Classification of java tea (Orthosiphon aristatus) quality using FTIR spectroscopy and chemometrics

R Heryanto1,2*, D I Pradono1,2, E Marlina2 and L K Darusman1,2
1Biopharmaca Research Center, Bogor Agricultural University, IPB Taman Kencana Campus, Jl. Taman Kencana No 3, Bogor, Indonesia.
2Department of Chemistry, Faculty of Mathematics and Natural Sciences, Bogor Agricultural University, Bogor, Indonesia.

E-mail: kemometrik@gmail.com

Abstract. Java tea (Orthosiphon aristatus) is a plant that widely used as a medicinal herb in Indonesia. Its quality is varying depends on various factors, such as cultivating area, climate and harvesting time. This study aimed to investigate the effectiveness of FTIR spectroscopy coupled with chemometrics for discriminating the quality of java tea from different cultivating area. FTIR spectra of ethanolic extracts were collected from five different regions of origin of java tea. Prior to chemometrics evaluation, spectra were pre-processed by using baselining, normalization and derivatization. Principal Components Analysis (PCA) was used to reduce the spectra to two PCs, which explained 73% of the total variance. Score plot of two PCs showed groupings of the samples according to their regions of origin. Furthermore, Partial Least Squares-Discriminant Analysis (PLSDA) was applied to the pre-processed data. The approach produced an external validation success rate of 100%. This study shows that FTIR analysis and chemometrics has discriminatory power to classify java tea based on its quality related to the region of origin.

1. Introduction
The use of java tea as a medicinal herb makes it a good potential in pharmaceutical world. Currently, people choose herbal remedies because they are rarely causing side effects and the price is lower than synthetic medicines. However, there are problems in the production of herbal medicines, such as availability and quality of raw materials, materials standardization, quality stability and control that are considered as challenging [1]. Bioactive components contents of medicinal herbs vary widely depends on the species, variety, geographical origin, cultivation, harvesting method, and post-harvest processing. These variations may cause inconsistency in the efficacy, quality, and safety of the herbal products. Therefore, it is needed a serious control in the determination of specifications and parameters for raw materials [2].

Qualitative method frequently used to analyze active compounds is FTIR (fourier transform infrared) spectroscopy providing information about the presence of functional groups in a sample and commonly used to see the fingerprint of a sample. Every type of active compound will give a unique IR spectrum that

* Corresponding author
may be used to see the response consistency of raw material quality standardization. Other methods generally used in the quality control of herbal medicines raw materials are HPLC, GC, GC-MS, and TLC. These methods, however, need sample preparation and a relatively long time of examination process.

Medicinal herb analysis using FTIR spectrophotometer generates a very complicated FTIR spectra since this spectra is a result of chemical compounds interaction in a complex sample matrix. This spectra is difficult to be directly interpreted that it is needed chemometrics, a method used to get both qualitative and quantitative information from the spectra. Combination of infrared spectrometry and chemometrics has been widely used as a rapid quality control method of herbal medicines with extensive variety and for example, it has been used for discrimination of *Gentiana rigescens* from different origins [3].

Multivariate analysis frequently used are PCA, PLSDA, and PLS, among others. PCA and PLSDA may be used to group samples, while PLSDA is a common regression technique for multivariate data that may be used to predict hidden information in mixed spectra [4].

This study aimed to conduct quality control of java tea raw materials from several regions having relatively different geographical conditions by grouping and making a prediction model to relate the IR spectra characteristics and java tea raw materials. The parameters of the raw materials and extracts tested are ash content, water content, total phenol, and total flavonoid.

2. Experimental Section

This study consisted of 3 main working steps, sample preparation, quality analysis, and chemometrically data processing. Sample preparation includes harvesting, drying, and sample extraction. Quality analysis was conducted to dried samples and to the obtained extracts in the sample preparation step. Dried samples were tested for water and ash contents, extract samples were analyzed for total phenol and total flavonoid contents. FTIR analysis was also performed to java tea extracts to obtain IR spectra. Absorbance data and the wavelength of the IR spectra were chemometrically processed using The UnscramblerX 10.3. Chemometric methods used were PCA and PLSDA. The results from the chemometrically processed data and quality data from the quality analysis generated a model of java tea classification from 5 different regions.

