JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY
ARTICLE

J. Agric. Food Chem. 2009, 57, 7651–7658 7651
DOI:10.1021/jf901508f

Metabolic Changes in Different Developmental Stages of *Vanilla planifolia* Pods

Tony Lionel Palama,[†,§] Alfi Khatib,[∥] Young Hae Choi,[§] Bertrand Payet,[‡] Isabelle Fock,[†] Robert Verpoorte,[§] and Hippolyte Kodja*[,†]

[†]UMR C53, Peuplements Végétaux et Bioagresseurs en Milieu Tropical and [‡]Laboratoire de Chimie des Substances Naturelles et des Sciences des Aliments, Faculté des Sciences et Technologies, Université de La Réunion, 15 avenue René Cassin, B.P. 7151, 97715 Saint Denis Messag, Cedex 9, La Réunion, France, [§]Division of Pharmacognosy, Section Metabolomics, Institute of Biology, Leiden University, Post Office Box 9502, 2333 CC Leiden, The Netherlands, and [∥]Faculty of Food Science and Technology, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

The metabolomic analysis of developing *Vanilla planifolia* green pods (between 3 and 8 months after pollination) was carried out by nuclear magnetic resonance (NMR) spectroscopy and multivariate data analysis. Multivariate data analysis of the ^{1}H NMR spectra, such as principal component analysis (PCA) and partial least-squares-discriminant analysis (PLS-DA), showed a trend of separation of those samples based on the metabolites present in the methanol/water (1:1) extract. Older pods had a higher content of glucovanillin, vanillin, *p*-hydroxybenzaldehyde glucoside, *p*-hydroxybenzaldehyde, and sucrose, while younger pods had more bis[4-(β-D-glucopyranosyloxy)-benzyl]-2-isopropyltartrate (glucoside A), bis[4-(β-D-glucopyranosyloxy)-benzyl]-2-(2-butyl)tartrate (glucoside B), glucose, malic acid, and homocitric acid. A liquid chromatography−mass spectrometry (LC−MS) analysis targeted at phenolic compound content was also performed on the developing pods and confirmed the NMR results. Ratios of aglycones/glucosides were estimated and thus allowed for detection of more minor metabolites in the green vanilla pods. Quantification of compounds based on both LC−MS and NMR analyses showed that free vanillin can reach 24% of the total vanillin content after 8 months of development in the vanilla green pods.

KEYWORDS: *Vanilla planifolia*; vanillin; pod development; metabolomic analysis; nuclear magnetic resonance spectroscopy

INTRODUCTION

Vanilla is one of the most popular flavoring ingredients used in food, beverages, cosmetics, and tobacco. To satisfy the ever-increasing demand for vanilla-flavored products, industries have begun to substitute natural vanilla extracts with vanillin. Vanillin (4-hydroxy-3-methoxybenzaldehyde) is the major compound of vanilla flavor and can be produced semi-synthetically by (bio)conversion of related natural products or via synthesis. Nowadays, less than 1% of the global production of vanillin (12 000 tons each year) is derived from vanilla pods (*1*). Indeed, synthesized vanillin is much cheaper (< $15/kg) than vanillin extracted from vanilla pods [estimated between $1200 and $4000 per kg (*1*)]. Nevertheless, there is a distinct difference between the flavor of vanilla extract, which is composed of more than 200 compounds, and that of pure synthesized vanillin (*1*). Most people prefer the flavor of the natural extract.

Natural vanilla is obtained from the pods of *Vanilla planifolia* G. Jackson, a member of the orchid family (Orchidaceae). Although originally from Mexico, this plant is now grown in various others tropical countries, such as Madagascar, Indonesia, Uganda, Papua New Guinea, India, and islands, such as Comoros, Tahiti, and La Réunion. Madagascar produces more than half of the world production (1000−1200 tons), and Indonesia is the second largest producer, with about 350 tons (*2*). Flavorless mature green pods are harvested about 8−9 months after pollination and submitted to a curing process, which allows the development of the vanilla aroma. The curing process used by vanilla producers in the Indian Ocean region is known as the Bourbon method, so-called because it originated in the former French colony of Bourbon Island, today known as La Réunion. At present, Madagascar is the main producer of Bourbon-type pods, while production in La Réunion and Comoros, although smaller, is still significant (*3*). The curing process of vanilla involves four steps: killing, sweating, drying, and conditioning, after which the vanilla pods turn a chocolate brown color. During this process, phenolic glucosides, such as glucovanillin, are hydrolyzed to free forms, which are responsible for the flavor. Phenomena such as oxidation or polymerization of compounds have also been observed (*4*−*6*). Variations in the curing process can lead to noticeable differences in the flavor of the pods (*7*). Although many studies have been conducted to optimize the

*To whom correspondence should be addressed. Telephone: +262-262-93-81-71. Fax: +262-262-93-81-19. E-mail: kodja@univ-reunion.fr.

Published on Web 08/14/2009

process and to identify all of the components of vanilla flavor, few studies have focused on the chemical composition of the green pods. Of these, most of them have focused on their composition in some phenolic glucosides, reputedly the major precursors of flavor compounds (*8*−*17*). However, thus far, no study of the wide range of metabolic changes occurring during vanilla pod development has been published.

Metabolomics is defined as both the qualitative and quantitative analysis of all primary and secondary metabolites of an organism (*18*). Chemical analysis techniques to be applied to metabolite profiling should be rapid, reproducible, and stable over time while requiring only simple sample preparation. Nuclear magnetic resonance (NMR) spectroscopy potentially meets these requirements. Recently, the combination of NMR and multivariate data analysis has been applied to the metabolic profiling of various plant species (*19*−*25*) and also to wine (*26*), coffee (*27*), fruit juice (*28*), and beer (*29*).

In this study, we report a ^{1}H NMR spectroscopy method coupled with multivariate analysis for the analysis of metabolic variation of developing *V. planifolia* pods. Although ^{1}H NMR spectroscopy allows for good quantization of a large range of metabolites, this technique cannot detect minor compounds (*18*). Additionally, we used liquid chromatography−mass spectrometry (LC−MS) to obtain additional information. This more targeted approach was applied to enhance the sensitivity and the specificity of the analysis on phenolic compounds in developing pods.

MATERIALS AND METHODS

Plant Material. Flowers of *V. planifolia* were hand-pollinated and labeled during the second week of November 2006 in a shade house in Saint-André, La Réunion. Four pods were collected simultaneously from five accessions of *V. planifolia* and immediately plunged into liquid nitrogen. Frozen pods were then freeze-dried to minimize possible enzyme degradation. The four pods of each accession were ground together to a fine powder with a mortar and pestle before storing in a cold room (−20 °C) until analysis. This protocol was performed at 6 different times (3−8 months) after pollination. Thus, 30 samples were analyzed in NMR and laser scan mass spectrometry (LS-MS) (5 accessions × 6 developmental stages). For each accession, the same plant was always used for harvesting pods submitted to metabolomic analysis.

