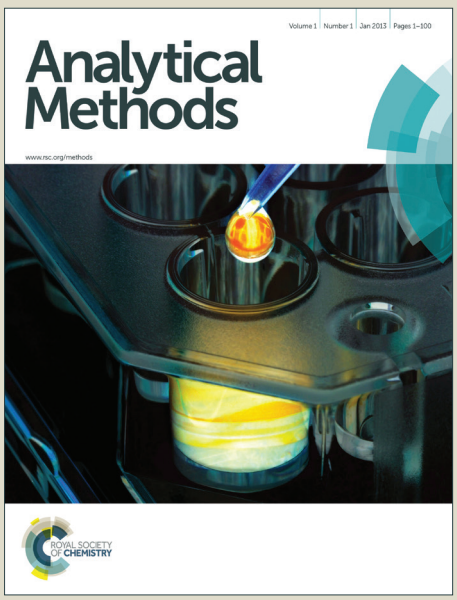
Analytical Methods

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/methods

HPLC-DAD method to metabolic fingerprinting to the phenotyping of sugarcane genotypes

Gabriel Mazzi Leme[a1], Isabel Duarte Coutinho*[a1], Silvana Creste[b], Ossamu Hojo[a], Renato Lajarim Carneiro[c], Vanderlan da Silva Bolzani[a], Alberto José Cavalheiro[a].

[a]Instituto de Química, Universidade Estadual Paulista "Julio de Mesquita Filho" (UNESP), Rua Prof. Francisco Degni, 55, CEP 14800-060, Araraquara, São Paulo, Brazil

[b]Instituto Agronômico - Centro de Cana, Rodovia Antonio Duarte Nogueira, km 321, CP206, CEP 14032-800, Ribeirão Preto, São Paulo, Brazil

[c]Departamento de Química, Universidade Federal de São Carlos (UFSCar), Rodovia Washington Luís, km 235, SP-310, São Carlos, São Paulo, Brazil

[1] These authors contributed equally to this work

* Corresponding author, E-mail address: isadcoutinho@hotmail.com

Abstract

Metabolite profiling techniques play an important role in the phenotyping and diagnostic analysis of plants. The complexity of plant metabolome makes the application of a single extraction and analytical method to all metabolites remains challenging. High Performance Liquid Chromatography coupled to Diode Array Detector (HPLC-DAD) has been the most popular analytical technique for analyzing natural products. Our group therefore aimed to develop relevant methods using experimental design techniques, using as a response the number of metabolites detectable by HPLC-DAD to compare sugarcane genotypes. Using this approach, the best extraction parameters were identified, including the number of extractions and the ratios of the solvents Water/Ethanol/Isopropanol. The chromatographic performance was optimized using a Kinetex column, methanol as solvent B and a gradient time of 60 minutes. The method was applied to the phenotyping of two sugarcane genotypes with different susceptibilities to orange rust disease. The metabolic fingerprinting was subjected to Principal Component Analysis (PCA), enabling discrimination of the genotypes based on their content of flavonoids.

1. Introduction

Plants produce a diverse variety of metabolites. It is estimated that as many as 100,000-200,000 different metabolites may be produced by plants[1,2]. A single plant can produce 5,000-25,000 compounds (*Arabidopsis* is estimated to produce approximately 5,000 compounds) at any given time[3]. Furthermore, the produced and/or modified metabolites (the metabolome) are the end products of gene expression and define the biochemical phenotype of a cell or tissue[4].

Metabolomics approaches enable the determination of a broad range of metabolites and provide great value for both the phenotyping and diagnostic analyses of plants[5]. Due to the complexity of the metabolome, in non-target metabolomic approaches have been used for metabolic phenotype analyses. A non-targeted approach can be divided in metabolic profiling that refers to analysis of a class of metabolites and metabolic fingerprinting can be defined empirically as the set of all metabolites or derivative products (identified or unknown) that are detected by analyzing a sample using a particular analytical technique[6,7].

A metabolic fingerprinting reflects the complex chemical composition of the sample and is obtained by chromatographic, spectroscopic or electrophoretic analysis[8]. These

methods are usually based on chromatographic separations. High Performance Liquid Chromatography coupled to Diode Array Detector (HPLC-DAD) is the most popular analytical technique for analyzing natural products[9]. Despite its many limitations compared to other sophisticated but costly techniques, HPLC-DAD presents many advantages for monitoring changes in the amounts of a wide range of metabolites having distinct chromophores. Indeed, depending on the chromophore, UV detection can be more sensitive and accurate than mass spectrometry methods[10].

The performance of a chromatographic fingerprint is highly dependent on the degree of chromatographic separation and the concentration distribution of the chemical components in the plant. Therefore, it is necessary to develop and optimize a methodology to create fingerprints. The resulting fingerprints should provide good selectivity, reproducibility and feasibility[11].

Sample preparation is the most important step for the development of fingerprinting, primarily because the metabolome consists of a vast number of chemical structures with diverse chemical and physical properties, including molecular weight and size, polarity, volatility, solubility, stability and pKa[3]. Therefore, the optimum extraction conditions also vary widely depending on compound type. The methods used to extract metabolites during a metabolomic study are the same as those commonly used for phytochemical studies[12]. However, additional aspects can be considered for the extraction and optimization of chromatographic conditions.

