

## Article

# An HPLC-DAD Method to Quantify Flavonoids in *Sonchus arvensis* and Able to Classify the Plant Parts and Their Geographical Area through Principal Component Analysis

Rifki Husnul Khuluk <sup>1</sup>, Amalia Yunita <sup>1</sup>, Eti Rohaeti <sup>1,2</sup>, Utami Dyah Syafitri <sup>2,3</sup>, Roza Linda <sup>4</sup>, Lee Wah Lim <sup>5</sup>, Toyohide Takeuchi <sup>5</sup> and Mohamad Rafi <sup>1,2,\*</sup>

- <sup>1</sup> Departement of Chemistry, Faculty of Mathematics and Natural Science, IPB University, Jalan Tanjung Kampus IPB Dramaga, Bogor 16680, Indonesia; husnulkhuluk08@gmail.com (R.H.K.); ayuamaliaaa05@gmail.com (A.Y.); etirohaeti@apps.ipb.ac.id (E.R.)
- <sup>2</sup> Tropical Biopharmaca Research Center, Research and Community Empowerment Institute, IPB University, Jalan Taman Kencana No. 3 Kampus IPB Taman Kencana, Bogor 16128, Indonesia; utamids@apps.ipb.ac.id
- <sup>3</sup> Department of Statistics, Faculty of Mathematics and Natural Science, IPB University, Jalan Meranti Kampus IPB Dramaga, Bogor 16680, Indonesia
- <sup>4</sup> Department of Chemistry Education, Faculty of Education, Riau University, Jalan Pekanbaru-Bangkinang KM 12.5 Kampus Bina Widya, Pekanbaru 28293, Indonesia; rozalinda@gmail.com
- <sup>5</sup> Department of Chemistry and Biomolecular Science, Faculty of Engineering, Gifu University, 1-1 Yanagido, Gifu 501-1193, Japan; lim@gifu-u.ac.jp (L.W.L.); take-t@gifu-u.ac.jp (T.T.)
- \* Correspondence: mra@apps.ipb.ac.id; Tel.: +62-2518624567



check for updates

**Citation:** Khuluk, R.H.; Yunita, A.; Rohaeti, E.; Syafitri, U.D.; Linda, R.; Lim, L.W.; Takeuchi, T.; Rafi, M. An HPLC-DAD Method to Quantify Flavonoids in *Sonchus arvensis* and Able to Classify the Plant Parts and Their Geographical Area through Principal Component Analysis. *Separations* **2021**, *8*, 12. <https://doi.org/10.3390/separations8020012>

Received: 27 December 2020

Accepted: 20 January 2021

Published: 23 January 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** A simple and efficient method has been developed for the simultaneous determination of eight flavonoids (orientin, hyperoside, rutin, myricetin, luteolin, quercetin, kaempferol, and apigenin) in *Sonchus arvensis* by high-performance liquid chromatography diode array detector (HPLC-DAD). This method was utilized to differentiate *S. arvensis* samples based on the plant parts (leaves, stems, and roots) and the plant's geographical origin. The chromatographic separation was carried out on a reverse-phase C18 column by eluting at a flow rate of 1 mL/min using a gradient with methanol and 0.2% aqueous formic acid. In the optimum conditions, the developed method's system suitability has met the criteria of good separation. The calibration curve shows a linear relationship between the peak area and analyte concentration with a correlation coefficient ( $r^2$ ) > 0.9990. The ranges for the analytes' limits of detection and quantitation were 0.006–0.015 and 0.020–0.052 µg/mL, respectively. Intra-day and inter-day precision expressed in terms of RSD values were <2%, and the accuracy range based on recovery was 97–105%. The stability of all analytes within 48 h was about 2%. By combining HPLC-DAD fingerprint analysis with chemometrics, the developed method can classify *S. arvensis* samples based on the plant parts and geographical origin.

**Keywords:** chemometrics; flavonoids; HPLC-DAD; PCA; *Sonchus arvensis*

## 1. Introduction

Indonesia is known as one of the countries with the highest biodiversity level. The organisms found in it include food and medicinal plants, which can be used for dietary supplements and herbal medicines to prevent or treat diseases. *Sonchus arvensis*, known as perennial sowthistle and *tempuyung* in Indonesia, has been used in commercial functional drinks and herbal medicines. In European countries, *S. arvensis* leaves have been used as a salad [1]. This plant is also used in traditional medicine to treat inflammation, coughing, bronchitis, gallstones, and asthma [2]. For medicinal purposes, the leaves are commonly used [3]. *S. arvensis* tissues are rich in flavonoids like kaempferol, luteolin-7-O-glucoside, and apigenin-7-O-glucoside. They also contain coumarin (scopoletin), taraxasterol, inositol, and phenolic acids (e.g., cinnamic acid, coumaric acid, and vanillic acid) [4]. Besides that, *Sonchus arvensis* also have essential oil [5], sesquiterpenes [6], and quinic acid esters [7].

The mentioned flavonoids and other secondary metabolites found in *S. arvensis* have an essential role in its biological activity.

Flavonoids are a large group of plant polyphenols found in many medicinal plants, vegetables, fruits, and crop products, mainly in flowers and leaves. These compounds play an important role in the human diet and have a beneficial effect on human health, showing antioxidants [8], anticancer [9], anti-inflammatory, antiviral, and antimicrobial [10] activities. Considering the importance of these biological activities of flavonoids, the identification and quantification of these compounds in plant tissues must be carried out by employing methods that can be used to certify the plant's quality and safety containing the mentioned flavonoids.