NMR Analysis. *Extraction*. Freeze-dried plant material (50 mg) was transferred to a 2 mL microtube. A volume of 1.5 mL of a mixture of KH_2PO_4 buffer (pH 6.0) in D_2O containing 0.05% trimethylsilylpropionic acid sodium salt (TMSP, w/w) and methanol-d_4 (1:1) was added to the plant samples. The mixture was vortexed at room temperature for 1 min, ultrasonicated for 20 min, and centrifuged at 13 000 rpm for 10 min. An aliquot of 0.8 mL was used for NMR analysis.

Measurements. All of the NMR parameters were the same as those used by Abdel-Farid et al. and Jahangir et al. (*24*, *30*). ^{1}H NMR and 2D J-resolved spectra were recorded at 25 °C on a 500 MHz Bruker DMX-500 spectrometer (Bruker, Karlsruhe, Germany) operating at a proton NMR frequency of 500.13 MHz. MeOH-d_4 was used as the internal lock. Each ^{1}H NMR spectrum consisted of 128 scans requiring a 10 min and 26 s acquisition time with the following parameters: 0.16 Hz/point, pulse width (PW) = 30° (11.3 μs), and relaxation delay (RD) = 1.5 s. A presaturation sequence was used to suppress the residual H_2O signal with low power selective irradiation at the H_2O frequency during the recycle delay. Free induction decays (FIDs) were Fourier-transformed with LB = 0.3 Hz. The resulting spectra were manually phased and baseline-corrected and calibrated to TMSP at 0.0 ppm, using XWIN NMR (version 3.5, Bruker). 2D J-resolved NMR spectra were acquired using 8 scans per 128 increments for F1 and 8K for F2 using spectral widths of 5000 Hz in F2 (chemical-shift axis) and 66 Hz in F1 (spin−spin coupling constant axis). A 1.5 s relaxation delay was employed, giving a total acquisition time of 56 min. Data sets were zero-filled to 512 points in F1, and both dimensions were multiplied by sine-bell functions (SSB = 0) prior to double complex FT. J-resolved spectra tilted by 45° were symmetrized about F1 and then calibrated, using XWIN NMR (version 3.5, Bruker).

^{1}H−^{1}H correlated spectroscopy (COSY) and heteronuclear multiple bonds coherence (HMBC) spectra were recorded on a 600 MHz Bruker DMX-600 spectrometer (Bruker). The COSY spectra were acquired with a 1.0 s relaxation delay and 6361 Hz spectral width in both dimensions. The window function for COSY spectra was sine-bell (SSB = 0). The HMBC spectra were obtained with a 1.0 s relaxation delay and 30 183 Hz spectral width in F2 and 27 164 Hz in F1. Qsine (SSB = 2.0) was used for the window function of the HMBC. The optimized coupling constant for HMBC was 8 Hz.

Data Analysis. The ^{1}H NMR spectra were automatically reduced to the ASCII file. Spectral intensities were scaled to total intensity and reduced to integrated regions of equal width (0.04), corresponding to the region of δ 0.30−10.02. The regions of δ 4.74−4.98 and 3.30−3.38 were excluded from the analysis because of the residual signal of HDO and CD_3OD, respectively. Bucketing was performed by AMIX software (Bruker) with scaling on total intensity. Principal component analysis (PCA) and partial least-squares-discriminant analysis (PLS-DA) were performed with the SIMCA-P software (version 12.0, Umetrics, Umeå, Sweden) with scaling based on Pareto and the unit variance method, respectively. The analysis of variation (ANOVA) for ^{1}H NMR signals was performed using the R software (*31*).

LC−MS Analysis. *Extraction*. The extraction procedure was derived from Odoux (*7*). Freeze-dried plant material (100 mg) was extracted with 8 mL of a 4 °C precooled phosphate buffer (pH 5) and kept at 4 °C on a platform shaker overnight. The mixture was then centrifuged for 30 min at 4 °C and 3000 rpm. The residue was extracted with a further 4 mL of a 4 °C precooled phosphate buffer using the same method. Supernatants were pooled, and the extract was rapidly heated to 80 °C for 10 min to inactivate endogenous β-glucosidase. After cooling, half of the extract was hydrolyzed with 10 units of almond β-glucosidase (Sigma, ref 49290) at 35 °C for 1 h for the enzymatic hydrolysis of glucosylated precursors. Both hydrolyzed and nonhydrolyzed extracts were studied. This extraction procedure was performed twice and analyzed separately by LC−MS.

Measurements. The LC system employed was an Agilent CPL/SM 1100 series (Massy, France) equipped with LC/MSD Chemstation software, degasser G1322A, binary pump G1312A, autosampler G1313A, thermostatted column oven G1316A, diode array detection system G1315B to monitor between 200 and 400 nm, and a MSD/VL detector with an electrospray source. A LiChrospher 100 RP-18 (250 × 4.6 mm inner diameter, s-5, 5 μm) (Merck, Darmstadt, Germany) column was used with a LichroCART 4-4 (Merck) guard column, at 30 °C. Samples (10 μL) were eluted with a two-solvent gradient of 100% of 0.1% acetic acid in water to 15% acetonitrile in 55 min and then back to original conditions in 5 min at a flow rate of 1.0 mL/min. The electrospray mass spectrometer conditions were as follows: negative ion mode; fragmentation voltage, 70 V; capillary voltage, 4000 V; drying gas (nitrogen) flow, 11 mL/min; nebulizer pressure, 60 psig; drying gas temperature, 350 °C; mode scan, m/z 50−1000. First, a library with the high-performance liquid chromatography (HPLC) retention times and ultraviolet diode array detector (UV-DAD) spectra was made with 12 compounds (*p*-hydroxybenzyl alcohol, *p*-hydroxybenzoic acid, *p*-hydroxybenzaldehyde, protocatechuic acid, protocatechualdehyde, vanillyl alcohol, acetovanillone, vanillin, vanillic acid, caffeic acid, ferulic acid, and *p*-coumaric acid), and a calibration table was constructed for each compound. The mass spectrum of each reference compound was also recorded and used to confirm identity.