Studies on the development of fingerprinting for the quality control of herbal medicines have employed experiment design techniques, such as statistical mixture design, factorial screening design and main effect plots, to optimize the proportions of solvents used in the extraction and the duration and temperature of the extraction[13,14]. The main aim of these studies is typically to increase the diversity and yield of the analytes obtained during sample preparation and to optimize their analysis. Fractional factorial designs have successfully been employed to discern relevant from irrelevant variables during the development of chromatographic fingerprints from *Bauhinia forticata* and *Casearia sylvestris*[15].

Sugarcane is one of the most important Brazilian crops. Currently, Brazil is the world's largest sugarcane producer, with almost 9 million hectares cultivated for this crop;

approximately 589 million tons were produced in the 2012/2013 crop season[16]. Sugarcane is primarily used for the production of sugar and ethanol, and, at smaller scales, sugarcane spirits, artisanal products and *in natura* products. Botanically, sugarcane belongs to the family Poaceae, and modern sugarcane cultivars were obtained primarily by crossing the wild species *Saccharum officinarum* and *Saccharum spontaneum*[17].

In this context, breeding programs have proposed genetic introgression among modern varieties and ancestral genotypes to promote the incorporation of new genes for sucrose production, biomass accumulation and resistance to biotic and abiotic stress. Metabolic fingerprinting has become a promising approach for the phenotyping of plants. Therefore, this work aimed to develop, optimize, validate and apply an HPLC-DAD fingerprinting technique for sugarcane leaves using experimental design and multivariate analysis.

2. Experimental

2.1 Plant material

For method development and validation, sugarcane leaves (cultivar RB3280) were collected from a commercial plantation in Araraquara city (São Paulo State, Brazil), dried in an oven (40 °C) for 48 hours and milled in a knife mill.

To evaluate the applicability of the method, culms of two varieties of sugarcane, IAC 95-5000 and SP 89-1115, were obtained from the Agronomic Institute (IAC) – Cane Center on September 13, 2010. Sections of single-bud culms of each genotype were planted in 180-mL pots containing Bioplant® (Nova Ponte, Minas Gerais, Brazil) nutrient medium and grown for 30 days in a greenhouse. Afterwards, two seedlings of each genotype were transferred to 3-L pots containing the same medium. The plants were maintained in a growth chamber at 24 °C under a photoperiod of 10 h of darkness and 14 h of light. Ten seedlings were randomly harvested twice a day (at 6:00 AM and 8:00 PM) at 30, 35, 40, 45, 60 and 75 days old from five pots and immediately frozen under liquid nitrogen and stored at -80 °C. Prior to extraction, the samples were lyophilized and milled in a cryogenic mill using a first step for sample freezing (1 minute) followed by 1 cycle with 1 stage of pulverization and cooling to obtain particles smaller than 60 μm. The ten seedlings harvested from each genotype were pooled for analysis.

2.2 Chemicals and reagents

Methanol (MeOH), ethanol (EtOH), isopropanol (IPA) and acetonitrile (ACN) HPLC grade were obtained from J.T. Baker SOLUSORB® (Xalostoc, México), and glacial acetic acid (HOAc) analytical grade (99.7%) was obtained from Qhemis® (São Paulo, Brazil). Water was purified using a Millipore Milli-Q® system (Millipore Corp., New Bedford, MA, USA). Luteolin-8-*C*-glucoside (93.3%), Luteolin-6-*C*-glucoside (97.8%) and apigenin-6-*C*-glucoside (96.5%) standards were purchased from ChromaDex® (Irvine, CA, USA).

2.3 Chromatographic instrumentation and method

Chromatographic analyses were performed in a Shimadzu system (Kyoto, Japan) equipped with two LC-20AT pumps, an SPD-M20A diode array detector, a 20A autosampler, an SCL-10AVP controller, a CTO-20AC column oven and LC Solution software, version 1.23 SP1. Initially, the chromatographic runs were performed using a Luna(2) column under an exploratory gradient elution (Water+0.1% HOAc/MeOH+0.1% HOAc 95:05 to MeOH+0.1% HOAc 100% in 35 minutes, followed by 7 minutes of isocratic elution with MeOH+0.1% HOAc 100%) at a flow rate of 1.5 $mL.min^{-1}$ in an oven thermostated at 40 °C. An inverse gradient from 100 to 5% MeOH+0.1% HOAc was applied over 3 minutes, followed by isocratic elution for 15 minutes for column re-equilibration. The injection volume was 20 µL. UV spectral data were recorded from 200 to 400 nm.

A Liquid Chromatograph hyphenated to Diode Array Detector and Mass Spectrometer (LC-DAD-MS), LCQ FLEET® ThermoFinnigan (San Jose, CA), equipped with an auto sampler and a column oven was employed to identify the metabolites. The MS acquisition was performed in negative ion mode using electrospray ionization. For the Electrospray Ionization (ESI), nitrogen was used as the nebulizing gas at a flow rate of 16 $\mu L.min^{-1}$, the capillary temperature was set at 275 °C, the capillary voltage was set at -20 V, the source voltage was set to 5 kV, and the source current was set to 100 µA. The chromatographic conditions used in this analysis are the same as those optimized in section 2.5. When available, reference compounds were used to compare the corresponding retention times, mass and UV/Vis spectral profiles.

