FINGERPRINT ANALYSIS OF UNFRACTIONATED Piper PLANT EXTRACTS BY HPLC-UV-DAD COUPLED WITH CHEMOMETRIC METHODS

FREDDY A. BERNAL, WILMAN A. DELGADO, LUIS E. CUCA*

Laboratorio de Investigación en Productos Naturales Vegetales, Departamento de Química, Universidad Nacional de Colombia,
Carrera 30 45-03, Bogotá, Colombia.

(Received: December 19, 2012 - Accepted: May 26, 2012)

ABSTRACT

23 crude ethanolic extracts of Piper plants were analyzed by combining HPLC fingerprints and chemometric methods including principal component analysis (PCA), hierarchical clustering analysis (HCA) and parallel factor analysis (PARAFAC). All statistical analysis were carried out both with and without previous preprocessing by counting with alignment of the fingerprints by correlation optimized warping (COW) and several normalization methods. Normalization processes of autoscaling, range scaling and vast scaling resulted to be the better offering the most relevant information. Both direct two-way methods (PCA and HCA) and HCA after data minimization employing a three-way method (PARAFAC) were demonstrated to be useful in the establishment of chemical composition relationships among Piper samples, and resulting in agreement with phylogenic relationships previously reported in literature to some Piper species.

Keywords: Piper, Piperaceae, fingerprints, HPLC, chemometrics, PCA, pest control potential

INTRODUCTION

Piper, Peperomia, Sarchorhachis, and Ottone are the four major genera found in tropical and sub-tropical regions at the Neotropics, all belonging to Piperaceae family, which are considered as one of the most arcanic pantropical flowering plants. The genus Piper is comprised by 1000 species approximately, so that it is one of the largest genera of basal angiosperms, and its species are located in all types of vegetation, mainly as components of pioneer vegetation. Although economic uses of Piper plants are limited to some Asian and African species (such as fruits of Piper nigrum, which is the source of black pepper, and the roots of P. methysticum as the source of Kava, which is used as traditionally narcotic beverage), it possesses a rich ethnobotanical and ethnomedical history almost incomparable to other plant families suggesting it as an interesting object of study. To date, there is a great number of papers related to Piper chemistry and biological activity, demonstrating that insecticidal potential is prominent and predominant constant to Piper plants.

On the other hand, in the world has been recognized the need to stop the growth of weeds, herbivorous insect pests and plant pathogens by employing chemical substances, and replace it by new solutions such as biopesticides. So, at present botanical products are being accepted as prominent alternatives for pests control. However, the biopotential of crude extracts as pest controllers is determined by chemical composition of the source plant, so that a possibility to search new plant extracts possessing a determined biological activity can be based on comparison of secondary metabolites profiles obtained by high resolution separation techniques (CG, HPLC, etc.) with those to previously verify active extracts, mainly using different statistical tools (principal component analysis, hierarchical clustering, similarity analysis, etc.), which can include analysis of n-way matrices (PARAFAC, Tucker, etc.), depending on acquired data. Therefore, by continuing our research on Colombian higher plants, in this paper are presented chromatographic and chemometric analysis of Piper plant extracts without any pretreatment by multi-wavelength-HPLC (the simplest hyphenated technique to characterization of plant natural products), including a comparison of results to PCA and HCA analysis to integrated HPLC profiles with results of PARAFAC and HCA analysis to HPLC-UV-DAD profiles, both with and without different data preprocessing (alignment and several normalization processes).

MATERIALS AND METHODS

Plant materials

23 samples belonging to genus Piper were collected on 2006 – 2007 (rely on at least 15 species and different plant tissues to some of these plants) at Sumapaz and Guavio provinces (Cundinamarca department, Colombia) and listed in Table 1. All of them were determined by biologist Adolfo Jara Muñoz at the Herbario Nacional Colombiano, Instituto de Ciencias Naturales, Universidad Nacional de Colombia, and a voucher specimen of each was deposited there.