RESULTS AND DISCUSSION

Visual Inspection of ^{1}H NMR Spectra and Assignments of Compounds. *V. planifolia* pods of six different ages (from 3 to 8 months after pollination) were analyzed by NMR to evaluate the effect of age on the metabolic profile. Different classes of metabolites, such as phenolic compounds, carbohydrates, and organic acids, were detected in the ^{1}H NMR spectra of the methanol/water (1:1) extract (**Figure 1**). The major differences occurred in the aromatic region (δ 6.0−8.0), where an increase of

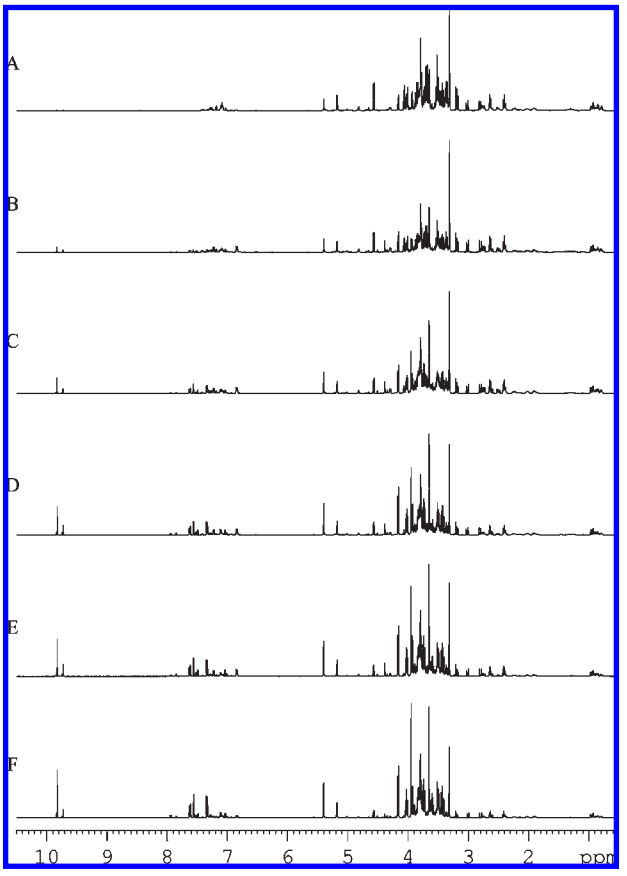

Figure 1. ^{1}H NMR spectra (methanol-d_4−KH_2PO_4 in D_2O at pH 6.0 extract) of *V. planifolia* pods of (A) 3 months old, (B) 4 months old, (C) 5 months old, (D) 6 months old, (E) 7 months old, and (F) 8 months old in the range of δ 0.5−10.5.

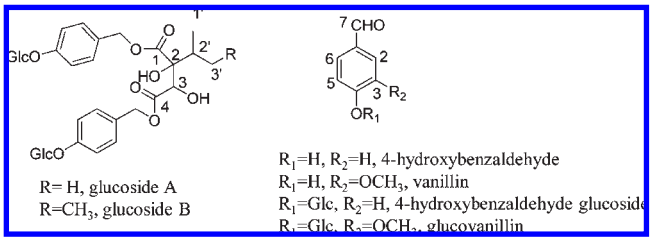

Figure 2. Chemical structures of phenolic metabolites identified by ^{1}H NMR in *V. planifolia* green pods.

phenolic compounds was observed during development. The ^{1}H NMR signals of this region were assigned by a comparison of ^{1}H NMR spectra of the reference compounds and ^{1}H−^{1}H COSY. Apart from the chemical-shift data of ^{1}H NMR, HMBC spectra can provide evidence for the identification of glycosides. The signals of glucovanillin (**Figure 2**) were clearly distinguished in the ^{1}H NMR spectra. Major signals corresponding to H-6 at δ 7.62 (dd, J = 8.5, 2.0 Hz), H-2 at δ 7.57 (d, J = 2.0 Hz), and H-5 at δ 7.34 (d, J = 8.5 Hz) were observed in the aromatic region of the green pod extract (**Figure 3**). In addition to these aromatic signals, other characteristic signals of glucovanillin, such as H-7 at δ 9.82 (s), OCH_3 of C-3 at δ 3.95 (s), and H-1′ of the glucose moiety at δ 5.19 (d, J = 7.8 Hz), were clearly identifiable in the spectra. Signals characteristic of vanillin, *p*-hydroxybenzaldehyde glucoside, and *p*-hydroxybenzaldehyde were also detected as other main phenolic compounds of *V. planifolia* green pods. In most cases, level of these compounds during pod development increased (**Figure 3**).

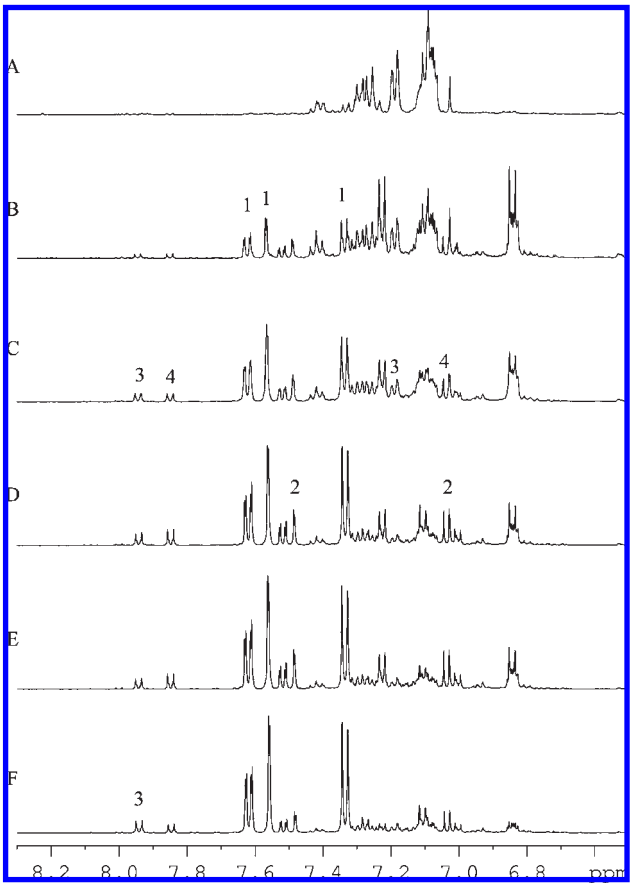

Figure 3. ^{1}H NMR spectra (methanol-d_4−KH_2PO_4 in D_2O at pH 6.0 extract) of *V. planifolia* pods of (A) 3 months old, (B) 4 months old, (C) 5 months old, (D) 6 months old, (E) 7 months old, and (F) 8 months old in the range of δ 6.55−8.25. Assignments: 1, glucovanillin; 2, vanillin; 3, *p*-hydroxybenzaldehyde glucoside; 4, *p*-hydroxybenzaldehyde.