2.4 Development of extraction and clean up procedures

The extracts were prepared by sonication, using 200 mg of dried material and 2.0 mL of solvent (according a Simplex-Centroid design, Figure 1) in an ultrasonic bath Unique®

(Indaiatuba, São Paulo, Brazil, model USC-2800, 40 kHz, 120 W, 7.8 L) for 15 minutes under ambient temperature.

The solvent extraction was determined based on a Simplex-Centroid mixture model[18,19] using water, ethanol and isopropanol. The number of peaks observed in the HPLC-DAD chromatogram (254 nm) was the optimized response in order to maximize the number of compounds detectable in the extract using HPLC-DAD.

The extracts were dried in a Speed Vac® rotary evaporator (Thermo, Waltham, USA), model SPD121P-115, equipped with a refrigerator, model RVT4104 (*Refrigerated Vapor Trap*), and a vacuum pump, model OFP400-115; the extracts were then redissolved in 1.0 mL of methanol/water 1:1 (v/v). The resulting hydroalcoholic solutions were purified by solid phase extraction using SampliQ® Silica C18 ODS SPE cartridges, 3 mL/500 mg (Agilent Technologies, Wilmington, USA), which had been previously conditioned using 1.0 mL of methanol and 1.0 mL of water. The samples were eluted with 4.0 mL of methanol. The eluates were dried using a rotary evaporator and redissolved in 1.0 mL of a methanol/water 1:1 (v:v) solution, after which they were filtered through a regenerated cellulose membrane (0.20 μm pore diameter, Agilent Technologies) and stored in 1.5 mL vials. The extracts were prepared in triplicate and analyzed using HPLC.

The remaining extraction parameters were evaluated using factorial design; the parameters included were the sample weight/solvent volume ratio (w/v), the sonication time and the number of extractions, in order to maximize the extraction yield (Table 1).

Table 1. Factors and levels studied in the factorial design. Sample amount: 200 mg.

Factor	Units	Codified levels		
		(-1)	(0)	(+1)
Volume	mL	2.0	3.0	4.0
# of extractions	-	1	2	3
Extraction time	min	15	30	45

2.5 Optimization of the chromatographic conditions

The optimum chromatographic conditions, including the mobile phase composition, gradient time, flow rate, injection volume and temperature, were optimized in two steps for the following three columns from Phenomenex (Torrance, USA): Luna(2)® C18 (250 x 4.6 mm, 5 μm), Onyx® monolithic C18 (100 x 4.6 mm) and Kinetex® C18(2) (150 x 4.6 mm, 2.6 μm). For each column, a central composite design[20,21] was used to optimize the gradient time and mobile phase composition (Table 2). The oven temperature, injection volume and flow rate were kept constant at 40 °C, 20 μL and 1.5 mL.min^{-1}, respectively.

Table 2. Codified levels for gradient time and mobile phase composition used in the central composite design.

Factors	Levels				
	-1.4	**-1**	**0**	**1**	**1.4**
Gradient time (min)	24	30	45	60	66
Solvent B composition*	100:0	85:15	50:50	15:85	0:100

** Proportions between MEOH and ACN.*

In a second step, after defining the optimum mobile phase conditions, the effects of temperature, injection volume and flow rate were evaluated using a factorial design (Table 3). The values assayed for flow rate were adjusted to the different types of columns to account for differences in their packing.

Table 3. Factors and levels evaluated in the factorial design for the studied columns.

Factor		Units	Level (-1)	Level (0)	Level (+1)
Temperature		ºC	30	40	50
Injection volume		μL	10	20	30
Flow rate	Luna(2)®	mL.min^{-1}	1.0	1.4	1.8
	Onyx® in series		1.5	3.0	4.5
	Kinetex®		0.5	1.0	1.5

2.6 Method validation

Because the proposed method aimed to achieve a metabolic fingerprinting for use in comparing samples, the validation parameters evaluated were equipment precision, sample

preparation, intra-day and inter-day repeatability[11] and sample stability under various storage conditions.

2.6.1 Intra and inter-days repeatability

Three extracts (named "1", "2" and "3") were prepared from the same vegetal sample and were analyzed by HPLC-DAD using the developed sample preparation protocol. The extracts were analyzed daily in random replicates over five days; extract "1" and "2" was analyzed in triplicate, and extract "3" was analyzed in quintuplicate. The repeatability of the analysis result was evaluated based on the Relative Standard Deviation (RSD) of the number of peaks and the total peak areas of the obtained chromatograms.

2.6.2 Sample stability

Three aliquots of the hydroalcoholic extracts of sugarcane leaves were subjected to three storage conditions for five days before analysis: dried extract stored in a refrigerator (-18 °C), extract dissolved in MeOH/Water 1:1 (v/v) (2.0 $mg.mL^{-1}$) and stored in a refrigerator (-18 ^{o}C), and extract dissolved in MeOH/Water 1:1 (v/v) (2.0 $mg.mL^{-1}$) and stored at ambient temperature protected from light (the dried extract was dissolved immediately before the analysis).

The sample stability was evaluated based on the RSD of the number of peaks and the total peak areas of the obtained chromatograms.

2.7 Data processing and multivariate data analysis

The chromatographic data were reduced to ASCII files. Each dataset was arranged in a X_{IxJ} matrix, where I corresponded to rows (60 samples) and J corresponded to columns (4001 variables). The variables comprised the absorbance values at 254 nm (in mV), which were recorded every 640 ms during the chromatographic run. The data preprocessing and Principal Component Analysis (PCA) were performed using Matlab 7.12.0 (MathWorks Co., Natick, MA).