Several analytical methods have been developed to identify flavonoids in some medicinal plants. They are based on thin-layer chromatography, high-performance liquid chromatography (HPLC) with ultraviolet or fluorescence detectors [11], UV-Vis spectrophotometry [12], and capillary electrophoresis [13]. High-performance liquid chromatography (HPLC) is one of the most common and efficient chromatographic methods used to separate flavonoids from each other. In particular, silica-based stationary phases are the most popular and commonly used in HPLC procedures to achieve the separation of flavonoids. This stationary phase is widely used, providing high separation efficiency and excellent mechanical stability. In addition, silica stationary phases modified by several types of bound phases using C4–C30 alkyl groups, such as octadecyl or C18, have been widely used in HPLC [14]. The stationary phase is used to produce more efficient separation results in combination with good selectivity.

In recent years, many studies have been published on the separation of the flavonoids present in medicinal plants, including studies conducted on *Lycium barbarum* [15], *Nelumbo nucifera* [16], and *Scutellaria* species [17]. In general, the flavonoids present in these plants can be separated well by HPLC by employing different compositions of the mobile phase. Khan [18] reported the separation of flavonoids like rutin, kaempferol, quercetin, myricetin, and hyperoside extracted from *S. arvensis*. Seal [19] only detected the presence of catechin, quercetin, and kaempferol in *S. arvensis*. On the other hand, Hua and Qin [20] reported the presence of luteolin and apigenin in *S. arvensis*. Both luteolin and apigenin compounds are typical compounds in *S. arvensis* [21]. Therefore, these two flavonoids should be quantified in *S. arvensis*. Furthermore, no analytical approach has been reported, enabling users to discriminate between plant parts to obtain the sample or predict the *S. arvensis* sample's geographical origin.

This study aimed to develop a method for the simultaneous identification and quantification of flavonoids in *S. arvensis*, namely orientin, hyperoside, rutin, myricetin, luteolin, quercetin, kaempferol, and apigenin. We also applied the developed method for discrimination of geographical origin and plant parts used from *S. arvensis*. Thus, in the present study, a combination of chromatographic fingerprint and chemometric analysis was used as a powerful tool for discriminating between plant parts used to prepare the samples and identify the *S. arvensis* sample's geographical origin.

## 2. Materials and Methods

### 2.1. Chemicals and Reagents

All solvents were obtained from Kanto Chemicals (Tokyo, Japan) in an analytical or HPLC grade. These include methanol, hydrochloric acid 6M (HCl), tertiary butyl hydroxyquinoline (TBHQ) solutions (prepare by dissolving 0.05 g TBHQ in methanol 62.5%), and 0.2% formic acid in distilled water. Purified water was produced in the laboratory using a water distillation system GS-590 device (Advantec, Tokyo, Japan). All pure standard solutions used in this work were prepared with methanol. The standard solutions used in this work are apigenin (purity  $\geq 95.0\%$ ), myricetin (purity  $\geq 96.0\%$ ), kaempferol (purity  $\geq 97.0\%$ ), luteolin (purity  $\geq 97.0\%$ ), rutin (purity  $\geq 94.0\%$ ), and quercetin (purity  $\geq 95.0\%$ ), which were obtained from Cayman Chemical (Ann, Arbor, MI, USA), whereas

orientin (purity  $\geq 97.0\%$ ) and hyperoside (purity  $\geq 97.0\%$ ) were obtained from Toronto Research Chemical (Canada).

## 2.2. Plant Materials

Plant samples were collected from two regions (May 2018), characterized by low- and high-altitude. The low-altitude samples were collected in Dramaga, Bogor, West Java, whereas the high-altitude samples were collected in Lembang, West Bandung, West Java, Indonesia. We separated and investigated *S. arvensis* roots, stems, and leaves samples. All samples were cut into small pieces, dried at room temperature, and milled to a 60-mesh size. Mr. Taopik Ridwan performed the sample's identification, and the voucher specimen was stored at the Tropical Biopharmaca Research Center of IPB University, Indonesia.

## 2.3. Sample Preparation and Standard Solutions

Accurately weighed powdered samples (1 g) were sonicated in the presence of HCl 6M (5 mL) and TBHQ solutions (20 mL) for one hour at room temperature. The sample extracts obtained following sonication were filtered through a 0.45- $\mu\text{m}$  membrane filter before being injected into the HPLC-DAD device. Standard stock solutions of apigenin, kaempferol, luteolin, myricetin, rutin, orientin, hyperoside, and quercetin were prepared in methanol at 1000  $\mu\text{g}/\text{mL}$  concentrations. An appropriate amount of each standard stock solution was mixed and diluted with methanol to obtain eight concentrations of the working standard solutions of the eight analytes for constructing the relevant calibration curves.

## 2.4. Chromatography Conditions

An HPLC LC-20 AD equipped with a diode array detector (Shimadzu, Kyoto, Japan) and Zorbax Eclipse Plus C18 column (150 mm  $\times$  4.6 mm i.d., 5- $\mu\text{m}$  particle size) was used to separate the compounds. The mobile phase consisted of methanol (A) and 0.2% formic acid in water (B), with the following gradient elution program: 15–30% A at 0–10 min, 30–50% A at 10–45 min, 50–80% A at 45–50 min, 80–95% A at 50–55 min, and 100% A at 55–60 min. After each run, the chromatographic system is set to 0% B for 5 min and balanced for 5 min with a flow rate of 1 mL/min. A diode-array detector monitors the column eluate. The quantification of flavonoids is based on a standard comparison of 270 nm for rutin and hyperoside, 330 nm for kaempferol, and 340 nm for orientin, myricetin, luteolin, quercetin, and apigenin. We used an injection volume of standard and sample solutions of about 10  $\mu\text{L}$ .

