Using Macroscopic, Microscopic, and FTIR Spectroscopy combined with Chemometrics to Authenticate Arabica Coffee from Antbush

David Fernando¹, Djoko Santosa¹,², Abdul Rohman¹,³*

¹ Centre of Excellence, Institute of Halal Industry and Systems, Universitas Gadjah Mada, Yogyakarta
² Department of Biological Pharmacy, Faculty of Pharmacy, Gadjah Mada University, Yogyakarta, Indonesia
³ Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Gadjah Mada University, Yogyakarta, Indonesia

ABSTRACT

Limited coffee stocks, high consumption, export demands, and increasing prices may contribute to counterfeiting. In other countries, including India and Africa, coffee forgery using antbush has been extensively established. Due to its abundance in Indonesia, unscrupulous coffee growers allegedly employ antbush (*Senna occidentalis*) as an adulterant in coffee to increase commercial earnings. This study aims to authenticate coffee from antbush using macroscopic and microscopic differences and FTIR spectroscopy combined with chemometrics. The arabica coffee samples from various regions were oven-roasted to a second crack and milled. The materials were then examined under macroscopic, microscopic, and infrared spectroscopic conditions. The obtained responses were used to monitor the qualitative and quantitative information in the targeted samples. Antbush in coffee samples is successfully identified microscopically by remnants of palisade tissue as well as the structural differences of the endosperm. In addition, FTIR spectroscopy combined with multivariate calibration can accurately estimate the concentration of antbush as adulterants in the target sample. Principle component regression (PCR) provides the best modeling for the relationship between the actual value. FTIR predicted the value of antbush with the lowest RMSEC and RMSEP values of 0.852 and 0.896, respectively, with the coefficient of determination ($R^2$) in calibration and validation models of 0.9996 and 0.9967, respectively. The combination of macroscopic, microscopic, and FTIR spectroscopy offered reliable tools to authenticate arabica coffee from antbush.

Keywords: antbush; arabica coffee; IR spectroscopy; macroscopy; microscopy

INTRODUCTION

As a product of the archipelago’s spices, coffee has also been acknowledged as one of Indonesia’s leading export commodities. Indonesia’s high consumption and export demand can trigger a need for coffee stocks from coffee supply companies. The lack of coffee stocks has forced some producers to commit fraud in counterfeiting coffee. Some unscrupulous coffee producers allegedly used antbush (*Senna occidentalis* (L.) Link) because of their circulation and abundance in Indonesia. Coffee counterfeiting using antbush has been widely reported abroad, especially in India and Africa (Shaheen et al., 2019).

Antbush is an invasive shrub that can be found growing wild on roadsides. Antbush originated from America and has spread to Sumatra (known as Andelan coffee) and Java (known as senting plant) (Yusarman, 2016). The seeds of this plant (dark-olive in color) are usually used as a coffee substitute. Because antbush can be consumed as coffee, this plant is also known as Senna coffee. Even though it is known to have good properties, these substitutions are often not accompanied by the inclusion of ingredient composition and the availability of adequate information. This phenomenon can be categorized as an act of coffee fraud or coffee counterfeiting (Mariod et al., 2017).

The coffee counterfeiting practice is detrimental to coffee lovers because the antbush belonging to the Leguminosae family is not a coffee plant. Plants with the Leguminosae family also contain specific proteins or allergens, which may have health implications for some consumers (Smits et al., 2021). This falsification can trigger an allergic reaction or even death if allergen information is not listed. In addition, the presence of bacteria, harmful chemicals, and a decrease in the nutritional content of food can cause further harm to coffee consumers (Visciano and Schirone, 2021). Thus, reliable analytical methods should be developed to mitigate such actions.