Peak alignment was implemented using the algorithm Correlation Optimized Warping (COW)[22], which was provided by Professor Rasmus Bro (Chemometrics Group, Department of Food Science, University of Copenhagen). The alignment parameters used were slack size 2 and segment length 120.

After the alignment, normalization was applied to reduce variations between the samples. PCA was performed to investigate the correlation between the chromatographic

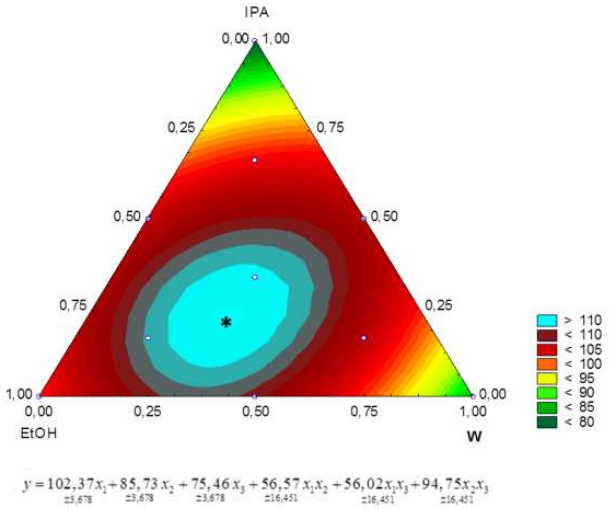
profiles of the two varieties of sugarcane and their resistance to *Puccinia kuehnii*. The region from 16.5 to 56.6 minutes was selected as the basis for the PCA.

3. RESULTS AND DISCUSSION

3.1 Sample preparation

To obtain an extract of sugarcane leaves with the greatest metabolic information, a solvent mixture was determined using an experimental Simplex-Centroid design using the number of peaks observed in the HPLC-DAD chromatogram (254 nm) as the response and considering only the peaks automatically identified by the software based on width and slope thresholds of 5 s and 1.000 $\mu V.min^{-1}$ (Signal/Noise > 3), respectively. The response surface (Figure 1), which was obtained using a quadratic model for a ternary mixture, presented low lack of fit at 95% confidence (r^2 0.947).

Figure 1. Response surface of the number of chromatographic peaks as a function of solvent proportion. Simplex-Centroid design - the experimental points are indicated with white dots, and the optimum mixture is indicated with an asterisk.

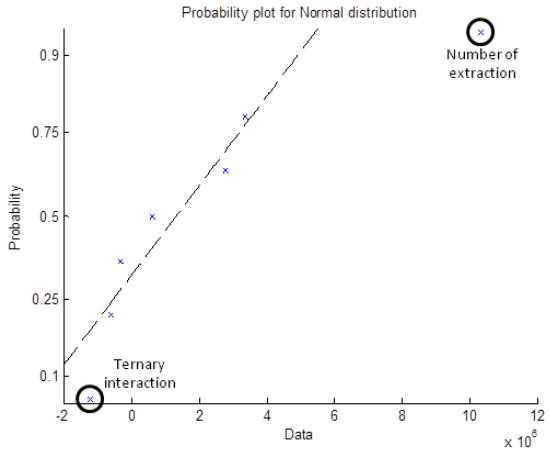


The proportion of solvents that extracted the largest number of compounds (as detected using DAD) was Water/EtOH/IPA 30:45:25 (v/v/v). Next, a two-level factorial design was used to evaluate the effect of sonication time, the number of extractions and the sample/solvent ratio for the purpose of improving the extraction yield. The values of the levels of the variables used in this study are presented in Table 1.

Based on the normal probability plot (Figure 2), it is evident that the number of extractions was the most important parameter for optimizing the extraction yield. The

ternary interaction exerted only a slight influence, and the other parameters were not significant in the experimental domain evaluated.

Figure 2. Normal probability plot of the two-level factorial design used to evaluate the sonication time, number of extractions and sample/solvent ratio for improving the extraction yield.



Hence, seven consecutive extractions were performed in triplicate. The sum of the total peak area and the extract weight of the seven extractions was taken as 100%. The first 3 extractions yielded 81.5 ± 2.6% of the total peak area and 80.9 ± 5.8% of the extract weight. Thus, the sample preparation was defined as follows: 200 mg of dried plant material were extracted three times with 2 mL of Water/EtOH/IPA (30:45:25) by sonication for 15 minutes. The supernatants were consolidated, dried using a Speed Vac rotary evaporator, dissolved in 1 mL of MeOH/W (1:1) and submitted to a clean up using an SPE-C18 column and eluted with 4 mL of MeOH. The eluate was dried using a Speed Vac rotary evaporator, dissolved in 1 mL of MeOH/Water (1:1), ultrafiltered and subjected to HPLC analysis.

3.2 Optimization of the chromatographic conditions

The extract from sugarcane leaves obtained using the optimized conditions was used to improve the chromatographic parameters in two steps. First, the mobile phase composition and the gradient time were optimized for each column using a central composite design that consisted of a 2^k factorial including $2k$ axial points and n central points, where k represents the number of factors to be evaluated. The real and codified values are presented in Table 2, and the results obtained using the various columns (i.e., the number of peaks observed in

the chromatograms) are shown in Table 4. The best results were obtained using the Kinetex® column, MeOH as solvent B and a gradient time of 60 minutes.