## 2.5. Analytical Performance

The analytical performance of our approach was evaluated by following the guidelines of the International Conference on Harmonization (ICH) [22] by determining system suitability, curve calibration linearity, the limit of detection (LOD), the limit of quantitation (LOQ), precision, accuracy, and stability, expressed as a relative standard derivation (RSD) from each of the analyte levels ( $\mu\text{g}/\text{g}$ ) obtained from the characteristic chromatogram peaks. Notably, the samples utilized to determine the analytical performance were *S. arvensis* leaves.

## 2.6. Data Analysis

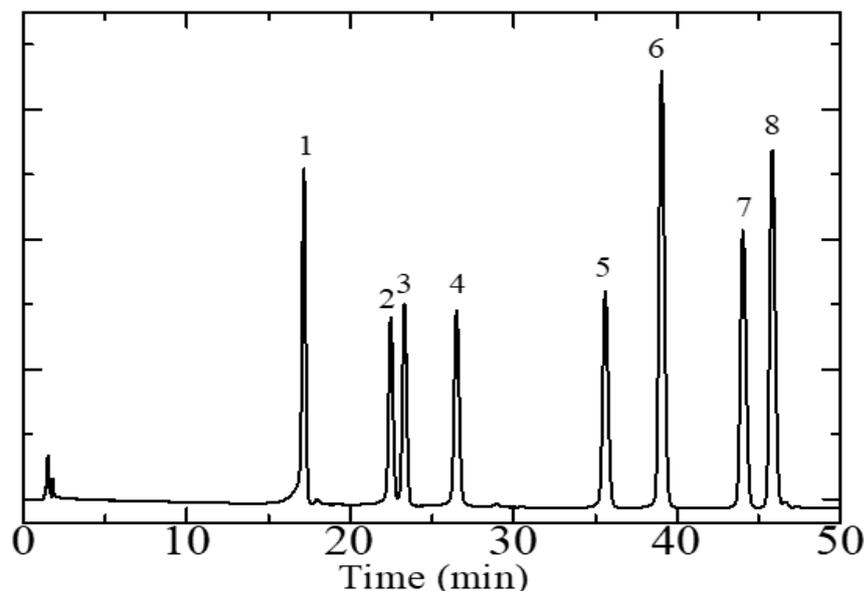
Principal component analysis (PCA) was used to classify the different plant parts and the geographical origin of *S. arvensis* samples. For this purpose, we used HPLC-DAD chromatogram data of *S. arvensis*. Before being subjected to the PCA, the chromatogram was aligned using correlation optimized warping (COW). PCA and COW were performed using the Unscrambler X software version 10.1 from Camo (Oslo, Norway).

# 3. Results and Discussion

## 3.1. Optimization of HPLC-DAD Conditions

In order to meet the best separation condition for quantitative analysis and a chromatographic fingerprint, the effect of the composition of the mobile phase was investigated,

and the detection wavelength was optimized. The resolution of each analyte and the total analysis time were the parameters used to optimize the chromatographic conditions. Representative chromatograms of the standard solutions of flavonoids are reported in Figure 1. The optimal chromatographic separation was achieved using a gradient elution of phase A, methanol, and an elution of phase B, 0.2% formic acid in the water, as described in the Materials and Methods section.



**Figure 1.** High-performance liquid chromatography chromatogram of a standard 10 µg/mL flavonoid solution obtained using a Zorbax Eclipse Plus C18 column. Mobile phase composition program: 0–10 min, 15–30% A, 10–45 min, 30–50% A, 45–50 min, 50–80% A, 50–55 min, 80–95% A, 55–60 min, 100% A. Flowrate: 1 mL/min. Analyte-associated peaks: 1, orientin, 2, hyperoside, 3, rutin, 4, myricetin, 5, luteolin, 6, quercetin, 7, kaempferol, 8, apigenin. Sample injection volume: 10 µL. Ultraviolet (UV) wavelength for analyte detection: 340 nm. Phase A: methanol. Phase B: 0.2% formic acid in water.

A satisfactory separation was achieved for all analytes with a separation factor of each peak being greater than 1.5 and the total time to perform the quantitative flavonoid analysis is 60 min. The wavelengths chosen for analytes quantification (corresponding to the absorption maxima of the analytes) were 270 nm in the case of rutin and hyperoside, 330 nm for kaempferol, and 340 nm for orientin, myricetin, luteolin, quercetin, and apigenin. These wavelengths were chosen because they afforded high sensitivity and signal intensities concerning the alternatives for each separated compound.

### 3.2. Evaluation of the Analytical Performance of the Developed Method

An analytical performance test is a process aimed at proving that a given analytical procedure provides satisfactory results in terms of its stated intent and purpose. This evaluation can provide information on the goodness, quality, reliability, and consistency of the analytical results. The parameters defining the outcome of the analytic performance test include system suitability, linearity, LOD, LOQ, accuracy, and stability.