Instrumentation and reagents

All HPLC-DAD analysis were performed using a Merck-Hitachi HPLC system equipped with a L-4500 diode array detector, a L-6200A intelligence pump, and a L-6000A interface, controlled with Merck-Hitachi D-7000 Chromatography Data Station Software, version 4.1. HPLC grade methanol and acetonitrile from Merck (Darmstadt, Germany) were used without previous filtration. Deionized water was purified by a Milli-Q system from Millipore (Bedford, MA, USA). Distilled ethanol was used for extracts preparation.

Sample preparation

All Piper samples were cut into smaller pieces and further ground into powder (only to leaves and inflorescences), and then each sample was exhaustively extracted with 96% ethanol at room temperature during two weeks (conventional maceration process), and solvent was evaporated under reduced pressure to obtain an extract for each sample. 10 mg of these extracts were accurately weighed and re-dissolved with 1 mL MeOH. These solutions were filtered through a 0.45 μm filter (PDVF). An aliquot of 20 μL filtrate was injected for HPLC analysis.

Chromatographic conditions

Separations were performed on a reversed-phase Phenomenex C30 Luna column (150 x 4.6 mm i.d., 5 μm), employing a mobile phase consisting of acetonitrile and water. The gradient program was developed as follows: 0 min, 20% CH3CN; 40 min, 100% CH3CN, 70 min, 100% CH3CN. The flow rate was kept at 0.5 mL/min. The UV spectra were acquired from 220 to 400 nm using 7 nm as spectral bandwidth and 800 ms of spectral interval.

Data analysis

Data were manipulated into two forms: 1) Data were exported as 2D ASCII files at all time scale each 10 nm to whole absorbance range to build a matrix for each sample (matrix contain 5251 x 24 points), i.e., retention time in one direction, and UV spectra in the other direction; 2) Data were exported as ASCII files to integrated chromatogram between 220 to 400 nm (a single profile by sample giving a matrix with 5251 points). All data operations (preprocessing, PCA, HCA, and PARAFAC) were performed using MATLAB R2007b (Mathworks) on a computer Intel Pentium 4 processor containing 500 MB RAM and running Microsoft Windows XP.

Data preprocessing: Raw data (both matrixes) were submitted to chemometric analysis without preprocessing as well as after different pretreatments, as follows:

Chromatographic peak alignment: Correlation optimized warping (COW) algorithm was employed to the alignment process which was performed by using the algorithm supplied by The Quality and Technology Website.

Centering and scaling: centering and scaling pretreatments were applied to the data set by means of algorithms coded by authors (to more information about scaling preprocessing, refer is remitted to van der Berg et al. and literature cited herein).

Principal component analysis (PCA) and Hierarchical clustering analysis (HCA): In this work, PCA and HCA were carried out for the 25 Piper samples employing an integrated chromatogram (from 220 to 400 nm), and results are shown as score plots to each preprocessing method used here. The
PCA were performed using the algorithm included into Statistical Toolbox for Matlab.

**Parallel factor analysis (PARAFAC):** The algorithm supplied by The Quality and Technology Website to perform the PARAFAC model was used here (to get information about PARAFAC model reader can be remitted to references1,2). In order to perform PARAFAC analysis, normalization of integrated chromatogram was used as preprocessing procedure. Prior to analysis, core consistency diagnostic was performed to provide an appropriate number of components. Six components were demonstrated to be able to describe the PARAFAC model. No restrictions on PARAFAC model were taken into account in the present research.