Table 4. Number of peaks as a function of experiment number for the central composite design (see the variable codifications presented in Table 2).

Exp.	Solvent B	Gradient time	Luna (2)®	Onyx®	Kinetex®
1	-1	-1	132	129	169
2	1	-1	132	114	151
3	-1	1	154	**159**	**220**
4	1	1	157	145	187
5	-1.4	0	143	146	198
6	1.4	0	142	132	173
7	0	-1.4	113	117	147
8	0	1.4	**164**	152	217
9*	0	0	154 (± 1.73)	142 (± 3.78)	189 (± 2.31)

**Central point in triplicate, average values and standard deviations are presented in parentheses.*

It was not possible to construct prediction models using the number of peaks as a function of solvent B and the gradient time for the Onyx® and Kinetex® columns because the data did not fit the models well (r^2 of 0.63 and 0.78, respectively). However, based on the empirical results, the best experiments were easily identified such that they could be used to increase the number of peaks in the chromatogram.

To conclude the discussion of the chromatographic method development, the effects of the injection volume, column oven temperature and mobile phase flow rate were evaluated for each column based on a two-level factorial design with a central point. The results are summarized and presented in Table 5. Even using the probability plot, it was impossible to identify the most important variables in the experimental domain clearly, and the linear models exhibited a poor fit. However, based on the empirical results presented in Table 5, the parameters evaluated followed the same trend, exhibiting the maximum number of peaks around the medium level.

Table 5. Number of peaks as a function of experiment number for the two-level factorial design (see variable codifications in Table 3).

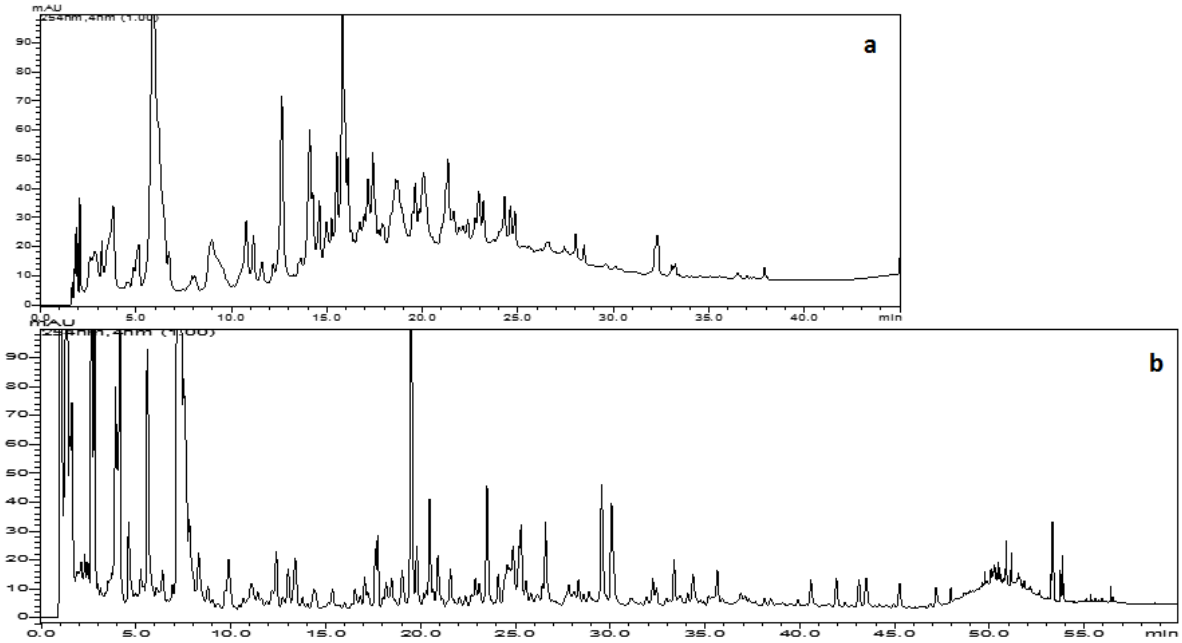
Exp.	Oven temp. (° C)	Inj. Vol. (µL)	Flow rate ($mL.min^{-1}$)	Peaks		
				Luna(2)®	Onyx®	Kinetex®
1	-1	-1	-1	141	166	212
2	1	-1	-1	171	164	187
3	-1	1	-1	147	176	184
4	1	1	-1	162	162	190
5	-1	-1	1	153	151	197
6	1	-1	1	155	141	222
7	-1	1	1	146	181	203
8	1	1	1	148	138	215
9*	0	0	0	**164 (± 4)**	**181 (±2)**	**215 (±3)**

**Central point in triplicate, average values and standard deviations are presented in parentheses.*

3.2.1 Gradient adjustments

The initial gradient condition was adjusted to 10% B to more rapidly eliminate the highly polar compounds that could not be separated using RP-C18 and those that were poorly resolved early in the chromatogram. At the end of the gradient, a fast gradient step (10 min) to 100% B was added and held isocratically for 5 minutes to complete the elution of the lipophilic compounds. The re-equilibration time used was 20 minutes (corresponding to approximately 10 column volumes, as recommended by Snyder et al, 2010)[23]. Figure 3 shows the chromatographic runs before and after the optimization process, clearly revealing the improvement in the method.