System suitability was assessed, performing five replicate analyses of a mixed standard solution of orientin, hyperoside, rutin, myricetin, luteolin, quercetin, kaempferol, and apigenin. The relative standard deviation (RSD) values of the retention time, peak area, capacity factor, tailing factor, and theoretical plate number were calculated to evaluate the system suitability (Table 1). The RSD values for all the parameters quantified in the system suitability tests were found to be less than 2%, which indicates that this method is very well suited for achieving the target separation of flavonoids.

**Table 1.** Result of the system suitability test performed on the analytical method employed for the separation and quantification of flavonoids in a mixed standard solution of orientin (OR), hyperoside (HP), rutin (RT), myricetin (MR), luteolin (LT), quercetin (QR), kaempferol (KM), and apigenin (AG).

Analyte <sup>a</sup>	OR	HP	RT	MR	LT	QR	KM	AG
Retention time (min)								
Mean	16.912	21.952	22.996	26.224	35.057	38.451	43.334	45.136
RSD (%)	0.03	0.21	0.05	0.24	0.03	0.03	0.1	0.06
Peak area								
Mean	276,087	1,769,818	275,124	246,515	247,129	454,891	331,047	373,706
RSD (%)	0.31	0.42	0.22	0.34	0.62	0.41	0.13	0.46
Capacity factor								
Mean	10.956	14.684	15.258	14.453	23.779	26.197	39.685	32.605
RSD (%)	0.14	0.17	0.07	0.44	0.12	0.04	0.13	0.51
Tailing factor								
Mean	0.819	1.036	1.014	1.141	0.972	0.934	1.291	1.125
RSD (%)	0.65	0.75	0.63	0.93	0.56	0.75	0.42	0.36
Theoretical plate number								
Mean	21,383	23,821	27,756	27,336	42,393	52,963	63,706	73,598
RSD (%)	0.38	0.32	1.83	1.92	1.22	2.17	1.78	2.01

<sup>a</sup> The concentration of each analyte was 10 µg/mL (n = 5). RSD: relative standard deviation.

The calibration curves were performed at eight concentration levels in the 0.078–10 µg/mL range with three replicate injections for each concentration level by plotting the peak areas versus each analyte’s concentration (Table 2). The linearity of the calibration curves was determined based on the value of the correlation coefficient (r<sup>2</sup>). Good linearity was obtained with a mean correlation coefficient value greater than 0.9995 for all analytes within the concentration levels. The LOD (S/N = 3.3) and LOQ (S/N = 10) were observed in the range of concentration levels of the calibration curve between 0.006–0.015 and 0.020–0.052 µg/mL, respectively. These results indicated that the proposed analytical method is characterized by good sensitivity.

**Table 2.** Analytical data of calibration curves.

Analyte	Regression Equation <sup>a</sup>	Correlation Coefficient (r <sup>2</sup> )	Standard Deviation	
			Intercept	Slope
OR	y = 27480x + 171.66	0.9991	193.74	264.86
HP	y = 35146x + 924.74	0.9998	249.50	135.91
RT	y = 17497x + 445.46	0.9996	547.45	197.55
MR	y = 24911x + 113.63	0.9994	304.48	94.98
LT	y = 25092x + 28.296	0.9995	113.34	217.54
QR	y = 45521x + 750.85	0.9997	184.52	746.81
KM	y = 33302x + 43.674	0.9996	19.43	86.88
AG	y = 36866x + 863.5	0.9992	128.96	166.70

<sup>a</sup> Concentration range for all analytes for calibration curve: 0.078–10 µg/mL. AG: apigenin. HP: hyperoside. KM: kaempferol. LT: luteolin. MR: myricetin. OR: orientin. QR: quercetin. RT: rutin. SD: standard deviation.

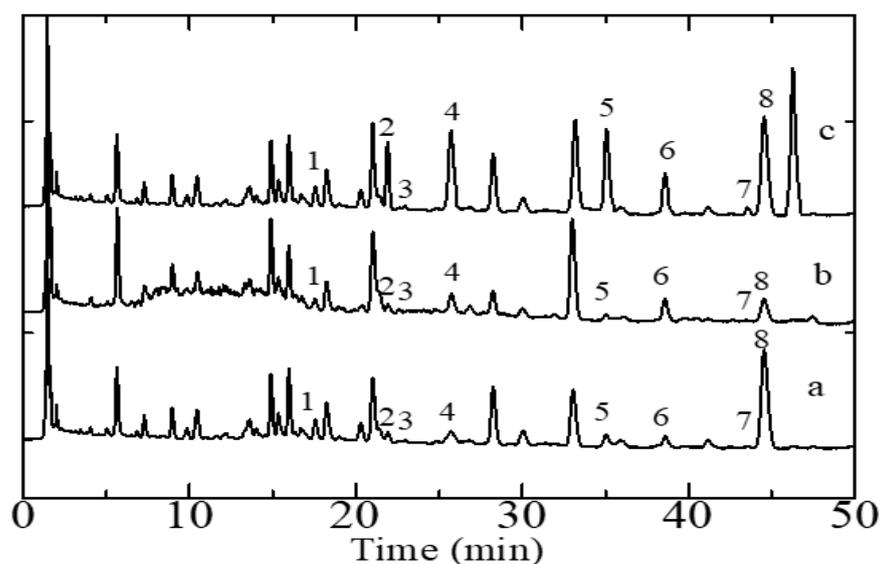
The precision of the method was determined by measuring the intra-day and inter-day repeatability of six individual sample analyses performed every day over three consecutive days. The method precision was expressed in terms of RSD, and the RSD values obtained for intra-day and inter-day analysis results were found to be less than 3%, confirming the good repeatability of the method. The accuracy of the method was evaluated by carrying out a recovery test that consisted of spiking known amounts of standard solutions of flavonoids into a test sample with three different levels of added standard solutions and

triplicate measurements at each level. The average percentages of recovery for all analytes were observed in the 97–105% range, with relevant RSD values below 4%. These results demonstrated that the developed analytical method was reliable and accurate.