### Table 1. Collection number and geographic distribution for *Piper* species

<table>
<thead>
<tr>
<th>Sample</th>
<th>Part of plant</th>
<th>Species</th>
<th>Origin</th>
<th>Collection Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Leaves</td>
<td><em>Piper eriocladum</em></td>
<td>Granada</td>
<td>COL 517694</td>
</tr>
<tr>
<td>2</td>
<td>Leaves</td>
<td><em>Piper sp.</em></td>
<td>Fusagasugá</td>
<td>COL 519816</td>
</tr>
<tr>
<td>3</td>
<td>Fruits</td>
<td><em>Piper aduncum L.</em></td>
<td>Fusagasugá</td>
<td>COL 515969</td>
</tr>
<tr>
<td>4</td>
<td>Leaves</td>
<td><em>Piper aduncum L.</em></td>
<td>Fusagasugá</td>
<td>COL 515969</td>
</tr>
<tr>
<td>5</td>
<td>Stems</td>
<td><em>Piper aduncum L.</em></td>
<td>Fusagasugá</td>
<td>COL 515969</td>
</tr>
<tr>
<td>6</td>
<td>Leaves</td>
<td><em>Piper amalago L.</em></td>
<td>Nocaima</td>
<td>COL 510519</td>
</tr>
<tr>
<td>7</td>
<td>Aerial</td>
<td><em>Piper arboretum</em></td>
<td>Fusagasugá</td>
<td>COL 519815</td>
</tr>
<tr>
<td>8</td>
<td>Aerial</td>
<td><em>Piper artanthe C.DC.</em></td>
<td>San Bernardo</td>
<td>COL 516759</td>
</tr>
<tr>
<td>9</td>
<td>Fruits</td>
<td><em>Piper bogotense C.DC.</em></td>
<td>Granada</td>
<td>COL 517696</td>
</tr>
<tr>
<td>10</td>
<td>Leaves</td>
<td><em>Piper bogotense C.DC.</em></td>
<td>Granada</td>
<td>COL 517696</td>
</tr>
<tr>
<td>11</td>
<td>Leaves</td>
<td><em>Piper cumanense</em></td>
<td>Fusagasugá</td>
<td>COL 518183</td>
</tr>
<tr>
<td>12</td>
<td>Stems</td>
<td><em>Piper cf. el-bancoanum Trel. &amp; Yunck.</em></td>
<td>Fusagasugá</td>
<td>COL 518182</td>
</tr>
<tr>
<td>13</td>
<td>Fruits</td>
<td><em>Piper cf. eriopodom (Miq.) C.DC.</em></td>
<td>Tibacuy</td>
<td>COL 516757</td>
</tr>
<tr>
<td>14</td>
<td>Leaves</td>
<td><em>Piper cf. eriopodom (Miq.) C.DC.</em></td>
<td>Tibacuy</td>
<td>COL 516757</td>
</tr>
<tr>
<td>15</td>
<td>Stems</td>
<td><em>Piper cf. eriopodom (Miq.) C.DC.</em></td>
<td>Tibacuy</td>
<td>COL 516757</td>
</tr>
<tr>
<td>16</td>
<td>Leaves</td>
<td><em>Piper hispidum Kunth.</em></td>
<td>Nocaima</td>
<td>COL 510518</td>
</tr>
<tr>
<td>17</td>
<td>Roots</td>
<td><em>Piper hispidum Kunth.</em></td>
<td>Nocaima</td>
<td>COL 510518</td>
</tr>
<tr>
<td>18</td>
<td>Aerial</td>
<td><em>Piper holtonii C.DC.</em></td>
<td>Arbelaez</td>
<td>COL 517184</td>
</tr>
<tr>
<td>19</td>
<td>Aerial</td>
<td><em>Piper marginatum Jacq.</em></td>
<td>Pandi</td>
<td>COL 518188</td>
</tr>
<tr>
<td>20</td>
<td>Aerial</td>
<td><em>Piper peltatum</em></td>
<td>Pandi</td>
<td>COL 512098</td>
</tr>
<tr>
<td>21</td>
<td>Aerial</td>
<td><em>Piper septulinnervium (Miq.) C.DC.</em></td>
<td>Granada</td>
<td>COL 517695</td>
</tr>
<tr>
<td>22</td>
<td>Inflorescences</td>
<td><em>Piper aff. subtomentosum Trel. &amp; Yunck.</em></td>
<td>Granada</td>
<td>COL 516758</td>
</tr>
<tr>
<td>23</td>
<td>Leaves</td>
<td><em>Piper aff. subtomentosum Trel. &amp; Yunck.</em></td>
<td>Granada</td>
<td>COL 516758</td>
</tr>
</tbody>
</table>