Figure 3. Cultivar RB3280 chromatographic analysis: a) initial chromatographic conditions and b) optimized chromatographic conditions (Kinetex® column).



3.3 Method validation

The repeatability of the sample preparation procedure was evaluated by analyzing three sugarcane extracts prepared with aliquots of the same plant material on different days. The relative standard deviation in the number of peaks and the total peak area were 0.95% and 7.21%, respectively. The relative standard deviations in the intraday repeatability were 1.34% and 6.33% for the number of peaks and total peak area, respectively, and the relative standard deviations in the interday repeatability were 1.61% and 6.11% for the number of peaks and the total peak area, respectively.

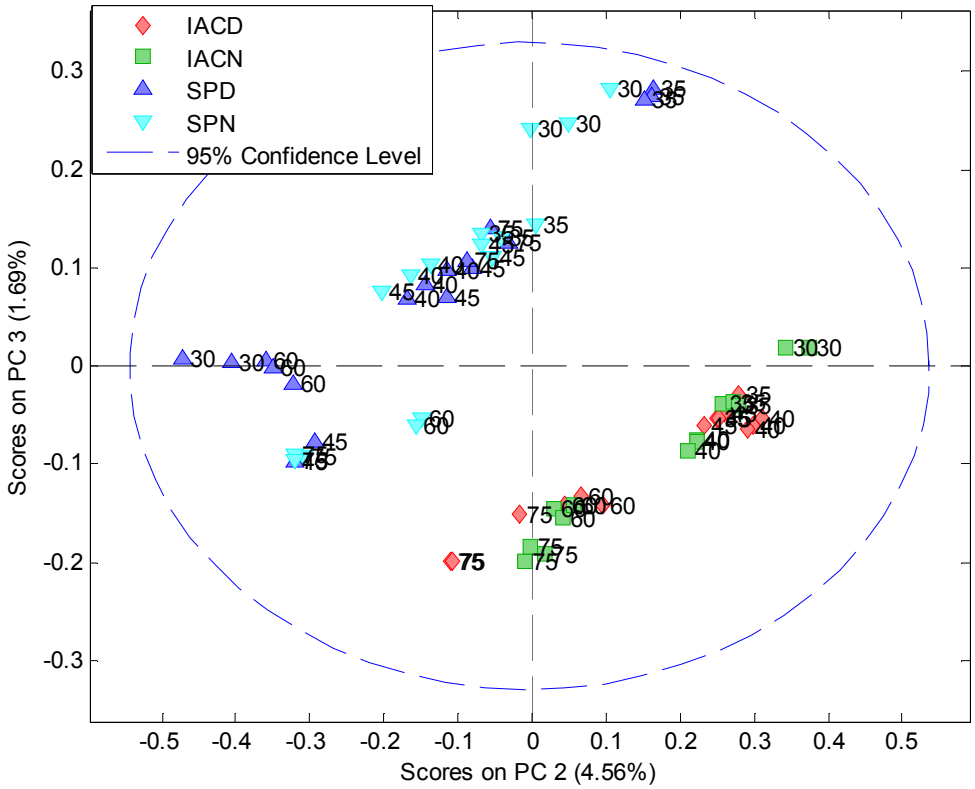
The sample stability was evaluated by analyzing samples stored under three conditions for five days. The highest relative standard deviation in the number of peaks was obtained for the sample stored in solution at ambient temperature (1.8%), followed by the sample solution stored under refrigeration (1.44%) and, finally, by the sample stored dried and refrigerated (1.09%). The opposite trend was observed for the total peak area, in which the dried and refrigerated extract exhibited the highest relative standard deviation (0.78%); the refrigerated solution and the ambient temperature solution exhibited RSD values of 0.65% and 0.14%, respectively. These results have important implications for the proposed chromatographic method because, typically, several samples are prepared and then stored in an autosampler, where they await their turn for analysis, sometimes for lengthy periods.

3.4 Evaluation of the application of the method

The sample preparation protocol and chromatographic method developed herein were used to obtain the metabolite fingerprint of two genotypes of sugarcane, SP 89-1115 and IAC 95-5000, which are resistant and susceptible to orange rust, respectively, during seedling development. The genotypes were harvested twice daily after 30, 35, 40, 45, 60 and 75 days from germination, at 6:00 AM and 8:00 PM. Principal Component Analysis was performed to investigate the main differences between the two genotypes. The first four principal components accounted for 97.29% of the data variance. The combination of PC2 and PC3 presented in Figure 4 reveal well-separated clusters, with the samples from resistant seedlings clustered in one group and the samples from susceptible seedlings clustered in another. A trend for developing seedlings can be observed in the scores plot: moving from positive to negative PC2 values was related to the seedling age.