The analytes stability in the sample solution was evaluated by performing analytes of the sample solution after 0, 4, 8, 12, 24, and 48 h of incubation at room temperature. The analytes were stable in the sample solution with RSD values ranging between 0.7% and 2.0% for all compounds. Analytical data for the LOD, LOQ, precision, accuracy, and stability are reported in Table 2.

### 3.3. Determination of Flavonoid Content in *S. arvensis*

The HPLC-DAD-based method for quantitative flavonoid analysis described above was applied for the simultaneous determination of flavonoids in six samples of *S. arvensis* consisting of several parts of the plant, namely two samples of roots, stems, and leaves from two plants collected in geographically distinct areas. Each sample analysis consisted of five replicate measurements of the amount of each analyte. Figure 2 show representative chromatograms of *S. arvensis* extracts whereby the peaks associated with specific target analytes are labeled as such. It can be seen that the eight target analytes contained in *S. arvensis* are well separated. The peaks of the flavonoid compounds were identified by two methods. The first method includes comparing the retention times of the peak samples with those of the reference compounds eluted under the same conditions and the second method includes spiking the sample with a standard solution of analytes.



**Figure 2.** Representative HPLC-DAD chromatograms of *Sonchus arvensis* extracts. (a) Root sample, (b) stem sample, and (c) leaf sample. Analyte-associated peaks: 1, orientin, 2, hyperoside, 3, rutin, 4, myricetin, 5, luteolin, 6, quercetin, 7, kaempferol, 8, apigenin. Operating conditions are the same as those detailed in the legend of Figure 1.

In addition to the eight peaks that have identified flavonoids, from the results obtained, there are still other peaks that have not been identified. The peak is likely another compound, such as other flavonoids. Then, from the results of these measurements, we analyzed the content of eight flavonoids using a calibration curve from the optimum condition of HPLC-DAD to see the sample difference based on the concentration levels obtained (Table 3).

**Table 3.** Analytical data of LOD, LOQ, precision, accuracy, and analyte stability for determination of eight flavonoids.

Analyte	Precision (RSD, %)		Accuracy <sup>a</sup>		Stability <sup>b</sup> (n = 6)
	Intra-Day (n = 6)	Inter-Day (n = 3)	Average Recovery (%)	RSD (%) (n = 3)	
OR	Day 1: 2.23 Day 2: 1.41 Day 3: 1.54	1.32	97.84	2.8	1.6
HP	Day 1: 0.94 Day 2: 1.07 Day 3: 0.92	0.93	103.6	2.31	1.38
RT	Day 1: 1.60 Day 2: 1.12 Day 3: 1.03	1.44	97.79	2.42	1.03
MR	Day 1: 0.90 Day 2: 1.00 Day 3: 0.96	0.81	104.5	1.34	1.47
LT	Day 1: 1.09 Day 2: 1.28 Day 3: 1.18	0.99	105.87	1.74	1.08
QR	Day 1: 0.87 Day 2: 0.76 Day 3: 1.89	0.87	99.22	1.39	1.15
KM	Day 1: 0.51 Day 2: 0.61 Day 3: 0.84	0.91	97.17	2.07	1.29
AG	Day 1: 0.69 Day 2: 0.98 Day 3: 0.65	0.75	105.1	0.89	0.99

<sup>a</sup> Three levels of added standard compounds (0.078, 0.625, 10 µg/mL) in the sample solution with each level measured in triplicate. <sup>b</sup> Five measurements performed 0, 4, 8, 12, and 24 h after sample preparation by extraction. AG: apigenin. HP: hyperoside. KM: kaempferol. LT: luteolin. MR: myricetin. OR: orientin. QR: quercetin. RT: rutin. RSD: relative standard deviation.

As can be seen in Table 4, flavonoid levels did not differ significantly due to either the geographical origin or type (roots versus stems versus leaves) of the sample. The amount of orientin, hyperoside, rutin, myricetin, luteolin, quercetin, kaempferol, and apigenin observed to be present across all samples fell in the following ranges, respectively: 16.36–24.91, 1.34–39.34, 1.73–46.59, 18.00–313.97, 4.03–91.87, 2.27–31.64, 0.91–1.27, and 123.22–272.34 µg/g. Apigenin was found to be the dominant compound concentration-wise, whereas kaempferol was the flavonoid characterized by the lowest concentration in most of the samples investigated. Evidence indicated that the concentrations of flavonoids present in *S. arvensis* samples originating from the Bogor area were generally higher than those observed in West Bandung samples.

The flavonoids were also found in very high concentrations in leaves samples, whereas their concentrations were very low in the plant stem samples. Based on these results, we could not discriminate them only on the amount of these eight bioactive compounds. For this purpose, therefore, we relied on the chemometric method, which is widely employed to discriminate between the different plant parts and the plant samples' geographical origin based on the samples' contents of suitable chemical compounds.