Differences were also observed in the anomeric signals of carbohydrates, such as δ 5.44 (d, J = 3.5 Hz), δ 5.18 (d, J = 3.5 Hz), and δ 4.57 (d, J = 7.8 Hz). These signals were assigned as anomeric protons of sucrose, α-glucose, and β-glucose. Another anomeric signal obtained from the fructose moiety of sucrose was also well-distinguishable at δ 4.17 (d, J = 7.8 Hz). In this region, development of the green pods was marked by an increase of sucrose. Signals corresponding to H-2 at δ 4.29 (dd, J = 7.8, 3.5 Hz), H-3 at δ 2.74 (dd, J = 16.0, 3.5 Hz), and H-3′ at δ 2.49 (dd, J = 16.0, 7.8 Hz) have been well-assigned to malic acid. Resonances at δ 3.01 (d, J = 16.0 Hz) coupled with δ 2.8 (d, J = 16.0 Hz) and at δ 2.64 (dd, J = 9.5, 7.5 Hz) coupled with δ 2.41 (dd, J = 9.5, 7.5 Hz) provided interesting information. An expanded ^{1}H NMR spectrum in the range of δ 2.3−3.1 is shown in **Figure 4**. Analysis of the HMBC spectra resulted in the assignment of these signals to homocitric acid (**Figure 4**), which is not common in plants because it is involved in the lysine biosynthesis in fungi (*32*). A possible explanation for its presence in plants is that they might receive variable amounts of carbon from mycorrhizal fungi. This idea is supported, for example, by recent studies that proved this in some green orchids (*33*). The term "mixotroph" could be applied to these plants, considering them as autotrophic organisms that combine their photosynthetic capacity with a partial heterotrophy for carbon sources. The presence of mycorrhizal fungi in *V. planifolia* has already been reported (*34*), providing an explanation for the detection of homocitric acid in developing green pods. An ion corresponding to homocitric acid has also

been detected in LS-MS chromatograms of our green pod extracts (data not shown).

Resonances at δ 2.20 (m), 0.92 (d, J = 7 Hz), and 0.86 (d, J = 7.0 Hz) were assigned to bis[4-(β-D-glucopyranosyloxy)-benzyl]-2-isopropyltartrate (glucoside A). Signals at δ 1.90 (m), 1.35 (m), 1.10 (m), 0.84 (d, J = 7.0 Hz), and 0.77 (t, J = 15.4 Hz) were assigned to bis[4-(β-D-glucopyranosyloxy)-benzyl]-2-(2-butyl)-tartrate (glucoside B). The assignments were performed on the basis of the analysis of 2D NMR spectra (J-resolved, COSY, and HMBC). These two glucosides have been reported in vanilla green pods (*13*−*15*). Analysis in MS/MS spectrometry confirmed the presence of ions of these glucosides in the green pod extract (data not shown). **Table 1** summarizes all compounds identified in ^{1}H NMR spectra with the chemical shifts and the coupling constants of the detected signals.

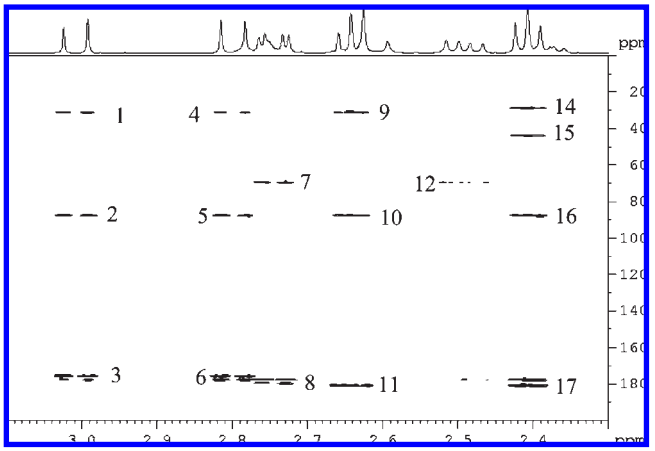

Figure 4. HMBC spectra of *V. planifolia* green pods: 1, correlation of H-2 and C-4 of homocitric acid; 2, correlation of H-2 and C-3 of homocitric acid; 3, correlation of H-2 and two carbonyl groups (C-1 and C-7) of homocitric acid; 4, correlation of H-2′ and C-4 of homocitric acid; 5, correlation of H-2′ and C-3 of homocitric acid; 6, correlation of H-2′ and two carbonyl groups (C-1 and C-7) of homocitric acid; 7, correlation of H-3 and C-2 of malic acid; 8, correlation of H-3 and two carbonyl groups (C-1 and C-4) of malic acid; 9, correlation of H-5 and C-4 of homocitric acid; 10, correlation of H-5 and C-3 of homocitric acid; 11, correlation of H-5 and a carbonyl group of homocitric acid (C-6); 12, correlation of H-3′ and C-2 of malic acid; 13, correlation of H-3′ and two carbonyl groups (C-1 and C-4) of malic acid; 14, correlation of H-4 and C-5 of homocitric acid; 15, correlation of H-4 and C-2 of homocitric acid; 16, correlation of H-4 and C-3 of homocitric acid; 17, correlation of H-4 and two carbonyl groups of homocitric acid (C-6 and C-7).

Multivariate Data Analysis. PCA is an unsupervised clustering method requiring no knowledge of the data set and acts to reduce the dimensionality of multivariate data while preserving most of the variance within it (*35*). The principal components can be displayed graphically as a score plot. This plot is useful for observing any groupings in the data set. PCA models are constructed using all of the samples in the study. Coefficients by which the original variables must be multiplied to obtain the PC are called loadings. The numerical value of a loading of a given variable on a PC shows how much the variable has in common with that component (*36*). Thus, for NMR data, loading plots can be used to detect the metabolites responsible for the separation in the data. Generally, this separation takes place in the first two principal components (PC1 and PC2).

Although an accession-based variation along PC2 can be observed mainly in the early stages of development, application of PCA to our data resulted in a good separation of young pods (three, four, and five months old) from older ones (six, seven, and eight months old) by PC1 (**Figure 5A**). However, from a more general point of view, we can consider the trend of developing pods as moving in the score plot from positive to negative values of PC1. This separation is due to the signals of sugars, organic acids, and phenolic compounds (**Figure 5B**). Younger pods contain more glucose, malic acid, homocitric acid, and glucosides A and B, while older pods contain more sucrose, glucovanillin, vanillin, *p*-hydroxybenzaldehyde glucoside, and *p*-hydroxybenzaldehyde. This can be concluded from the loading plot of PC1, in which a higher PC1 value indicates higher signals at δ 4.60 (glucose), 2.74 (malic acid), 2.41 (homocitric acid), 0.92 (glucoside A), and 0.77 (glucoside B) and lower signals at δ 7.62 (glucovanillin), 7.52 (vanillin), 7.94 (*p*-hydroxybenzaldehyde glucoside), 7.85 (*p*-hydroxybenzaldehyde), and 5.44 (sucrose). An ANOVA test was performed for the signals that were not in a crowded region of the ^{1}H NMR spectra. Some compounds significantly change during pod development (**Table 1**).