2.1 Sample Preparation [5,6]

The leaves of white-flowered java tea from 5 different regions were harvested in the afternoon. The leaves were dried using oven at 50°C for ±24 hours. Powdered samples were prepared and sieved using 40-meshed sieve. Ten gram of sample was then extracted in 5 replicates for each region using 100 mL ethanol 40%. The extract was then put in a water bath shaker for 120 minutes at 65°C. Subsequently, the resulting filtrate was filtered by using filter paper. This extract was concentrated with rotary evaporator and then kept in a bottle and put in refrigerator.

2.2 Total Ash Assay [7]

Two gram of sample was put in a pre-combusted and pre-weighted silicate crucible. The tested sample was slowly combusted until the charcoal was no longer left, ashed at 600°C until complete ashing was achieved, then cooled and weighted. Combustion was repeated until it was obtained a constant weight. Total ash content was calculated as %w/w of sample.

2.3 Water Content Assay [8]

Empty dish was dried in an oven at 105°C for 3 hours, transfer to desiccator to cool, and weighted. Two gram of sampel (a) was weighted and put in the dish, and dried at 105°C for 3 hours, transfer to desiccator to cool and weighted. The dish and the sample were reweighted until it was obtained a constant weight (b).
Water Content (%) = \frac{a-b}{a} \times 100\%

2.4 Total Phenol Assay [9]
Total phenol content in an extract was determined by using Follin-Ciocalteu (FC) reagent and external calibration curve was created by using gallic acid. 10 mg Extract was weighted, dissolved in 5 mL of water, and diluted 20X. To 3 mL aquadest, 2 mL absolute ethanol, and 0.5 mL 50% FC reagent (v/v) it was added 2 mL extract. The mixture was left for 5 minutes and then 1 mL 5% Na$_2$CO$_3$ (w/v) was added. The mixture was homogenized and then incubated in the dark for 1 hour. The absorbance was measured with UV-VIS spectrophotometer at 725 nm.

2.5 Flavonoid Content Assay [7]
Two hundred milligram of extract was hydrolized with 1 mL 0.5% hexamethyltetraamine 0.5% (w/v), 20 mL aceton, and 2 mL HCl 25% in water. Hydrolisis was performed by refluxing for 30 minutes. The hydrolized mixture was filtered with cotton into a 100 mL volumetric flask and aceton was added up to the measuring mark. 20 mL Filtrate was then transferred into a separatory funnel, added with 20 mL aquadest, and extracted 3 times, each with 15 mL ethyl acetate. Ethyl acetate fractions were collected into a 50 mL volumetric flask and added up with ethyl acetate up to the measuring mark. Spectrophotometry analysis was started by transferring 10 mL of the ethyl acetate fraction solution into a 25 mL volumetric flask and 1 mL 2% AlCl$_3$ in 5% glacial acetic acid (in methanol) was added. 5% Glacial acetic acid was added up to the measuring mark and the absorbance was measured at 425 nm. Pure quercetin solution in acetic acid was used as standard solution.

2.6 FTIR Analysis
In order to make a pellet, 2 mg of ethanol extract was mixed with 200 mg KBr. The pellet was made by using a hand press with pressure of 80 kN for 10 minutes. The spectrum was measured using FTIR spectrophotometer in the range of 4000-400 cm$^{-1}$. The spectrum data was normalized that the lowest data value was set to 0, while the highest value was set to 2. The result of normalization was baselining-corrected to make the spectra baseline in the 0 absorbance, followed by first derivatization and Savitsky Golay method smoothing.

2.7 Chemometrically Data Analysis
FTIR spectra were saved in OPUS format containing data in the absorbance mode. There were total thirty five samples spectra where seven samples came from each origin. Twenty five spectra were used as training data and two spectra from each origin were used for test data. Samples grouping were performed by using PCA and PLSDA. PCA and PLSDA analysis of the resulting FTIR spectra used The UnscramblerX 10.3 software.

2.8 Statistical Data Analysis
The mean values of the data from water content, ash content, extract yield, total phenol and total flavonoid were determined and standard deviations were also determined. Further analysis was performed for variance and Duncan’s Multiple Range Test.