Previous studies have discussed the analytical methods for coffee authentication, highlighting differences between arabica coffee and antbush beans (Vionita et al., 2019; Hall et al., 2022). The fingerprinting analysis technique based on FTIR spectroscopy has been widely employed to quantify adulterants in coffee samples (Fernando et al., 2022). Pizarro et al., 2007 used this technique...
to predict the levels of robusta coffee in roasted coffee blends. In this study, the multivariate calibration of partial least square regression (PLSR) using the variable of the first derivative spectra provides a better model than normal spectral providing low RMSEC and RMSEP values. Suhandy et al., 2018 have developed fingerprint analysis techniques combined with PLSR to predict the levels of civet coffee mixed with arabica coffee or robusta coffee. The developed model provided low RMSEC and RMSEP and high values of $R^2$ either in calibration or validation samples. The model can reliably quantify Civet coffee in a mixed sample (test set). In addition, Correia et al., 2018 also used the FTIR spectroscopic method to quantify arabica coffee adulterated with robusta coffee. However, the studies on applying macroscopic, microscopic, and FTIR spectroscopy combined with chemometrics are very limited; therefore, this study aims to identify the differences between arabica coffee and antbush macroscopically and microscopically. This study also seeks to predict the content of antbush in coffee using FTIR spectroscopy combined with multivariate calibration.

**MATERIALS AND METHODS**

**Materials**

Eight samples of arabica coffee beans were obtained from the local markets around Yogyakarta and Central Java, Indonesia. Antbush beans were obtained from the Yogyakarta area as wild type. The eight different types of arabica coffee beans originated from other regions: Gayo, Gayo Natural, Halu, Kalosi, Flores, Papua, Kintamani Arabica, and Java Arabica. The authenticity of coffee beans was obtained in the Department of Pharmaceutical Biology under Prof. Dr Erna Prawita, Faculty of Pharmacy, Gadjah Mada University, Indonesia. Unless otherwise stated, the solvents and reagents employed were of analytical grade.

**Sample preparation**

The arabica coffee beans were roasted using various intensities. The coffee beans are oven-roasted until they undergo a second burst (Table 1). Antbush beans were washed with flowing water (defects or floating seeds were eliminated) and sun-dried for seven days before being oven-roasted until cracking. Then, using a coffee grinder with a grind size setting of 5, the arabica coffee and antbush beans were grounded. The ground beans were sealed into an airtight silica gel pack container. The samples were used up to three months after the grinding process.

**Macroscopic and microscopic evaluations**

Macroscopy and microscopic methods are considered conventional pharmacognosy methods widely applied in plant identification (Kumar et al., 2011). "Organoleptic observation" refers to the macroscopy practice of observing the seed or beans’ scent, color, size, and other physical qualities without using tools. On the other hand, microscopy evaluation is usually done using the help of a microscope. Arabica coffee beans and antbush seeds were soaked in distilled water for 24 hours, scraped, and then cut longitudinally and transversely. They were then drizzled with a chloral hydrate solution (2:1), heated briefly in a spirit burner, and examined under a microscope. The heating aims to hasten the chloral hydrate solutions during the bleaching process. To determine the fragments in each sample, the tissue from the antbush seed and the arabica coffee bean are compared under magnifications of 40x, 100x, and 400x, respectively. Target samples are created using forgers in varying concentrations ranging from 0-100% of antbush in arabica coffee.

**Scanning of FTIR spectra**

Ground coffee and antbush are analyzed using FTIR. FTIR spectra of the studied samples were scanned using FTIR with an ATR diamond detector (Thermo Nicolet iS10 instrument). The diamond detector is cleaned with 96% acetone before the sample is placed on top of it. The spectral acquisition was performed using 32 scanning at 4000-400 cm$^{-1}$ with a resolution of 4 cm$^{-1}$. The choice of a small resolution aims to make the peaks visible because the smaller the resolution, the more the peaks will be seen. The background spectrum of air was measured before scanning samples to avoid variations in the spectra from one sample to another, repeated every 120 minutes, and the sample readings were taken one by one. Each sample was replicated in three readings.

**Chemometrics analysis**

PCA was conducted using Minitab® version 19 (Minitab Inc., USA). The variables inputted are absorbances at certain wavenumbers (using the peaks detected). The wavenumbers used for modeling are based on recommendations from the software and were selected manually based on the peak intensity between the arabica coffee and antbush powder. The pre-processing of spectral data was carried out using standard spectral data, first derivative, and second derivative. In addition, baseline corrections were made for all FTIR spectra during modeling with the "linear removal" feature. The data is disregarded if one of the
components has a significant standard deviation. Data normalization using the software’s mean centering mode was also applied. The evaluation of multivariate calibrations of classical least square (CLS), stepwise multiple linear regression (SMLR), partial least square regression (PLSR), and principal component regression (PCR) was conducted using TQ Analyst™ version 9 (ThermoFisher Scientific Inc, USA) to predict the quantity of adulterant in target samples.