**Table 4.** Concentration of eight flavonoids detected in *Sonchus arvensis* samples.

Analyte	Geographical Origin of Sample					
	Bogor			West Bandung		
	Sampel Type					
	Root	Stem	Leaves	Root	Stem	Leaves
Concentration ( $\mu\text{g/g} \pm \text{SD}$ ), $n = 5$						
OR	22.70 $\pm$ 0.35	17.89 $\pm$ 0.10	24.91 $\pm$ 0.35	17.80 $\pm$ 0.35	16.99 $\pm$ 0.35	16.36 $\pm$ 0.45
HP	83.40 $\pm$ 0.15	1.83 $\pm$ 0.22	39.34 $\pm$ 0.19	1.38 $\pm$ 0.22	1.34 $\pm$ 0.25	18.64 $\pm$ 0.18
RT	46.59 $\pm$ 1.03	1.75 $\pm$ 0.03	40.01 $\pm$ 0.30	2.72 $\pm$ 0.04	1.73 $\pm$ 0.01	22.03 $\pm$ 0.15
MR	35.55 $\pm$ 0.37	27.63 $\pm$ 0.42	43.07 $\pm$ 0.22	18.00 $\pm$ 0.28	20.57 $\pm$ 0.26	313.97 $\pm$ 0.15
LT	91.87 $\pm$ 2.03	82.65 $\pm$ 1.52	12.57 $\pm$ 0.19	12.79 $\pm$ 0.10	4.03 $\pm$ 0.06	32.68 $\pm$ 1.67
QR	9.77 $\pm$ 0.21	15.88 $\pm$ 0.23	31.64 $\pm$ 0.22	2.27 $\pm$ 0.03	12.41 $\pm$ 0.18	21.60 $\pm$ 0.31
KM	0.91 $\pm$ 0.52	1.10 $\pm$ 0.67	1.24 $\pm$ 0.71	0.97 $\pm$ 0.81	1.19 $\pm$ 0.51	1.27 $\pm$ 0.22
AG	157.22 $\pm$ 2.02	234.58 $\pm$ 1.57	255.39 $\pm$ 1.26	272.34 $\pm$ 1.76	123.22 $\pm$ 1.31	206.65 $\pm$ 1.16

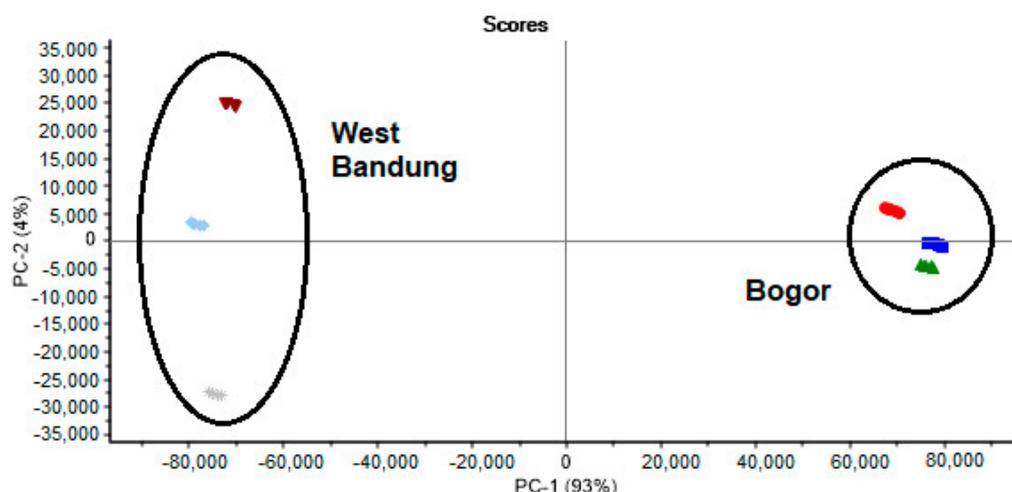
AG: apigenin. HP: hyperoside. KM: kaempferol. LT: luteolin. MR: myricetin. OR: orientin. QR: quercetin. RT: rutin. SD: standard deviation.

### 3.4. Clustering of *S. arvensis* Samples from the Different Geographical Origin and Plant Parts

In developing a method for distinguishing between *S. arvensis* samples based on the plant part they originated from and on the geographical origin of the plant, we employed a combination of HPLC-DAD fingerprint analysis and the chemometric method. This combination has become one of the most used approaches to the classification, authentication, and discrimination of medicinal plants to recognize the plant’s geographical origin, the detection of adulteration, and the discrimination of closely related species. One of the multivariate analysis techniques we used in this study was PCA, which is a well-known unattended pattern recognition method. PCA is useful for reducing data and extracting information to find combinations of variables to describe the data set [23]. The original variable is transformed into a new and uncorrelated variable called the principal component (PC) [24].

As a variable, we used the HPLC-DAD fingerprint chromatogram data and pre-processed first by alignment using correlation optimized warping (COW). COW is a transformation to align chromatogram data that has an x-axis misalignment between one sample and another. The COW method works by aligning the sample chromatogram to the reference chromatogram using dynamic programming by stretching or compressing the sample segments using linear interpolation [25,26]. Segment length and slack size are two parameters optimized by the COW automated algorithm [27]. The segment length parameter determines the number of segments, while the slack parameter controls how these segments can be expanded or compressed [26]. When the number of time points in the sample and reference segment is different, the first to be linearly interpolated to make the segment length the same is the slack parameter [28]. The optimal slack size and segment length obtained were 70 and 14, respectively.

