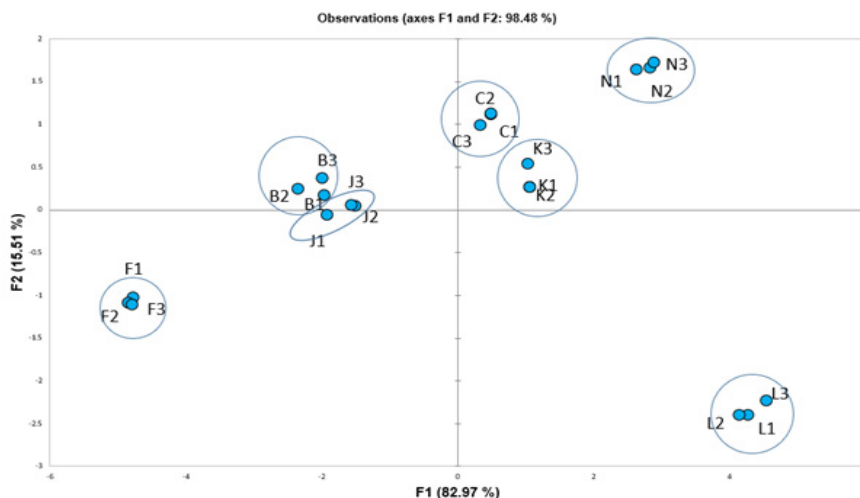
Figure 4. DPCA biplot of seven aromatic rice varieties.

By looking at the plot of scores generated by PCA using the ten wavelength ranges presented in Figure 4, qualitative recognition of different groups of samples was allowed. The figure shows the scatter plot of the first two Principal Components (F) for the seven aromatic rice samples, which together account for 98.48% of the total variability. Three groups can be suggested based on the distribution of samples along Principal Component 1 (F1). First, imported rice samples, Basmati and Jasmine rice, together with *Milagrosa* had negative F1 scores. Second are the two in-bred rice varieties NSIC Rc342 (*Mabango* 4) and NSIC Rc344 (*Mabango* 5), and *Dinorado* having positive F1 and F2 scores. Third, NSIC Rc148 (*Mabango* 2), although with a positive F1 score, has a negative F2 score. The F1 accounted for 82.97% of the total variation compared to the 15.51% posed by Principal Component 2 (F2). This indicates that the majority of the contributions to the pattern separation are contributed by F1.

Conclusion

Gas-phase FTIR spectroscopy coupled with chemometric data analysis demonstrated its effectiveness in differentiating aromatic rice varieties. Unsupervised pattern recognition techniques depicted obvious clustering of aromatic rice varieties in the AHC dendrogram and PCA plot. Both AHC and PCA provided clear discrimination of each aromatic rice variety from the others. The results show the efficiency of the gas-phase FTIR in differentiating the local and imported aromatic rice varieties and imply its potential utility in identifying aromatic rice.

The gas-phase FTIR is a rapid, non-destructive, and cost-efficient option to assess aromatic rice quality. The technique has potential uses for fast screening of analysis of aromatic rice varieties. However, it needs to be accompanied by established rice quality analysis for critical inspections of suspected fraudulent rice samples.

Acknowledgment

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization, F.B.S. and M.E.S.; methodology, F.B.S. and M.E.S.; data collection M.E.S.; analysis and interpretation of data, F.B.S. and M.E.S.; original draft preparation, M.E.S.; review and editing of the draft, F.B.S. All authors have read and agreed to the final version of the manuscript.

Institutional Review Board Statement

Not applicable.

References

[1] Sliwinska-Bartel M, Burns DT, & Elliott C. Rice fraud a global problem: A review of analytical tools to detect species, country of origin and adulterations. Trends in Food Science and Technology **2021**; 116, 36-46.

[2] Hu ZQ, Yang H, Chaima M, Gnag CY, Lu L, Hu XQ, Du B, Zhu Z, and Huang JY. A visualization and quantification method to evaluate the water-absorbing characteristics of rice. Food Chemistry **2020**, 331:127050.

[3] Ogawa Y. Chapter 17 - Quality Evaluation of Rice. Computer Vision Technology for Food Quality Evaluation (Second Edition) **2016**, 413-437.

[4] Ahmed A, Bilal M, Jabbar S, Mumtaz A, Khalid N, Asghar M, Amir RM, Ameer K & Malik AM. Physicochemical and sensory evaluation of indigenous rice genotypes. Journal of Agricultural Research **2020**, 58(2):89-96.

[5] Haas J, & Mizaikoff B. Advances in Mid-Infrared Spectroscopy for Chemical Analysis. Annual Review of Analytical Chemistry **2016**; 9, 45-68.

[6] Beebe KR, Pell RJ, & Seasholtz MB. Chemometrics: A Practical Guide. **1998**.

[7] Borges EM, Gelinski JMLN, de Oliveira Souza VC, Barbosa F, & Batista BL. Monitoring the authenticity of organic rice via chemometric analysis of elemental data. Food Research International **2015**; 77, 299-309.

[8] Yasui A, & Shindoh K. Determination of the geographic origin of brown-rice with trace-element composition. Bunseki Kagaku **2000**; 49, 405-410.

[9]   Kamal-Eldin A, & Andersson R. A multivariate study of the correlation between tocopherol content and fatty acid composition in vegetable oils. Journal of the American Oil Chemists' Society **1997**; 74, 375–380.

[10] Aganda KCC, Nonato MG, Sevilla FB, & Santiago KS. Headspace Fourier Transform infrared spectroscopy for the differentiation of Pandanus species. Talanta **2017**; 164, 439-444.

[11] Zhou B, &Wang J. Discrimination of different types damage of rice plants by electronic nose. Biosystems Engineering **2011**, 109(4):250-257.

[12] Yan C, & Lu A. A deep learning method combined with electronic nose to identify the rice origin. Journal of Instrumentation **2022**, 17:P08016.

[13] Serafico ME, & Sevilla FB. Differentiation of Philippine aromatic rice varieties using electronic nose. Sensors and Transducers **2018**; 28, 19-25.

[14] Movasaghi Z, Rehman S, & Rehman I. Fourier transform infrared (FTIR) spectroscopy of biological tissues. Applied Spectroscopy Reviews **2008**; 43, 134–179.

[15] Naumann A, Heine G, & Rauber R. Efficient discrimination of oat and pea roots by cluster analysis of Fourier transform infrared (FTIR) spectra. Field Crops Research **2010**; 119, 78–84.

[16] Kuhnen S, Ogliari JB, Dias PF, Boffo EF, Correia I, Ferreira AG, Delgadillo I, & Maraschin M. ATR-FTIR spectroscopy and chemometric analysis applied to discrimination of landrace maize flours produced in southern Brazil. International Journal of Food Science and Technology **2010**; 45, 1673–1681.

[17] Lu L, Deng S, Zhu Z, & Tian S. Classification of rice by combining electronic tongue and nose. Food Analytical Methods **2015**; 8, 1893-1902.

[18] Bryant RJ, & McClung AM. Volatile profiles of aromatic and non-aromatic rice cultivars using SPME/GC–MS. Food Chemistry **2011**; 124, 501–513.

[19] Cho S, & Kays SJ. Aroma-active compounds of wild rice (*Zizania palustris* L.). Food Research International **2013**; 54, 1463–1470.

[20] Lin JY, Fan W, Gao YN, Wu SF, & Wang SX. Study on volatile compounds in rice by HS-SPME and GC-MS. Julius-Kühn-Archiv **2010**; 425, 125-134.