The binary selections were deployed for the modeling purpose with the concentration ranges of antbush of 0-50% wt/wt. The best model was then evaluated using the Paired-Samples T Test (IBM SPSS® Statistics version 20) and regression modeling using Excel 2018 (Microsoft Inc, USA).

RESULTS
FTIR spectroscopy combined with chemometrics

Figure 1 revealed FTIR spectra of arabica coffee [A] and antbush beans [B] scanned at mid-infrared regions. The functional groups responsible for IR absorptions are peak [A] around 3340 cm⁻¹ coming from -OH stretch (hydrogen bonding), peak [B] at 3009 cm⁻¹ due to cis -C=CH stretch, [C] at 2929 cm⁻¹ due to -CH₃ asymmetric, [D] at 2859 cm⁻¹ due to -CH₂ symmetric, [E] at 1746 cm⁻¹ comes from -C=O stretch, [F] 1657 cm⁻¹ is due to cis C=C, [G] 1519 cm⁻¹ is due to -NH bend, [H] 1455 cm⁻¹ due to CH₂ bend, [I] 1375 cm⁻¹ due to -OH bend, [J] 1240 cm⁻¹ due to C-O stretch, [K] 1153 cm⁻¹ due to C=OH stretch, [L] at 1062 cm⁻¹ in coffee arabica is due to -OH stretching, while [L] at 710 cm⁻¹ in antbush is due to CH-CH vibration, [M] at 1036 cm⁻¹ is due to C-O, and [N] at 714 cm⁻¹ is coming from CH-CH vibration (Coates, 2006; Pavia et al., 2001). The FTIR spectra of coffee arabica powder can be differentiated from those of antbush powder, mainly in 3 different regions (3000–2800 cm⁻¹, 1800–1600 cm⁻¹, and 1500–1000 cm⁻¹), as revealed in Figure 1C. The absorbance values of antbush and arabica coffee powder in all regions were then selected for principal component analysis (PCA) using Minitab® version 19 software.

After the selection process, the absorbance values at 43 wavenumbers were selected as the variable for PCA, and the score plot was obtained, as depicted in Figure 2A. Antbush is grouped in the left quadrant (III), while arabica coffee forms a group in the right quadrant (II and IV). Eigenanalysis data from PCA revealed that the variable influencing the score plot is the absorbance value at 2925 cm⁻¹. PC1 and PC2 contributed to 97.4% variances. PC1 is the loading of absorbance obtained from a wavelength of 2925 cm⁻¹, which can be seen from the most considerable Eigenvector value in column PC1 (0.481). PC2 is the loading of absorbance value obtained from wavenumbers of 2925 cm⁻¹, which can be seen from the most considerable Eigenvector value in column PC2 (0.238). This Eigenvector value proves that the absorbance value at certain wavenumbers is very influential in the differentiation between samples of arabica coffee (PC2 on the y-axis) and samples of antbush against arabica coffee (PC1 on the x-axis), as shown in Figure 2B. The results show that arabica coffee has more compounds containing the vibration of -CH₂ and -CH₃ stretching than those of antbush samples.

The absorption of the raw mixed data was then plotted using TQ Analyst™ version 9 software using the multivariate calibrations of CLS, SMLR, PLSR, and PCR modeling algorithms to predict adulterant levels (antbush beans) in arabica coffee. The performance of multivariate calibrations using the FTIR spectra previously subjected to preprocessing along with statistical parameters of coefficient of determination in calibration model (R²-cal), coefficient of determination in validation model (R²-val), RMSEC, and RMSEP was compiled.
Figure 1. Spectra of different varieties of arabica coffee powder [A] and spectra of antbush powder [B] scanned using mid infrared region (4000-650 cm⁻¹), along with FTIR spectral difference between arabica coffee and antbush bean [C]. Spectral difference between the two is circled in red.
Using Macroscopic, Microscopic, and FTIR Spectroscopy combined with Chemometrics