### RESULTS AND DISCUSSION

**HPLC-UV-DAD fingerprints of *Piper* extracts**

Chromatographic conditions (mobile and stationary phases) were optimized using sample 1 as a pattern, according to number of components and their resolution on thin layer chromatography assay (results not shown). A gradient-based reversed-phase HPLC method was developed as follows, enabling the observation of the major number of well-defined peaks at 250 nm (monitoring wavelength): Phenomenex C30 Luna column, and mobile phase consisting of acetonitrile – water; LiChrospher® 100 RP-18 endcapped (250 x 4.6 mm, 5 µm) as stationary phase (methanol – water, methanol – 0.1% formic acid, and acetonitrile – 0.1% formic acid as mobile phases were attempted too). To the best of our knowledge this study correspond to the first report on separation of secondary metabolites from *Piper* plant extracts employing HPLC with a reversed-phase C30 column.

In this study, 23 samples belonging to genus *Piper* originates from Cundinamarca department were analyzed under the optimized chromatographic conditions. According to the presence of at least 15 different *Piper* species (table 1) in the present research, the fingerprints showed to be very different (Fig. 1). Thereof, *Piper* species studied here can possess different biological activities and even several potencies for a single activity, thus it is very important to determine which species become to be related to other from the secondary metabolite content.

**Multivariate analysis**

Several preprocessing methods, including mean centering, autoscaling, range scaling, pareto scaling, vast scaling and level scaling, were applied to the dataset according to literature12,13. Before the scaling and centering...
processes a peak alignment was tested. Because of the great variation of the HPLC fingerprints to the *Piper* species, peak alignment turned out to be far from evident (it results almost impossible because of very high impossibility to define a single or a group of peaks as reference compound when alignment algorithm should be applied). Nonetheless, warping process on integrated chromatographic profiles using a correlation optimized warping algorithm (COW algorithm)\(^{16}\) was performed (warped chromatograms are not shown). Sample 1 was used as a reference, and all samples were warped to this one.

To evaluate the differences on both HPLC profiles unaligned and warped, PCA and HCA were performed. The alignment process did not result in an improvement of observed correlations to the evaluated species. As an example, Fig. 2 show the 2D score plots by comparing above mentioned data sets. On basis of these results, in the rest of the paper discussions will be given on the unaligned HPLC profiles.

[Fig. 2. Score plots of PC1 vs PC2 for raw data (a) and warped data (b) sets.]

Regarding to the effect of the pretreatment processes on data analysis, PCA and HCA were performed to each matrix data (after a single preprocessing) such as it is illustrated in Fig. 3 where it is compared the score plots. It is possible to determine what preprocessing data is able to provide the better grouping of samples. So that, normalization by centering, pareto and level scaling together with raw data didn’t really give any valuable information about any possible grouping of samples (species) at least according to if these belongs to the same species. Therefore, the above mentioned normalization processes were discarded, and no conclusions could be inferred on data results provided by these (to see Fig. 3). On the other hand, normalization by autoscaling, range scaling and vast scaling let us to find some relationships between studied samples as it is shown in Fig. 3. Last mentioned normalization processes gave different grouping results, from which vast scaling normalization resulted to be the least clear. However, it is possible to see two group; the first one conformed by samples 3, 4 and 5 which correspond to samples belonging to the specie *P. aduncum* (different plant parts), and the second one with the other samples, excepting sample 1 that appears as an isolated sample.

[Fig. 3: Continued]
The statistical comparison on secondary metabolites from *Piper* plants achieved here could be paralleled with those to DNA sequence or related data in the same way as those to morphological characters published before to establish a phylogeny of genus *Piper*. However, many species studied in the present paper has not been included into the most recent publications about phylogenetic relationships to *Piper* species. Moreover, the high species diversity within *Piper* is unique among the traditional Magnoliidae, constituting this genus as a complicated case of study. Thus, *Piper* has been submitted to different classification systems along history, which including several clades and subclades. Specifically, the neotropical clade consists of at least six well-supported subclades, being the most complex clade belonging to genus *Piper*.

To accomplish more reliable conclusions around any possible grouping into *Piper* plants from Cundinamarca (Colombia), a hierarchical clustering analysis to data after normalization processes of autoscaling, range scaling and vast scaling were carried out. Nevertheless, the most relevant information extracted from HCA corresponded with those inferred from PCA (data not shown).