PLS-DA analysis was also applied to the ^{1}H NMR data. Two classes based on the developmental stage of green pods were created from the samples: young (from 3- to 5-month-old pods) and older (from 6- to 7-month-old ones). Despite a variation observed in the early stages of development, a distinct separation between these two classes was observed from the score plot of PLS component 1 versus 2 (**Figure 6**). The result of the cross-validated (CV)-ANOVA suggested that the model was highly significant, having a p value of 1.07×10^{-8} (*37*). Analysis of the loading column plot of this PLS-DA model gave identical results to those obtained from the PCA model. Younger pods were also

Table 1. ^{1}H Chemical Shifts (δ), Coupling Constants (Hz), Age Effect, and Results of ANOVA Test of *V. planifolia* Pod Metabolites Identified by References and Using 1D and 2D NMR Spectra (Methanol-d_4−KH_2PO_4 in D_2O at pH 6.0)

compound	chemical shifts and coupling constants	age effect[a]	ANOVA test (p value)[b]
glucovanillin	δ 9.82 (s), 7.62 (dd, J = 8.5, 2.0 Hz), 7.57 (d, J = 2.0 Hz), 7.34 (d, J = 8.5 Hz), 5.19 (d, J = 7.8 Hz), 3.95 (s)	+	2.6×10^{-4}
vanillin	δ 9.73 (s), 7.52 (dd, J = 8.5, 2.0 Hz), 7.49 (d, J = 2.0 Hz), 7.04 (d, J = 8.5 Hz), 3.93 (s)	+	5.3×10^{-4}
p-hydroxybenzaldehyde glucoside	δ 9.84 (s), 7.94 (d, J = 9 Hz), 7.29 (d, J = 9 Hz), 5.01 (d, J = 8 Hz)	+	2.7×10^{-3}
p-hydroxybenzaldehyde	δ 9.75 (s), 7.85 (d, J = 8.5 Hz), 7.04 (d, J = 8.5 Hz)	+	5.6×10^{-4}
sucrose	δ 5.4 (d, J = 3.5 Hz), 4.17 (d, J = 7.8 Hz)	+	4.4×10^{-4}
glucose	δ 5.18 (d, J = 3.7 Hz), 4.57 (d, J = 7.9 Hz)	−	nc[c]
malic acid	δ 4.29 (dd, J = 7.8, 3.5 Hz), 2.74 (dd, J = 16.0, 3.5 Hz), 2.49 (dd, J = 16.0, 8.0 Hz)	−	7.4×10^{-5}
homocitric acid	δ 3.01 (d, J = 16.0 Hz), 2.8 (d, J = 16.0 Hz), 2.64 (dd, J = 9.5, 7.5 Hz), 2.41 (dd, J = 9.5, 7.5 Hz)	−	4.9×10^{-4}
bis[4-(β-D-glucopyranosyloxy)-benzyl]-2-isopropyltartrate (glucoside A)	δ 2.20 (m), 0.92 (d, J = 7.0 Hz), 0.86 (d, J = 7.0 Hz)	−	nc
bis[4-(β-D-glucopyranosyloxy)-benzyl]-2-(2-butyl)tartrate (glucoside B)	δ 1.90 (m), 1.35 (m), 1.10 (m), 0.84 (d, J = 7.0 Hz), 0.77 (t, J = 15.0 Hz)	−	nc

[a] (+) increase with development and (−) decrease with development. [b] On the basis of six classes. [c] nc = not calculated due to the overcrowded signal region.

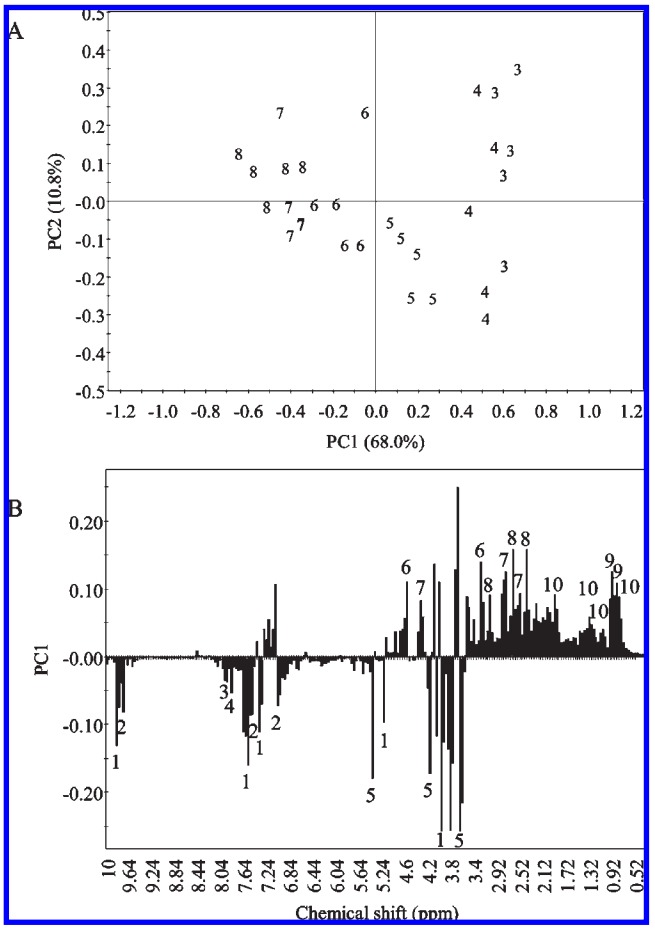

Figure 5. (A) Score plot (PC1 versus PC2) and (B) loading column plot of PCA results obtained from ^{1}H NMR spectra of five accessions of *V. planifolia* pods of (3) 3 months old, (4) 4 months old, (5) 5 months old, (6) 6 months old, (7) 7 months old, and (8) 8 months old. Assignments in B: 1, glucovanillin; 2, vanillin; 3, *p*-hydroxybenzaldehyde glucoside; 4, *p*-hydroxybenzaldehyde; 5, sucrose; 6, glucose; 7, malic acid; 8, homocitric acid; 9, bis[4-(β-D-glucopyranosyloxy)-benzyl]-2-isopropyltartrate (glucoside A); 10, bis[4-(β-D-glucopyranosyloxy)-benzyl]-2-(2-butyl)tartrate (glucoside B).

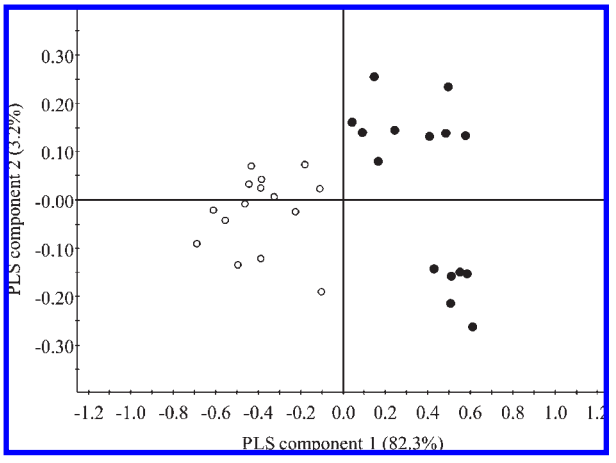

Figure 6. Score plot (PLS components 1 and 2) of PLS-DA results obtained from ^{1}H NMR spectra of five accessions of *V. planifolia* pods of (●) 3, 4, and 5 months old and (○) 6, 7, and 8 months old.

shown here to contain more glucose, malic acid, homocitric acid, and glucosides A and B, and older pods were shown here to contain more sucrose, glucovanillin, vanillin, *p*-hydroxybenzaldehyde glucoside, and *p*-hydroxybenzaldehyde.