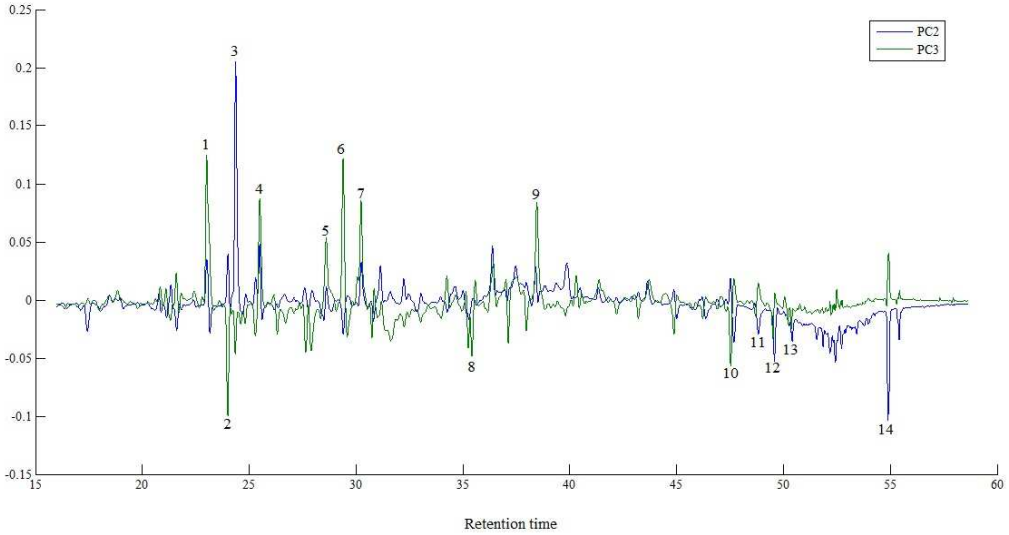
Based on the score (Figure 4) and loading (Figure 5) plots for PC2 versus PC3, the second component primarily accounted for the variance through variables 3 and 14, whereas the third component contributed to the variance through variables 1, 2 and 4-13.

Figure 4 – Scores plot for PC2 vs. PC3. IACD: genotype IAC 95-5000 harvested at 6:00 AM; IANC genotype IAC 95-5000 harvested at 8:00 PM; SPD: genotype SP89-1115 harvested at 6:00 AM; SPN SP89-1115 harvested at 8:00 PM. The numbers shown on the plot represent the number of days after germination.



The variables responsible for the clustering and discrimination between the samples were identified using an "in-house" databank constructed by our group (unpublished data) that contains the LC-DAD-MS and retention time data for the metabolites identified in the culm, leaves and juice of sugarcane[24-29]. Using this information, the samples located on the positive side of PC2 and PC3 exhibited higher concentrations of luteolin-8-*C*-glucoside (3), whereas the SP89-1115 located at positive PC3 and negative PC2 exhibited a greater abundance of the metabolites apigenin-8-*C*-arabinosyl-6-*C*-glucoside (2), diosmetin-8-*C*-glucoside (4), apigenin-6-*C*-glucoside (5) and tricin-7-*O*-glucosyl-rhamnoside (9). The other compounds have not been identified yet.

Figure 5 – Loadings *versus* scores plot (PC2 and PC3).



The discovery of high levels of luteolin-8-*C*-glucoside in the resistant sugarcane variety suggested that the IAC 95-5000 genotype contains constitutive antifungal agents. It is well established that flavonoids contribute to disease resistance, either as pre-formed antifungal agents or as phytoalexins[30]. Flavone *C*-glycosides have already been described as phytoalexins after their discovery in cucumbers infected with powdery mildew fungus[31]. The flavone glycosides found in high levels in IAC95-5000 could be considered phytoanticipins (pre-formed), which are stored as inactive molecules that can be released and activated following biotic attack. However, studies with sugarcane varieties inoculated with *Puccinia kuehnii* would be necessary to prove this hypothesis.

4. Conclusions

On the basis of the above results, our work led to the development of a new protocol for sample preparation for metabolic fingerprinting analyses. We applied a well-designed set of experiments to extract the largest number of compounds detectable in HPLC-DAD from sugarcane leaves. The best solvent extractor was defined as a mixture of water, ethanol and isopropanol, and this solvent enabled the extraction of a wide range of compounds, which were used to obtain a representative fingerprint of the sample. In addition, HPLC conditions were proposed after the evaluation of three C18 columns using factorial design to optimize the mobile phase parameters. The Kinetex® column exhibited the greatest efficiency for complex matrix analysis, and use of this column provided more peaks than the other columns evaluated.

The proposed method was validated in terms of the repeatability of the sample preparation and chromatographic analysis by monitoring the number of peaks and total peak area. The sample stability was also evaluated, revealing that the sample remained stable over the studied period (8 days) under various storage conditions, including dried and refrigerated, dissolved and refrigerated, and dissolved at ambient temperature, thus demonstrating that immediate analysis is not necessary, allowing the simultaneous preparation of several samples and the use of automatic samplers during the chromatographic analysis. These results validated the use of the sample preparation protocol and HPLC conditions for obtaining and comparing the metabolic fingerprints of sugarcane leaves.

Using this method, the seedling development of two varieties was compared; aided by PCA analysis, we were able to verify that variety IAC 95-5000 contains higher levels of luteolin-8-*C*-glucoside, possibly related to its resistance to orange rust. This information could be applied to the development of new resistant varieties, considering that Brazilian genetic breeding programs have not evaluated metabolite fingerprints during the selection of resistant traits.

We point out that the proposed method focus on the analysis of sugarcane, however the strategy for the method development is applicable to any sample and offers a rational alternative to chromatographic optimization instead of the usually employed "trial and error" approaches.