Phylogenetic studies reported to *Piper* has described some relationship degree to *P. aduncum* and *P. hispidum* both belonging to Radula subclade, and *P. marginatum* with *P. peltatum*, which are representative species of Pothomorpho subclade to the best of our knowledge additional species studied here has not been reported into phylogenetic analysis. However, our research let us to conclude different relationships as follows: *P. subtomentosum* (flowers) and *P. artanthe* (aerial part); *P. hispidum* (leaves and roots) and *P. marginatum* (aerial part); *P. arboresum* (aerial part), *P. amalago* (leaves) and *P. holtonii* (aerial part); and *P. eriopodom* (fruits and leaves) and *P. cumanense* (leaves); and *P. peltatum* (aerial part), *P. septuplinervium* (aerial part) and *Piper* sp. All cases mentioned above can be understood as grouping according to some similarity in samples respect to compounds (secondary metabolites) which possess a defined chromatographic behavior (under given conditions). Therefore, each cluster or group should be constituted by species that contain compounds whose structural properties confer them similar elution time and UV absorption intensity (i.e. presence or absence of polar substituents, chromophoric groups, relative concentration), and it can be useful to the establishment of new botanical insecticides by comparing only HPLC profiles among active *Piper* plants and those which posses unknown biological properties. And, although chromatographic profiles similarities can be insufficient to guarantee any possible biological activity performance since different chemical structures could result in a very similar retention time (including comparable UV absorption coefficients) under established conditions, enclosed *Piper* plants should not be neglected because of similar chemical richness (several compounds with defined chromatographic behavior and relative concentrations) could display particular synergistic properties to a biological activity as insecticides, and besides this richness is very interesting from a viewpoint of researchers focused on natural product chemistry.

On the other hand, PARAFAC analysis were computed employing 6 components as it was indicate by core consistency diagnostic (supported by Matlab n-way toolbox). The low number of components that support the PARAFAC model of the dataset can be due to different correlations on the components (direct or reverse), so that it is a relevant parameter inside relationships focused on antimicrobial potential because of UV spectra of single components on the chromatographic profile can be very similar. PARAFAC was only able to differentiate 8 and 22 as isolated samples along PC1-PC2 plane, while 3, 4 and 5 were grouped along to PC3 axis (Fig. 5). The remaining samples were not discriminated by means of PARAFAC analysis, so that PCA on PARAFAC scores to these only samples was tested (Fig. 6). Latter analysis gave as result an almost none discrimination along PC1-PC2 plane, although with aid of the third principal component, it was possible to determine some relationships (grouping level) between 13 and 14; 9 and 10; 1, 16 and 17; 7, 19, 20 and 21; 2, 6 and 12; 15, 18 and 23. According to explained variance to PCA (Fig. 6d) it was mandatory to analyze the results related to the third principal component although not relevant information was gained. It is important to note that results could be modified in order to get major information by applying non-negativity restriction to PARAFAC model as a next step in our research.
Fig. 5. Score plots from PARAFAC analysis.

Fig. 6. PCA score plots from PARAFAC scores excluding samples 3-5, 8 and 22.