The *V. planifolia* green pod is a typical fruit in which phenolic compounds (flavor precursors) are produced during the development. The fresh weight of the pod has been observed to increase at a high rate during the early stages (0−3 months) and then slow down during later stages (3−8 months) (*17*). Thus, an increase of sucrose levels (3−8 months) might be more likely due to the conversion of starch rather than an accumulation from the plant, as observed in other ripening fruits, such as mango (*38*) and banana (*39*). Nevertheless, detection of starch was not possible with the method performed in our study to confirm this hypothesis.

LC−MS Analysis. A more targeted approach to phenolic compounds was performed by analyzing the samples in LC−MS. These compounds were identified on the basis of their retention times, UV-DAD spectra, and electrospray mass spectra and quantified according to the calibration table. When chromatograms of nontreated and β-glucosidase-treated extracts were compared and the increase in areas of the aglycone peaks after hydrolysis was quantified, it was possible to estimate the aglycone/glycoside ratio of each compound in the developing vanilla pods (**Figure 7**). The major compounds identified in LC−MS were vanillin and *p*-hydroxybenzyl alcohol. Coinciding with previous reports (*17*), vanillin was detected as both an aglycone and a glycoside in green pods of *V. planifolia* (**Figure 7D**), appearing first in 4-month-old pods and increasing gradually to 340 μmol of total vanillin/g of dry weight at 8 months. At this developmental stage, the free vanillin amount represents 24% of the total vanillin content. This pattern of vanillin content had already been observed by different authors (*13*−*15*, *17*, *40*), despite higher ratios of aglycone/glucoside observed in our study. Residual endogenous glucosidase activity during the LC−MS protocol extraction could explain these higher ratios (**Figure 7**). Nevertheless, some compounds are detected only after hydrolysis by the exogenous glucoside (panels **A** and **G** of **Figure 7**). Furthermore, comparative quantification performed by either LC−MS or ^{1}H NMR analysis gave similar results (**Table 2**). Thus, the ratio observed in our experiments might be correct and reveal the composition of both aglycone and glucoside in the green pods. Further experiments are in progress in our laboratory and could explain the controversial results on the ratio of aglycone/glucoside obtained from the literature.

The *p*-hydroxybenzyl alcohol was detected only after hydrolysis with β-glucosidase, meaning that it was only present as a glucoside in green pods (**Figure 7A**). Its concentration decreased from 3 to 5 months (192−64 μmol/g of dry weight), increased at 6 months (93 μmol/g of dry weight), and decreased again at 8 months. Kanisawa et al. obtained a similar pattern of variation of *p*-hydroxybenzyl alcohol concentration during the development of green pods (*13*−*15*). Several studies have reported glucovanillin, *p*-hydroxybenzyl alcohol glucoside, and the glucosides A and B as the major phenolics in developing vanilla green pods (*13*−*15*), while some minor glucosides were reported after hydrolysis with β-glucosidase. In our study, the vanilla green pod content of *p*-hydroxybenzoic acid, *p*-hydroxybenzaldehyde, protocatechuic acid, protocatechualdehyde, vanillyl alcohol, acetovanillon, vanillic acid, caffeic acid, ferulic acid, and *p*-coumaric acid was also investigated. However, we did not detect acetovanillon nor protocatechuic, caffeic, ferulic, and *p*-coumaric acids in vanilla green pods probably because their levels of concentration are too low. More amounts of the sample and purification of phenolic compounds might be necessary to detect them. Nevertheless, protocatechualdehyde, *p*-hydroxybenzaldehyde, vanillic acid, *p*-hydroxybenzoic acid, and vanillyl alcohol were identified

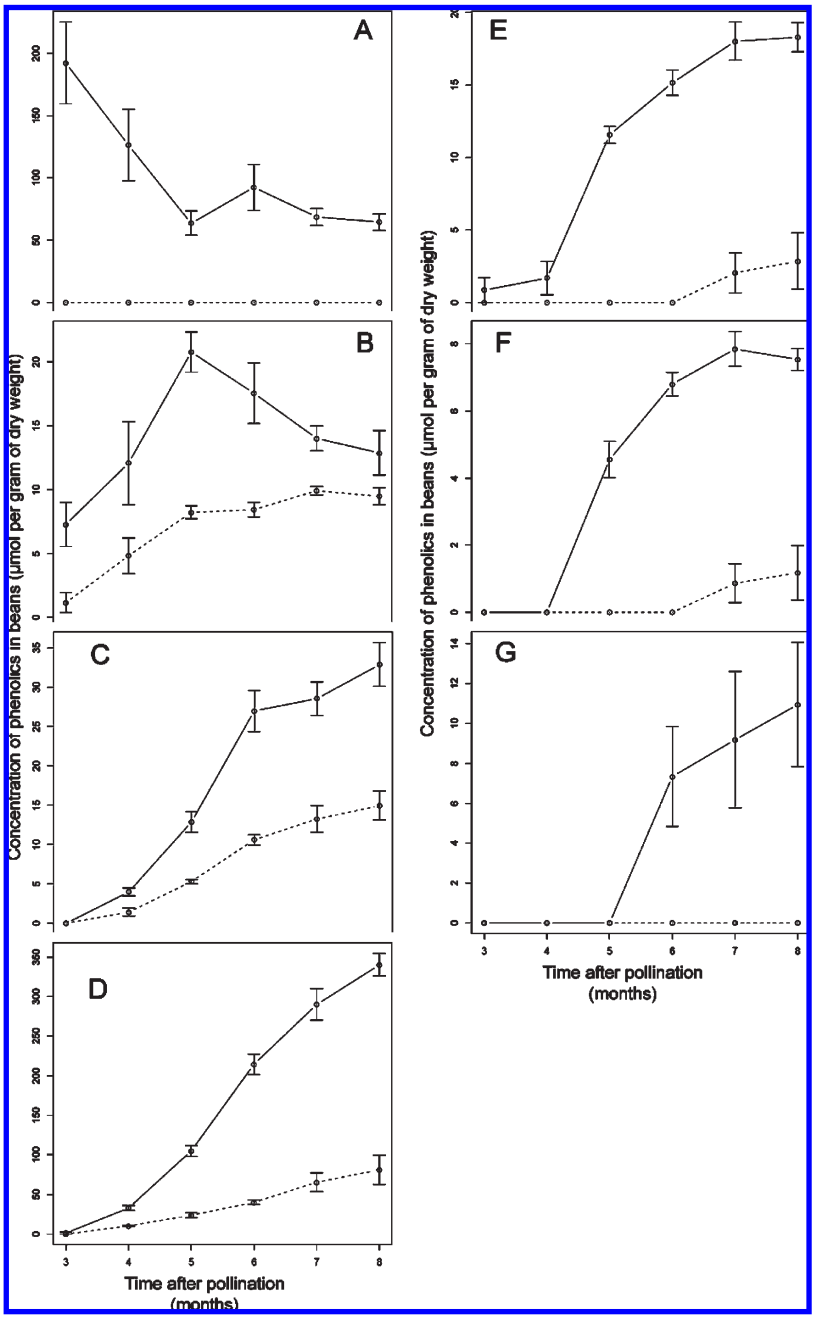

Figure 7. Amounts of phenolic compounds in pod extracts as a function of time after pollination: (A) *p*-hydroxybenzylalcohol, (B) protocatechualdehyde, (C) *p*-hydroxybenzaldehyde, (D) vanillin, (E) vanillic acid, (F) *p*-hydroxybenzoic acid, and (G) vanillyl alcohol. (- - -) Aglycon form and (—) total form after hydrolysis by β-glucosidase. Errors bars represents standard error (SE).