In this study, PCA was used to group samples based on differences associated with geographical origin. The score plot for the first two principal components (PC1 and PC2) is usually used most in the analysis because these two PCs contain the most data variations. The results of the score plots used the first two PCs originating from PCA. These components accounted for 97% of the total variance (PC1 = 93% and PC2 = 4%) (Figure 3). The PCA score plot explains that all samples can be classified into their respective groups based on their geographic origin and plant part. In this case, based on the results described above, the HPLC-DAD fingerprint chromatogram data combined with PCA can be reliably used to infer what plant part was utilized to prepare the sample and the geographical origin of the *S. arvensis* sample.



**Figure 3.** Principal component analysis (PCA) score plot of *S. arvensis* extract: root (■), stem (●), and leaves (▲) from Bogor and root (◆), stem (▼), and leaves (⊕) from West Bandung.

#### 4. Conclusions

In this study, the simultaneous quantification of eight flavonoids in different *S. arvensis* samples originating from different plant parts and geographical areas was achieved by employing an HPLC-DAD-based methodology, which was successfully developed a high-accuracy and high-reliability qualitative and quantitative analysis of the mentioned compounds. Evidence indicated that flavonoid content in the leaves of *S. arvensis* was higher than in the plant's roots and stem. A combination of HPLC-DAD fingerprint analysis and the chemometric method enabled us to group the samples into six groups successfully. Therefore, the developed method could be used for the quality control of *S. arvensis*.

**Author Contributions:** Conceptualization, R.H.K., E.R., R.L., L.W.L., T.T., and M.R. Methodology, R.H.K., E.R., U.D.S., R.L., L.W.L., T.T., and M.R. Software, M.R. Validation, R.H.K., E.R., and M.R. Formal analysis, R.H.K., A.Y., and U.D.S. Investigation, R.H.K. and A.Y. Resources, E.R., L.W.L., T.T., and M.R. Data curation, R.H.K. Writing—original draft preparation, R.H.K. and A.Y. Writing—review and editing, E.R., U.D.S., R.L., L.W.L., T.T., and M.R. Visualization, R.H.K. and A.Y. Supervision, E.R., U.D.S., R.L., L.W.L., T.T., and M.R. Project administration, E.R. Funding acquisition, E.R. and M.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Directorate of Research and Community Service, Ministry of Research and Technology/National Agency for Research and Innovation, the Republic of Indonesia, *Penelitian Dasar Unggulan Perguruan Tinggi*, grant No. 4016/IT3.L1/PN/2020.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable in this article.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### References

1. Holm, L.; Doll, J.; Holm, E.; Pancho, J.; Herberger, J. *World Weeds: Natural Histories and Distribution*; John Wiley and Sons Inc.: New York, NY, USA, 1997.
2. Li, X.M.; Yang, P.L. Research progress of *Sonchus* species. *Int. J. Food Prop.* **2018**, *21*, 147–157. [[CrossRef](#)]
3. Ministry of Health Republic of Indonesia. *Indonesia Herbal Pharmacopoeia*, 1st ed.; Ministry of Health Republic of Indonesia: Jakarta, Indonesia, 2010; pp. 138–140.
4. Manoi, F. Effect fineness extraction of materials and old quality extract (*Sonchus arvensis* L.). *J. Appl. Agric. Res.* **2015**, *15*, 156–161.