Acknowledgments

The authors wish to thank Pedro Luís da Costa Aguiar Alves for providing access to the growth chamber and the agencies CNPQ and CAPES for financial support.

References

1 J.W. Allwood, R. Goodacre, *Phytochem. Anal*., 2010, 21, 33-47.

2 K-M., Oksman-Caldentey, D. Inzé, *Trends Plant Sci*., 2004, 9, 433-440.

3 D. M. Beckles, U. Roessner, *in Plant metabolomics: Applications and opportunities for agricultural biotechnology*, ed. A. Altman, P. Hasegawa, Elsevier, Oxford, 2012, ch. 5, pp. 67-81.

4 L. W. Sumner, P. Mendes, R. A. Dixon, Phytochemistry, 2003, 62, 817-836.

5 A. R. Fernie, N. Schauer, *Trends Genet*., 2008, 25, 39-48.

6 Villas-Bôas, S. G., Rasmussen, S.; Lane, G. A., 2005. Metabolomics or metabolite profiles? Trends Biotechnol. 23, 385.

7 Ibáñez, C.; Garcia-Cañas, V.; Valdés, A.; Simó, C., *Trends in Anal. Chem*., 2013, 52, 100-111.

8 C. Tistaert, G. Chataigné, B. Dejaeghera, C. Rivière, N. Nguyen Hoaic, M. Chau Van, J. Quetin-Leclercq, Y. Vander Heyden, *J. chromatogr. B*, 2012, 910, 103-113.

9 C. Tistaert, B. Dejaegher, Y. V. Heyden, *Anal. Chim. Acta*, 2011, 690, 148-161.

10 A. Ishihara, F. Matsuda, H. Miyagawa, K. Wakasa, *Metabolomics*, 2007,3, 319-334.

11 Y-B. Ji, Q-S. Xu, Y-Z. Hu, Y. V. Heyden, *J. Chromatogr. A*, 2005, 1066, 97-104.

12 H. K. Kim, R. Verpoorte, *Phytochem. Anal*., 2010, 21,4-13.

13 F. Delaroza, I. S. Scarminio, *J. Sep. Sci*., 2008, 31, 1034-1041.

14 D. X. Soares, I. S. Scarminio, R. E. Bruns, *Anal. Chim. Acta*, 2011, 702, 288-294.

15 C. Funari, R. Carneiro, A. Andrade, E. Hilder, A. J. Cavalheiro, *J. Sep. Sci.*, 2014, 37, 37-44.

16 CONAB. Acompanhamento da Safra Brasileira. URL (http://www.conab.gov.br/OlalaCMS/uploads/arquivos/14_04_10_09_00_57_boletim_cana_portugues_-_4o_lev_-_13.pdf) (accessed June, 2014).

17 A. Cheavegatti-Gianotto, *Trop. Plant Biol*., 2011, 4, 62-89.

18 A. A. Almeida, I. S. Scarminio, *J. Sep. Sci*., 2007, 30, 414-420.

19 P. K. Soares, R. E., Bruns, I. S. Scarminio, *J. Sep. Sci*., 2009, 32, 644-352.

20 B. Barros Neto, I. S. Scarminio, R. E. Bruns, in *Como fazer experimentos: pesquisa e desenvolvimento na ciência e na indústria*, ed. B. Barros Neto, I. S. Scarminio, R. E. Bruns, Bookman, Porto Alegre, Brazil, 4th edn., 2010, ch. 6, pp. 294.

21 E. V. C. Galdámez, Master's Thesis, University of São Paulo, 2002.

22 G. Tomasi, T. Skov, F. V-Berg, Dynamic time warping (DTW) and correlation optimized Warping (COW), (http://www.models.kvl.dk/source/DTW_COW/index.asp), (January, 2014).

23 L. R. Snyder, J. J. Kirkland, J. W. Loyd, in *Introduction to modern liquid chromatography*, 3rd edn., John Wiley & Sons, Hoboken, New Jersey, 2010, ch. 9, pp. 448.

24 R. Colombo, J. H. Yariwake, M. McCullaghb, *J. Braz. Chem. Soc*., 2008, 19, 483-490.

25 R. Colombo, J. H. Yariwake, E. F. Queiroz, K. Ndjoko, K. Hostettmann, *J. Braz. Chem. Soc*., 2009, 20, 1574-1579.

26 R. Colombo, J. H. Yariwake, E. F. Queiroz, K. Ndjoko, K. Hostettmann, *J. Chromatogr. A*, 2005, 1082, 51-59.

27 R. Colombo, J. H. Yariwake, E. F. Queiroz, K. Ndjoko, K. Hostettmann, *Phytochem. Anal*. 2006, 17, 337-343.

28 J. M. Duarte-Almeida, G. Negri, A. Salatino, J. E. Carvalho, F. M. Lajolo, *Phytochemistry*, 2007, 68, 1165-1171.

29 J. M. Duarte-Almeida, A. V. Novoa, A. F. Linares, F. M. Lajolo, M. I. Genovese, *Plant Food Hum. Nutr*., 2006, 61, 187-192.

30 J. B. Harbone, C.A. Williams, *Phytochemistry*, 2000, 55, 481-504.

31 D. J. McNally, K. V. Wurms, C. Labbe, R. R. Belanger, *Physiol. Mol. Plant. Pathol*., 2003, 63, 293-303.