3. Results and Discussion
3.1 In Water Content, Ash Content, and Extract Yield
Water and ash content assays in this study were performed by gravimetric methods. The water and ash contents of dried samples from the five regions had fulfilled the maximum requirements in FHI (2008),
which are less than 10%. Water content shows the water contained in the sample materials. Water content assay may assist in the determination of the material actual weight and used in the extract yield calculation. The lower the water content, the higher the material stability and the lower the material destruction [10]. On the other hand, ash content shows the internal and external mineral contained in the material and related to the purity of the materials as well as materials contamination [11]. The values of water content, ash content, and extract yield are significantly different (p<0.05) for some particular regions (Figure 1).

The lower the ash content, the lower the possibility of the material to be contaminated. The heavier the extract yield, the more chemical compounds extracted from the raw materials. In general, quality is said to be good if the water content is low, the ash content is low, and the yield is high. Samples with the lowest water content comes from the sample from Nagrak with value significantly different from the samples from the other regions. The lowest water content value from the samples from Cigombong is indistinguishable with the samples from Leuwiliang and has the highest extract yield compared to the other regions (Figure 1).

![Figure 1. Water content, ash content and extract yield. Values are given in mean ± sd. Values marked with (a-d) are significantly different (p<0.05)](image)

### 3.2 Metabolites Content and Antibacterial Activity of Java Tea Leaves Extract

Analysis of the chemical compounds content in java tea leaves extracts was performed to the total phenol and total flavonoid contents. In java tea, phenol is an important compound since it has a significant role as an antioxidant. There are 20 types of phenol in java tea: 9 lipophilic flavons, 2 flavonol glycosides, and 9 caffeic acid derivatives [12]. Flavonoid is a main group of phenol in plants and sinensetin belongs to the group of methoxy flavon or lipophilic flavonoid [13].

The samples from Nagrak, based on the results of this study, has the highest total phenol content and significantly different from the samples from the other regions (Figure 2). The samples from Cimanggu has the highest total flavonoid content but insignificantly different with the samples from Nagrak. In general, the samples from Nagrak and Cimanggu have higher chemical content quality compared to other samples. On the contrary, the samples from Pacet has the lowest quality.

The results of chemical compounds content analysis correspond with the geographical characteristics of respective regions. Java tea grows very well in the altitude of 100-1000 mdpl, tropical climate, average rain fall of 3000 mm/year, and full sunlight [14]. Nagrak and Cimanggu meet the requirements for the growth of java tea. Cigombong has lower rain fall, Leuwiliang has lower temperature, while Pacet either has lower temperature or height that achieve 1400 mdpl. Further, the planting pattern of Pacet and Leuwiliang is polyculture, while the others monoculture. According to [10], other factors that may influence the quality of herbal medicines such as species variation, harvesting time, the part of the plant
used, and post-harvest treatments are considered as do not give any effects because, in this study, they are same for each region.

![Figure 2](image)

**Figure 2.** Total Phenol (I) and Total Flavonoid (II) contents of sample extracts from 5 regions. The values are given in mean ± sd. The values marked with (a-d) are significantly different (p<0.05)

Linear correlation between total phenol content and total flavonoid content in the sample can be seen from the results of Pearson Correlation test. The coefficient of correlation between total phenol and total flavonoid is 0.744, showing that the total phenol and total flavonoid contents are positively correlated and this correlation is relatively strong. Total phenol contents in the samples are high and so are the total flavonoid contents (Table 1).

| Total Phenol | Total Flavonoid |
|--------------|----------------|
| Total Phenol | 1              |
| Total flavonoid | 0.744       |
|               | 1              |

The correlation value between phenol and flavonoid is relatively strong, which is due to most flavonoids belong to phenol. However, there are also caffeic acid derivatives, tannins, and other compounds in the samples that are not belong to phenol.

### 3.3 FTIR Spectra of Java Tea Leaves Extracts

Every compound in a medicinal herb has an important role in a complicated mixed system for its effect to the efficacy of the plant. FTIR spectra generated (Figure 3) are the absorbances of various chemical components in the java tea leaves extract. The FTIR spectra of the samples from the five regions are not visually different showing that the chemical components contained are generally the same.