5. Kanaani, S.; Sani, A.M. Chemical composition of essential oils and in vitro antibacterial activity of methanolic extract of *Sonchus arvensis* and *Eremurus spectabilis* against food-borne pathogenic bacteria. *J. Essent Oil-Bear Plants* **2015**, *18*, 1093–1099. [[CrossRef](#)]
6. Xia, Z.; Yao, J.; Liang, J. Two new sesquiterpene lactones from *Sonchus arvensis*. *Chem. Nat. Compd.* **2012**, *48*, 47–50. [[CrossRef](#)]
7. Xu, Y.J.; Sun, S.B.; Sun, L.M.; Qiu, D.F.; Liu, X.J.; Jiang, Z.B.; Yuan, C.S. Quinic acid esters and sesquiterpenes from *Sonchus arvensis*. *Food Chem.* **2008**, *111*, 92–97. [[CrossRef](#)]
8. Karamian, K.; Asadbegy, M. Antioxidant activity, total phenolic and flavonoid contents of three *Onobrychis* species from Iran. *J. Pharm. Sci.* **2016**, *22*, 112–119. [[CrossRef](#)]
9. Raffa, D.; Maggio, B.; Raimondi, M.V.; Plescia, F.; Daidone, G. Recent discoveries of anticancer flavonoids. *Eur. J. Med. Chem.* **2017**, *142*, 213–228. [[CrossRef](#)] [[PubMed](#)]
10. Wang, T.Y.; Li, Q.; Bi, K.S. Bioactive flavonoids in medicinal plants: Structure, activity and biological fate. *Asian J. Pharm. Sci.* **2018**, *13*, 12–23. [[CrossRef](#)] [[PubMed](#)]
11. Hajimehdipour, H.; Kondari, B.M.; Amin, G.R.; Adib, N.; Rastegar, H.; Shekarchi, M. Development of a validated HPLC method for the simultaneous determination of flavonoid in *Cuscuta chinensis* L. by ultra-violet detection. *Daru J. Pharm. Sci.* **2012**, *20*, 57–62. [[CrossRef](#)] [[PubMed](#)]
12. Ramos, R.T.M.; Bezerra, I.C.F.; Ferreira, M.R.A.; Soares, L.A.L. Spectrophotometric quantification of flavonoids in herbal material, crude extract, and fractions from leaves of *Eugenia uniflora* Linn. *Pharm. Res.* **2017**, *9*, 253–260.
13. Wang, W.; Lin, P.; Ma, L.; Xu, K.; Lin, X. Separation and determination of flavonoids in three traditional Chinese medicines by capillary electrophoresis with amperometric detection. *J. Sep. Sci.* **2016**, *39*, 1357–1362. [[CrossRef](#)] [[PubMed](#)]
14. Zhao, L.; Liu, L.; Li, J. Qualitative and quantitative analysis of five bioactive flavonoids in *Salix bordensis* turcz. by HPLC-DAD and HPLC-ESI-MS. *Am. J. Anal. Chem.* **2014**, *5*, 851–860. [[CrossRef](#)]
15. Magiera, S.; Zareba, M. Chromatographic determination of phenolic acids and flavonoids in *Lycium barbarum* L. and evaluation of antioxidant activity. *Food Anal. Methods* **2015**, *8*, 2665–2674. [[CrossRef](#)]
16. Naveen, P.; Lingaraju, H.B.; Anitha; Prasad, K.S. Simultaneous determination of rutin, isoquercetin, and quercetin flavonoids in *Nelumbo nucifera* by high-performance liquid chromatography method. *Int. J. Pharm. Investig.* **2017**, *7*, 94–100. [[CrossRef](#)] [[PubMed](#)]
17. Gharari, Z.; Bagheri, K.; Danafar, H.; Sharafi, A. Simultaneous determination of baicalein, chrysin and wogonin in four Iranian *Scutellaria* species by high performance liquid chromatography. *J. Appl. Res. Med. Aromat. Plants* **2019**, *16*, 100232. [[CrossRef](#)]
18. Khan, R.A. Evaluation of flavonoids and diverse antioxidant activities of *Sonchus arvensis*. *Chem. Cent. J.* **2012**, *6*, 126–132. [[CrossRef](#)]
19. Seal, T. Quantitative HPLC analysis of phenolic acids, flavonoids and ascorbic acid in four different solvent extracts of two wild edible leaves, *Sonchus arvensis* and *Oenanthe linearis* of North-Eastern region in India. *J. Appl. Pharm. Sci.* **2016**, *6*, 157–166. [[CrossRef](#)]
20. Huo, B.S.; Qin, M.J. Content analysis of flavonoids in five species of *Sonchus* L. *Zhiwu Ziyuan Yu Huanjing Xuebao* **2008**, *17*, 77–78.
21. BPOM (Badan Pengawasan Obat dan Makanan) RI. Fingerprint of *Sonchus arvensis* Leaf as the Basis of Standardization of Natural Medicinal Plant Extracts. 2010. Available online: [http://www.pom.go.id/ppid/2016/ringkasan\\_riset2010.pdf](http://www.pom.go.id/ppid/2016/ringkasan_riset2010.pdf) (accessed on 14 August 2020).
22. ICH (International Conference on Harmonization). *Validation of Analytical Procedures: Methodology Q2 (R1)*; ICH: Geneva, Switzerland, 2005. Available online: [http://www.ich.org/fileadmin/Public\\_Web\\_Site/ICH\\_Products/Guidelines/Quality/Q2\\_R1/Step4/Q2\\_R1\\_Guideline.pdf](http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q2_R1/Step4/Q2_R1_Guideline.pdf) (accessed on 14 August 2020).
23. Liu, X.; Wu, Z.; Yang, K.; Ding, H.; Wu, Y. Quantitative analysis combined with chromatographic fingerprint for comprehensive evaluation of Danhong injection using HPLC DAD. *J. Pharm. Biomed.* **2013**, *76*, 70–74. [[CrossRef](#)]
24. Xu, X.; Jiang, J.; Liang, Y.; Yia, L.; Cheng, J. Chemical fingerprint analysis for quality control of Fructus Aurantii Immaturus based on HPLC-DAD combined with chemometric methods. *Anal. Methods* **2010**, *2*, 1–9. [[CrossRef](#)]
25. Skov, T.; Berg, F.V.D.; Tomasi, G.; Bro, R. Automated alignment of chromatographic data. *J. Chemom.* **2006**, *20*, 484–497. [[CrossRef](#)]
26. Kumar, K. Chemometric assisted correlation optimized warping of chromatograms: Optimizing the computational time for correcting the drifts in chromatographic peak positions. *Anal. Methods* **2018**, *10*, 1006–1014. [[CrossRef](#)]
27. Jiao, L.; Bing, S.; Wang, X.; Xue, Z.; Li, H. The application of automated correlation optimized warping to the quality evaluation of Radix Puerariae thomsonii: Correcting retention time shift in the chromatographic fingerprints. *Quim. Nova* **2015**, *38*, 8–13.
28. Ni, Y.; Song, R.; Kokot, S. Analysis of HPLC fingerprints: Discrimination of raw and processed Rhubarb samples with the aid of chemometrics. *Anal. Methods* **2012**, *4*, 171–176. [[CrossRef](#)]