Majalah Obat Tradisional, 29(1), 2024

The multivariate calibration of PCR using absorbance values of first derivative FTIR spectra at wavenumbers of 2922-1472 cm\(^{-1}\) provides the best model for predicting the concentration of antbush beans as an adulterant in arabica coffee. The model was selected based on the lowest values of RMSEC and RMSEP obtained (0.852 and 0.896, respectively) and the highest values of \(R^2\)-cal and \(R^2\)-val (0.9996 and 0.9967, respectively). The smaller the value of RMSEC and RMSEP, the better the regression model. Models are generally considered acceptable if the RMSEP value is not more than 1.2x RMSEC (Zhan et al., 2017). Thus, the PCR model for making the relationship between the actual and FTIR predicted value of antbush is good because the RMSEP value is not more than 1.2x RMSEC.

PCR modeling is then used to estimate the concentration of antbush in the targeted sample. Table III compiled the prediction results of antbush levels as an adulterant in arabica coffee, with a level of as low as 1% wt/wt could be predicted. The statistical test applying Paired-Samples t-test was applied to compare the actual and expected contents of antbush as predicted using FTIR spectra combined with PCR. The significance value \(P\) of 0.659 \((P > 0.05)\) was obtained, indicating that there is no significant difference between the actual value of antbush and the predicted value. This nominal value implies the success and effectiveness of the PCR model in estimating adulterant levels in arabica coffee. In addition, the predictive efficacy of PCR modeling was also tested using the regression method, and the evaluation of the predictive capability of the model was based on the coefficient of determination \(R^2\)-value. The correlation between the actual value \((x\text{-axis})\) and predicted value \((y\text{-axis})\) of antbush provided the regression equation of \(y = 0.9969x - 0.5114\) with an \(R^2\) value of 0.9923. If the \(R^2\) value > 0.8, then the regression model is reliable in predicting the level of counterfeiters in the targeted sample (Suhandy et al., 2018). This model proves that FTIR spectra using the PCR model with first-derived data are reliable and effective in predicting the levels of counterfeiters in the target sample.

**Macroscopic and microscopic evaluation of arabica coffee adulterated with antbush**

Macroscopically, arabica coffee beans have a distinctive shape, half-oval and flat. One side of the seed has a flat contour, while the other is a convex contour. The surface of the seeds has a smooth and glossy texture. Arabica coffee beans have a distinctive aroma and taste bitter and sour. Seed length ranges from 1 to 1.5 cm with a width ranging from 0.9 to 1 cm. Unlike arabica coffee beans, antbush beans are flattened circular objects with tapering ends on both sides. Antbush seeds have a smooth and silky surface and a dark brown tint. The seeds smell and taste like peanuts and have a splitting shape. The width of a seed is between 0.1 and 0.2 cm, while its length is between 0.3 and 0.4 cm.

The color of the two beans is similar because the roasting procedure has eliminated these differences. Arabica coffee and antbush beans are easy to distinguish when mixed, given the differences in size and shape of the beans with their characteristics. However, when ground into a powder, the perceptible difference depends only on the smell and taste because the size and shape factors are eliminated during the roasting and grinding. In the sample mixture of arabica coffee and antbush (50:50), the sharper taste and smell of
coffee masked the taste and smell of antbush. Thus, macroscopic identification is difficult when the powders from the two seeds have been mixed.

From microscopic evaluation, the transverse and longitudinal slices of coffee beans revealed that coffee beans comprise three types of tissue: sclerenchyma, parenchyma, and endosperm. Sclerenchyma tissue and parenchyma tissue are the constituents of silver skin in coffee. This finding aligns with that reported by Ukers.
Using Macroscopic, Microscopic, and FTIR Spectroscopy combined with Chemometrics

Majalah Obat Tradisional, 29(1), 2024

Table III. The correlation between actual value of antbush and calculated value as predicted using FTIR spectra and principle component regression.