A strong and wide absorbance observed at the wave number of about 3400 cm⁻¹ shows the presence of hydroxyl group like in polyphenol, and at 2962 cm⁻¹ and 2872 cm⁻¹ show C-H of aldehyde, at 1600 cm⁻¹-1700 cm⁻¹ show C=O, 1600-1420 cm⁻¹ shows phenyl, 1456 cm⁻¹ and 1382 cm⁻¹ show C-H of CH₃, and 1270-1150 cm⁻¹ shows ester (C-O) [15].
3.4 FTIR Spectra Grouping using PCA and PLSDA

FTIR spectra groupings of java tea leaves extracts from the five regions were performed by PCA. This method is not able to group the initial sample spectra or spectra without any pre-processes (Figure 4a). This is because the initial spectra are still influenced by baseline shifts, difference in the number of samples analyzed, and the noise produced by the detector. This influence may be resolved by pre-processing the spectra including baseline correction, normalization, and derivatization. These pre-processing techniques could improve PCA capability to group samples without loosing much information with 73% total variance obtained.

Figure 4b shows samples grouping from the five regions and each region was marked with different color. Samples from Pacet (P) found in quadrant I, samples from Cigombong (G) found in quadrant II and III, samples from Leuwiliang (L) and Cimanggu (C) found between quadrant I and IV, while samples from Nagrak (N) found in quadrant IV. Samples from Cigombong exhibits a very clear separation from the samples of the other four groups of regions. This could mean that there are possibilities the Cigombong samples have quality characteristics and chemical contents different from the samples of the other regions.
Quality groupings in Table 2 aim to facilitate the quality evaluation of the samples from each region. The quality group I is a group of samples with the best criteria compared to other samples based on significance test and the quality tested. The best quality comes from the samples from Nagrak because all test results are in the quality group I, except for ash content. Samples from Cimanggu is in the second place after Nagrak, and the lowest is the samples from Pacet. These results correspond with the results of sample grouping using PCA. The sample with the best quality lies in quadrant IV, medium quality lies between quadrant I and IV, and the lowest quality lies in quadrant I. Samples from Cigombong has good physical quality, but low in chemical content quality.

| Group of Quality |
|------------------|
| Water Content   |
| Ash Content     |
| Extract Yield   |
| Total Phenol    |
| Total Flavonoid |

Samples grouping by using PLSDA was performed to 2 matrix, i.e. the absorbance data from the results of FTIR analysis as X matrix and response matrix for every sample’s region as Y matrix. Response value of 1 is given for the samples successfully predicted to their own region of origin and 0 for the samples predicted not to their own region of origin. Calibration model was well built for having R² approaching 1 and RMSE (Root Mean Square Error) approaching 0 for all samples from each region. Prediction models of samples from Cimanggu and Pacet are relatively good compared to the other 3 regions (Table 3).

| Sample | Callibration R² | Callibration RMSEC | Prediction R² | Prediction RMSEP |
|--------|-----------------|--------------------|---------------|-----------------|
| G      | 0.9739          | 0.0646             | 0.9061        | 0.1313          |
| C      | 0.9792          | 0.0577             | 0.7578        | 0.2109          |
| P      | 0.9948          | 0.0289             | 0.9304        | 0.1131          |
| L      | 0.9789          | 0.0581             | 0.8762        | 0.1508          |
| N      | 0.9790          | 0.0579             | 0.8561        | 0.1626          |