Table 2. Quantification in Micromoles per Gram of Green Pod Dry Weight of the Different Forms of Vanillin during Development of the Green Pod, Performed by Either Using a Calibration Curve in the LC−MS Method or Integration of a Single Peak in the ^{1}H NMR Spectra

	LC−MS[a]		^{1}H NMR[a]	
months	vanillin (free)	vanillin (total)	glucovanillin	vanillin
3	0	1.4 ± 0.9	0.9 ± 0.1	0.8 ± 0.1
4	10.4 ± 0.6	32.9 ± 3.2	19.1 ± 3.3	8.5 ± 1.1
5	23.9 ± 3.2	105.1 ± 7.0	67.1 ± 5.4	28.0 ± 0.4
6	40.1 ± 2.7	214.3 ± 13.3	140.1 ± 10.1	49.0 ± 2.4
7	65.5 ± 12.2	289.8 ± 20.1	198.7 ± 6.5	73.7 ± 5.1
8	81.2 ± 18.5	339.8 ± 14.2	271.0 ± 25.8	74.2 ± 8.3

[a] Data are shown as the mean of concentration ± SE, with $n = 5$.

before and/or after hydrolysis with β-glucosidase (panels **B**, **C**, **E**, **F**, and **G** in **Figure 7**, respectively). While protocatechualdehyde increased from 3 to 5 months and then decreased until 8 months (**Figure 7B**), other compounds such as *p*-hydroxybenzaldehyde, vanillic acid, *p*-hydroxybenzoic acid, and vanillyl alcohol showed an increase of concentration all along pod development. The increase of *p*-hydroxybenzaldehyde proved to be very similar to vanillin as both an aglycone as well as the sum of aglycones and glucosides. **Figure 7** also shows that most of the phenolic compounds detected in mature pods were present as glucosides, underlining the need of the curing process to ensure the release of the aglycones to increase the vanilla flavor.

In this study, ^{1}H NMR combined with different multivariate statistical methods was successfully applied to analyze developing

V. planifolia green pods, showing the trends of the variation of their metabolic profile throughout this process, based on the analysis of the content of various distinctive metabolites. In addition, a more targeted approach on phenolic compound content was performed using LC−MS. These results provide clues and may serve as a base to clarify the biosynthetic pathway of vanillin in developing vanilla pods and to contribute to control their quality. Studies on the expression of various enzymes involved in this biosynthesis are in progress in our laboratory.

ACKNOWLEDGMENT

The authors thank Bertrand Côme (Coopérative Provanille, La Réunion) who provided the biological material, Frédéric Chiroleu and Sylvaine Jégo for their help in the statistical analysis, and Marie-Paule Gonthier, Evelyne Tarnus, and Erica G. Wilson for their comments.

LITERATURE CITED

(1) Walton, N. J.; Mayer, M. J.; Narbad, A. Molecules of interest—Vanillin. *Phytochemistry* **2003**, *63*, 505–515.

(2) Korthou, H.; Verpoorte, R. Vanilla. In *Flavours and Fragrances—Chemistry, Bioprocessing and Sustainability*; Berger, R. G., Ed.; Springer: Berlin, Germany, 2007; pp 203−217.

(3) Ranadive, A. S. Vanilla—Cultivation, curing, chemistry, technology and commercial products. In *Spices, Herbs and Edible Fungi*; Charalambous, G., Ed.; Elsevier: Amsterdam, The Netherlands, 1994; pp 517−576.

(4) Dignum, M. J. W.; Kerler, J.; Verpoorte, R. Vanilla curing under laboratory conditions. *Food Chem.* **2002**, *79*, 165–171.

(5) Pérez-Silva, A. Contribution à l'étude de la génèse des composés d'arôme au cours du procédé mexicain de transformation de la vanille (*Vanilla planifolia* G. Jackson). Université de Montpellier II, Montpellier, France, **2006**.

(6) Schwarz, B.; Hofmann, T. Identification of novel orosensory active molecules in cured vanilla beans (*Vanilla planifolia*). *J. Agric. Food Chem.* **2009**, *57*, 3729–3737.

(7) Odoux, E. Changes in vanillin and glucovanillin concentrations during the various stages of the process traditionally used for curing *Vanilla fragrans* beans in Réunion. *Fruits* **2000**, *55*, 119–125.

(8) Dignum, M. J. W.; van der Heijden, R.; Kerler, J.; Winkel, C.; Verpoorte, R. Identification of glucosides in green beans of *Vanilla planifolia* Andrews and kinetics of vanilla β-glucosidase. *Food Chem.* **2004**, *85*, 199–205.

(9) Leong, G.; Archavlis, A.; Derbesy, M. Research on the glucoside fraction of the vanilla bean. *J. Essent. Oil Res.* **1989**, *1*, 33–41.

(10) Leong, G.; Uzio, R.; Derbesy, M. Synthesis, identification and determination of glucosides present in green vanilla beans (*Vanilla fragrans* Andrews). *Flavour Fragrance J.* **1989**, *4*, 163–167.

(11) Odoux, E. Glucosylated aroma precursors and glucosidase(s) in vanilla bean (*Vanilla planifolia* G. Jackson). *Fruits* **2006**, *61*, 171–184.

(12) Negishi, O.; Ozawa, T. Determination of hydroxycinnamic acids, hydroxybenzoic acids, hydroxybenzaldehydes, hydroxybenzyl alcohols and their glucosides by high-performance liquid chromatography. *J. Chromatogr.* **1996**, *756*, 129–136.

(13) Tokoro, K.; Kawahara, S.; Amano, A.; Kanisawa, T.; Indo, M. Glucosides in vanilla beans and changes of their contents during maturation. In *Flavour Science and Technology*; Bessiere, Y., Thomas, A. F., Eds.; John Wiley and Sons: New York, 1990; pp 73−76.

(14) Kanisawa, T. Flavor development in vanilla beans. *Kouryou* **1993**, *180*, 113–123.