HCA on PARAFAC scores were attempted too (Fig. 7). This analysis showed a well-defined groups to 15, 23, 12 and 2; 19, 20 and 7; 16, 17 and 21; and 3, 4 and 5. To HCA, 8, 22 and 11 remain as isolated samples (high differentiation from all samples studied here). In other words, the most representative grouping was demonstrated to all different *P. aduncum* samples (stems, fruits and leaves); to *P. eriopodum* (stems), *P. subtomentosum* (leaves), *P. el-bancoanum* (wood) and *Piper* sp.; to *P. marginatum* (aerial part), *P. peltatum* (aerial part) and *P. arboreum* (aerial part); and to *P. hispidum* (leaves and roots) with *P. septuplinervium* (aerial part). Due to mathematical description of PARAFAC (n-way analysis), the result obtained after to perform HCA on PARAFAC scores gave groups of samples which has presented some correlation in both time and wavelength scales. Thus, the groups or clusters originated from latter HCA suggest *Piper* plants whose composition kept a narrow relationship as for structure core is referred. Consequently, the specimens of *P. marginatum*, *P. peltatum* and *P. arboreum* analyzed in the present report should contain compounds belonging to the same metabolite type; for example, phenylpropanoids and its derivatives has been reported as one of the main type of compound presented in *P. marginatum*\textsuperscript{19-22}, so that, we could expect of these metabolites are displayed in *P. arboreum* and *P. peltatum* too. Besides, it is possible to assure that each cluster will contain some defined metabolite types because samples such as *P. aduncum* and *P. hispidum* were grouped oneself even its different samples (from different tissues of a single specimen) exhibited different chromatographic profiles.
Some secondary metabolites from *Piper* species has been described as insecticidal compounds, which are can be used as novel insecticides or which are used in traditional control of different insects (vectors of diseases and damage storage crops)\(^2\). However, the mentioned reports have been developed on Asian species so that it is unable to correlate published results of insecticidal activity with the present findings. Some examples about biological activity of *Piper* are found published on literature\(^3,23-26\). Regarding to *Piper* plants studied here, reports on their biological activity have demonstrated to be almost inexist. The limited literature related to that is described as follows: the essential oil of *P. aduncum* has exhibited repellent activity against *Aedes aegypti*\(^2\); and the essential oil of *P. marginatum* has showed potent activity against same insect mentioned above (vector of dengue)\(^25\). Although it is important to take into account that essential oil potential can be very different to those for the crude extracts, to species enclosed to *P. marginatum* can be expected some insecticidal activity. Moreover, *P. aduncum, P. hispidum* and *P. arboresum*\(^29\) has been reported to contain amides for which these species could display insecticidal activity rather than some *Piper* amides has been proved to display a synergistic effect on herbivorous including *Spodoptera frugiperda*\(^25\). Taking into account literature\(^1\), to *P. septulineurum* extract can be expected to find amides whose structures are related to those of *P. hispidum*, which should possess fungicidal activity in some extension and possibly insecticidal activity too. In the same way, *P. peltatum* and *P. marginatum* should contain amides with high fungicidal potential, and more even insecticidal activity according to the possible content of amides into their ethanolic extracts.

**CONCLUSIONS**

Both two-way and three-way chemometric methods were applied to HPLC fingerprints of crude extracts of Colombian *Piper* species in order to achieve chemical component relationships among some species belonging to this genus. Autoscaling of chromatographic profiles without initial alignment process was demonstrated as the better result, so that this was able to exhibit relevant correlations amongst different *Piper* plants. Meanwhile, other normalization pretreatments and peak alignment (by correlation optimized warping) were discarded because of their lack of useful information offered, which was attributed to high complexity and variability among the *Piper* samples analyzed here.

PCA and HCA were performed on the choice treatmet processes resulting in suitable exploratory methods to find similarities among *Piper* extracts regarding to structural properties of secondary metabolites contained from these and which resulted responsible of their chromatographic behavior. Since authors are interested in the search of new potential botanical insecticides, enclosed *Piper* extracts should be considered as starting-point to make up a biological activity study (to higher similarity of HPLC profiles among samples resulting in suitable exploratory methods to find similarities among samples analyzed here).

On the other hand, PARAFAC was employed as a reduction data method obtaining some clustering of samples which contain similar response surfaces. In other words, enclosed *Piper* extracts correspond to samples enriched in compounds (secondary metabolites) with comparable UV spectra, and thereof, these samples could not display neither same nor similar biological activity since similar UV spectra would only indicate structural similarity which is not sufficient to ensure the activity (small structural modifications on active compounds could result in a either lost or increase of the interested property).

**ACKNOWLEDGEMENTS**

Authors are grateful with Colciencias for the financial support to the present research provided by project entitle “Bioprospección participativa de comunidades vegetales asociadas a la familia Piperaceae en la región del Sumapaz bajo y medio occidental” code 1101-05-17783 contract No. 319-2005. Universidad Nacional de Colombia and Universidad de Cundinamarca are thanked too for financial and technical collaborations.

**REFERENCES**