The results of prediction of 2 samples randomly selected from each region can be seen in Table 4. The prediction values will approach 1 if the samples from one region tested to the model of its own region and vice versa. The results of prediction are not sufficiently good for the samples from Leuwiliang and Nagrak, replicate 1. This may be due to the low prediction capability of the model and difference in the FTIR spectra characteristics of the samples tested.
| Sample     | Replicate | Prediction Value in PLSDA Model |
|------------|-----------|---------------------------------|
|            |           | Cigombong | Cimanggu | Pacet | Leuwiliang | Nagrak |
| Cigombong  | 1         | 0.9029    | 0.3485   | -0.0291 | -0.1930 | -0.0292 |
|            | 2         | 0.7750    | 0.1535   | 0.0521   | -0.0327 | 0.0521   |
| Cimanggu   | 1         | -0.0290   | 1.2399   | -0.0320 | -0.1168 | -0.0621 |
|            | 2         | 0.0853    | 0.7219   | -0.1987 | 0.1798   | 0.2117   |
| Pacet      | 1         | 0.0319    | -0.4598  | 0.7702   | 0.2986   | 0.3590   |
|            | 2         | 0.0395    | 0.1823   | 1.0255   | -0.2018  | -0.0455  |
| Leuwiliang | 1         | 0.1085    | 0.2847   | 0.0624   | 0.5712   | -0.0268  |
|            | 2         | -0.0823   | 0.0820   | -0.0722  | 1.0709   | 0.0016   |
| Nagrak     | 1         | -0.1232   | 0.3936   | 0.1720   | -0.1053  | 0.6628   |
|            | 2         | 0.0446    | 0.1379   | -0.0059  | -0.0807  | 0.9041   |

4. Conclusion

Extract samples of java tea leaves from 5 different regions are successfully grouped based on the results of PCA analysis and the physicochemical quality. The samples from Nagrak have better quality than the other regions, in which the score plot lies in quadrant IV. Medium quality lies in quadrant IV and I, which is for the samples from Cimanggu and Leuwiliang. The quality for the samples from Pacet is not sufficiently good and lies in quadrant I. The samples from Cigombong tend to be different from other samples, which lies between quadrant II and III and have low total phenol and total flavonoid contents. Samples grouping was also performed using PLSDA and almost all samples could be predicted based on their respective region of origin.

References
[1] Bandaranayake WM 2006 Modern Phytoedicine. Turning Medicinal Plants into Drugs. (Weinheim: WILEY-VCH Verlag GmbH & Co. KGaA)
[2] Komarawinata HD 2008 Budidaya dan pasca panen tanaman obat untuk meningkatkan kadar bahan aktif. (Bandung: Unit Riset dan Pengembangan, PT Kimia Farma (Persero) Tbk.)
[3] Zhao Y, Zhang J, Jin H, Zhang J, Shen T, Wang Y 2015 J. AOAC Int. 98 22-6
[4] Brereton RG 2003 Chemometric: Data Analysis for the Laboratory and Chemical Plant (England: John Wiley & Sons. Ltd.)
[5] Abdullah S, Ahmad MS, Shaari AR, Johar HM, and Noor NFM 2011 Int. Conference on Food Engineering and Biotechnology vol 9 p 305-9.
[6] Chew KK, Khoo MZ, Ng SY, Thoo YY, Wan Aida, WM, and Ho CW 2011 Int. Food Res. J. 18 1427-35.
[7] [Depkes] Departemen Kesehatan 2008 Farmakope Herbal Indonesia ed 1 Jakarta (ID): Depkes
[8] [AOAC] Association of Official Analytical Chemistry 2005 Official Method of Analysis (Arlington: The Association of Official Analytical Chemist. Inc.)
[9] Santosoj J, Anwariyah S, Runiantin RO, Putri AP, Ukhty N, and Stark YY 2012 J. Coastal Dev. 15 189-96
[10] Kunle OF, Egharevba HO, and Ahmadu PO 2012 Int. J. Biodivers. Conserv. 4 101-12
[11] Emilan T, Kurnia A, Utami B, Diyani LN, and Maulana A 2011 Konsep Herbal Indonesia: Pemastian Mutu Produk Herbal [bibliografi] (Depok: Universitas Indonesia)
[12] Akowuah GA, Zhari I, Norhayati I, Sadikun A, and Khamsah SM 2004 Food Chem. 87 559-66
[13] Hossain MA and Ismail Z 2012 Arab J. Chem. 2 1–4
[14] Sembiring BS, Rizal M, and Suhirman S 2012 Budidaya dan Pascapanen Kumis Kucing (Orthosiphon stamineus Benth) (Bogor: Balai Penelitian dan Pengembangan Pertanian, Pusat Penelitian dan Perkebunan, Balittro)

[15] Pavia DL, Lampman GM, and Kriz GZ 2001 Introduction to Spectroscopy (Washington (US): Thomson Learning Inc.)

Acknowledgement
This research was carried out with the aid of a research grant from SINAS of Ministry of Research, Technology and Higher Education.