<table>
<thead>
<tr>
<th>Samples</th>
<th>Actual value (% wt/wt)</th>
<th>Calculated value (% wt/wt)</th>
<th>Diff x Path</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample A</td>
<td>12,5</td>
<td>10,45</td>
<td>-2,05</td>
</tr>
<tr>
<td>Sample B</td>
<td>50</td>
<td>42,76</td>
<td>-7,24</td>
</tr>
<tr>
<td>Sample C</td>
<td>87,5</td>
<td>86,08</td>
<td>-1,42</td>
</tr>
<tr>
<td>Sample D</td>
<td>25</td>
<td>24,94</td>
<td>-0,06</td>
</tr>
<tr>
<td>Sample E</td>
<td>100</td>
<td>103,08</td>
<td>3,08</td>
</tr>
<tr>
<td>Sample F</td>
<td>0</td>
<td>3,54</td>
<td>3,54</td>
</tr>
<tr>
<td>Sample G</td>
<td>1</td>
<td>0,72</td>
<td>0,28</td>
</tr>
</tbody>
</table>

(Ukers, 2009). In contrast to the tissue observed in arabica coffee beans, antbush seeds were composed of 5 types: epidermal tissue, palisade tissue, hourglass cells, mesophyll tissue, and endosperm tissue. In cross-sectional and longitudinal sections, a layer of hourglass cells covers the mesophyll layer of palisade tissue (Figure 3). The antbush fragments observed were brown due to the roasting process. This finding agrees with what Az reported (Az et al., 2013).

Arabica coffee beans and antbush were ground using a mortar to observe the identification fragments under a microscope. The oil cells, parenchyma cells, sclerenchyma, and endosperm fragments were identified in the arabica coffee powder sample (Figure 4). The intensity of the essential oil observed is small because the volatile oil quickly evaporates during the roasting process. The easily recognizable fragments in the antbush seed sample were palisade, mesophyll, and endosperm. The arabica coffee and antbush ground samples differed microscopically because the antbush had easily recognizable palisade tissue fragments. In addition, the endosperm tissue of the two samples differed structurally. Antbush has endosperm, composed of spherical cells, while arabica coffee has endosperm tissue with a web-like structure. Arabica coffee has no epidermal (palisade) and mesophyll structure like antbush.

Eight targeted samples (marked with I to VIII) were prepared to test the effectiveness of the microscopic test in identifying the samples of arabica coffee adulterated with antbush. Target samples were made with different concentrations of forgers up to a concentration of forgers of 1%. The microscopic proved the presence of antbush in the target sample up to an adulteration concentration of 1% wt/wt (Sample VIII). Figure 5 indicates the presence of identification fragments as a palisade in targeted samples I, II, III, IV, V, and VIII.

**DISCUSSION**

Despite having the same functional group interpretations, the peaks from the two sets of samples (arabica coffee vs antbush) were found to be different under FTIR. Further chemometrics analysis proves FTIR as an effective tool in differentiating fingerprint profiles of oven-roasted arabica coffee against antbush. Adulterants can be reliably detected using the microscopy method down to 1% wt/wt. Although microscopic identification can identify adulterated samples with an accuracy of 100%, this method is only limited to qualitative observations. It is probable that other types of adulterants, such as rice, peanuts, and soybeans, could be detected using the same methods.

This result, however, might not represent realistic conditions from an industrial point of view, considering that the roasting process is carried out by oven and the grinding process was done using a manual mortar. In addition, the number of data points (contamination concentrations) is few, and the contaminant was used at too large a concentration, thus making the chemometrics model work. Further research must be done to provide data from an industrial point of view (i.e., roasting, grinding, and sieving process) and to prove that the chemometrics model works for smaller concentrations and ranges of contaminants.

**CONCLUSION**

Arabica coffee beans can be distinguished by sight through macroscopic and microscopic approaches. Once it is powdered and mixed, the macroscopic approach is challenging. The microscopic approach can detect the adulteration of arabica coffee grounds by antbush to a down level of 1% wt/wt through the identification fragments of each ingredient. Furthermore, FTIR spectroscopy combined with the PCR model can
quantitatively estimate the concentration of adulterant (antbush) in arabica coffee samples. PCR using the first derivative FTIR provides accurate and precise modeling as indicated by low RMSEC and RMSEP and high values of R² in calibration and validation models. The combination of macroscopic, microscopic, and FTIR spectroscopy-PCR are practical tools for analyzing the arabica coffee adulteration with antbush.
CONFLICT OF INTEREST
During this research, the authors declared no conflict of interest.

ETHICAL APPROVALS
This research does not involve any animal or human subjects.

REFERENCES


Majalah Obat Tradisional, 29(1), 2024