(15) Kanisawa, T.; Tokoro, K.; Kawahara, S. Flavor development in the beans of *Vanilla planifolia*. In *Proceedings of the International Symposium*; Kurihara, K., Suzuki, N., Ogawa, H., Eds.; Springer: Tokyo, Japan, 1994; pp 268−270.

(16) Sagrero-Nieves, L.; Schwartz, S. J. Phenolic content of *Vanilla planifolia* as affected by harvest period. *J. Food Compos. Anal.* **1988**, *1*, 362–365.

(17) Brodelius, P. E. Phenylpropanoid metabolism in *Vanilla planifolia* Andr. V. High-performance liquid-chromatographic analysis of phenolic glycosides and aglycones in developing fruits. *Phytochem. Anal.* **1994**, *5*, 27–31.

(18) Verpoorte, R.; Choi, Y. H.; Kim, H. K. NMR-based metabolomics at work in phytochemistry. *Phytochem. Rev.* **2007**, *6*, 3–14.

(19) Ward, J. L.; Harris, C.; Lewis, J.; Beale, M. H. Assessment of ^{1}H NMR spectroscopy and multivariate analysis as a technique for metabolite fingerprinting of *Arabidopsis thaliana*. *Phytochemistry* **2003**, *62*, 949–957.

(20) Choi, Y. H.; Kim, H. K.; Hazekamp, A.; Erkelens, C.; Lefeber, A. W. M.; Verpoorte, R. Metabolic differentiation of *Cannabis sativa* cultivars using ^{1}H NMR spectroscopy and principal component analysis. *J. Nat. Prod.* **2004**, *67*, 953–957.

(21) Choi, Y. H.; Tapias, E. C.; Kim, H. K.; Lefeber, A. W. M.; Erkelens, C.; Verhoeven, J. T. J.; Brzin, J.; Zel, J.; Verpoorte, R. Metabolic discrimination of *Catharanthus roseus* leaves infected by phytoplasma using ^{1}H-NMR spectroscopy and multivariate data analysis. *Plant Physiol.* **2004**, *135*, 2398–2410.

(22) Choi, H.-K.; Choi, Y. H.; Verberne, M.; Lefeber, A. W. M.; Erkelens, C.; Verpoorte, R. Metabolic fingerprinting of wild type and transgenic tobacco plants by ^{1}H NMR and multivariate analysis technique. *Phytochemistry* **2004**, *65*, 857–864.

(23) Kim, H. K.; Choi, Y. H.; Erkelens, C.; Lefeber, A. W. M.; Verpoorte, R. Metabolic fingerprinting of *Ephedra* species using ^{1}H-NMR spectroscopy and principal component analysis. *Chem. Pharm. Bull.* **2005**, *53*, 105–109.

(24) Abdel-Farid, I. B.; Kim, H. K.; Choi, Y. H.; Verpoorte, R. Metabolic characterization of *Brassica rapa* leaves by NMR spectroscopy. *J. Agric. Food Chem.* **2007**, *55*, 7936–7943.

(25) Choi, H.-K.; Yoon, J.-H. Metabolomic profiling of *Vitis vinifera* cell suspension culture elicited with silver nitrate by ^{1}H NMR spectrometry and principal components analysis. *Process Biochem.* **2007**, *42*, 271–274.

(26) Brescia, M. A.; Caldarola, V.; De Giglio, A.; Benedetti, D.; Fanizzi, F. P.; Sacco, A. Characterization of the geographical origin of Italian red wines based on traditional and nuclear magnetic resonance spectrometric determinations. *Anal. Chim. Acta* **2002**, *458*, 177–186.

(27) Charlton, A. J.; Farrington, W. H. H.; Brereton, P. Application of ^{1}H NMR and multivariate statistics for screening complex mixtures: Quality control and authenticity of instant coffee. *J. Agric. Food Chem.* **2002**, *50*, 3098–3103.

(28) Vogels, J. T. W. E.; Terwel, L.; Tas, A. C.; van den Berg, F.; Dukel, F.; van der Greef, J. Detection of adulteration in orange juices by a new screening method using proton NMR spectroscopy in combination with pattern recognition techniques. *J. Agric. Food Chem.* **1996**, *44*, 175–180.

(29) Duarte, I.; Barros, A.; Belton, P. S.; Righelato, R.; Spraul, M.; Humpfer, E.; Gil, A. M. High-resolution nuclear magnetic resonance spectroscopy and multivariate analysis for the characterization of beer. *J. Agric. Food Chem.* **2002**, *50*, 2475–2481.

(30) Jahangir, M.; Kim, H. K.; Choi, Y. H.; Verpoorte, R. Metabolomic response of *Brassica rapa* submitted to pre-harvest bacterial contamination. *Food Chem.* **2008**, *107*, 362–368.

(31) R Development Core Team. R: A language and environment for statistical computing, **2006**, http///www.r-project.org/.

(32) Zabriskie, T. M.; Jackson, M. D. Lysine biosynthesis and metabolism in fungi. *Nat. Prod. Rep.* **2000**, *17*, 85–97.

(33) Selosse, M.-A.; Roy, M. Green plants that feed on fungi: Facts and questions about mixotrophy. *Trends Plant Sci.* **2009**, *14*, 64–70.

(34) Porras-Alfaro, A.; Bayman, P. Mycorrhizal fungi of vanilla: Diversity, specificity and effects on seed germination and plant growth. *Mycologia* **2007**, *99*, 510–525.

(35) Goodacre, R.; Shann, B.; Gilbert, R. J.; Timmins, E. M.; McGovern, A. C.; Alsberg, B. K.; Kell, D. B.; Logan, N. A. Detection of the dipicolinic acid biomarker in *Bacillus* spores

using curie-point pyrolysis mass spectrometry and Fourier transform infrared spectroscopy. *Anal. Chem.* **2000**, *72*, 119–127.

(36) Massart, D.; Vandeginste, B.; Deming, S.; Michotte, Y.; Kauffman, L. *Chemometrics: A Textbook*; Elsevier: Amsterdam, The Netherlands, 1988.

(37) Eriksson, L.; Trygg, J.; Wold, S. CV-ANOVA for significance testing of PLS and OPLS models. *J. Chemom.* **2008**, *22*, 594–600.

(38) Castrillo, M.; Kruger, N. J.; Whatley, F. R. Sucrose metabolism in mango fruit during ripening. *Plant Sci.* **1992**, *84*, 45–51.

(39) Zhang, P.; Whistler, R. L.; BeMiller, J. N.; Hamaker, B. R. Banana starch: Production, physicochemical properties, and digestibility—A review. *Carbohydr. Polym.* **2005**, *59*, 443–458.

(40) Ranadive, A. S.; Szkutnica, K.; Guerrera, J. G.; Frenkel, C. Vanilla biosynthesis in vanilla beans. In *Proceedings of the IX International Congress of Essential Oils*; Singapore, 1983; pp 147−154.

Received May 7, 2009. Revised manuscript received July 22, 2009. Accepted August 01, 2009. This study was supported by the Regional Council of La Réunion and the European Social Fund.