Contributions to the quality control of two crops of economic importance : hops and yerba mate

Wilson, E.G.

chapter 8

Classification of Ilex Species Based on Metabolomic Fingerprinting Using Nuclear Magnetic Resonance and Multivariate Data Analysis

Young Hae Choi [a], Sarah Sertic [a], Hye Kyong Kim [a], Erica G. Wilson [a], Filippos Michopoulos [a], Alfons W. M.Lefeber [b], Cornelis Erkelens [b], Sergio D. Prat Kricun [c] Robert Verpoorte [a]

[a] *Division of Pharmacognosy, Leiden University.*

[b] *Division of NMR, Leiden University.*

[c] *EEA Cerro Azul INTA, Misiones, Argentina.*

Published in:
Journal of Agricultural and Food Chemistry, 53, 1237-1245 (2005)

Abstract

The metabolomic analysis of 11 *Ilex* species, *I. paraguariensis* var. *paraguariensis*, *I. argentina*, *I. brasiliensis*, *I. brevicuspis*, *I. dumosa* var. *dumosa*, *I. dumosa* var. *guaranina*, *I. integerrima*, *I. microdonta*, *I. pseudobuxus*, *I. taubertiana*, and *I. theezans* was carried out by NMR spectrometry and multivariate data analysis. The analysis using principal component analysis (PCA) and classification of the ^{1}H NMR spectra showed a clear discrimination of those samples based on the metabolites present in the organic and aqueous fractions. The major metabolites that contribute to the discrimination are arbutin, caffeine, phenylpropanoids, and theobromine. Among those metabolites, arbutin, that has not been reported yet as a constituent of *Ilex* species was found to be a biomarker for *I. argentina*, *I. brasiliensis*, *I. brevicuspis*, *I. integerrima*, *I. microdonta*, *I. pseudobuxus*, *I. taubertiana*, and *I. theezans*. This reliable method based on the determination of a large number of metabolites makes it possible to perform chemotaxonomical analysis of *Ilex* species.

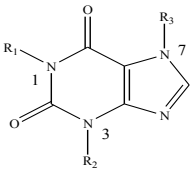
8.1 Introduction

The genus *Ilex* (Aquifoliaceae) comprises more than 500 species of dioecious trees and shrubs distributed throughout temperate and tropical regions of the world. The main areas of extant diversification are East Asia and South America. Among the species, Yerba mate, a tea-like infusion of *Ilex paraguariensis* St. Hill., is drunk in many parts of South America (Graham, 1984; Giberti,1989). It is also consumed socially in the Middle East by the Druze of Lebanon, Syria and the Golan Heights in northern Israel. An average of 300,000 t of mate is produced each year in South America. In addition to its standing as a popular beverage, mate is used as a tonic, diuretic, and a stimulant to reduce fatigue and aid gastric function in herbal medicine systems throughout South America and other parts of the world (Gonzalez *et al*, 1993;Gugliucci *et al*, 1996; Filip *et al*, 2001).

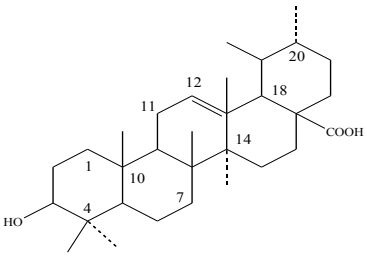
Studies concerning the metabolites present in *I. paraguariensis*, differ according to the material analyzed, that is, fresh plant material or mate, the commercial product. A caffeine content of 0.9-2.2% was determined in leaves, and this was found to depend on the age of the leaves (Graham, 1984; Bertoni *et al.,* 1992). Clifford and Ramirez-Martinez examined five commercial samples of two types of yerba mate of South American origin, using high performance liquid chromatography (HPLC) finding the caffeine content to be between 0.89 and 1.73% and the theobromine to be between

0.45 and 0.88%, with very small quantities of other purines (Clifford & Ramirez-Martínez, 1990).

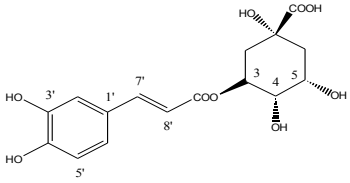
However, the caffeine content of commercial samples can be very variable since it depends greatly on the industrial treatment of the raw material (Schmalko *et al.*, 2001). In plant material, Reginatto *et al.* found the methyl xanthines content of *Ilex paraguariensis* var. *paraguariensis*, to be ~ 1.8% in old leaves using HPLC analysis (Reginatto *et al*, 1999). The caffeine content was estimated ~ 0.65% in old leaves and up to 1.4% in young leaves, while theobromine varied from 0.02% (in old leaves) to 0.27% in young leaves. Apart from caffeine analogues, caffeic acid, chlorogenic acid, and the three dicaffeoylquinic acids were found in all the species. Among the substitutes or adulterants assayed in recent studies, *I. brevicuspis* showed the highest total caffeoyl derivatives content (1.9%) followed by *I. argentina* (0.73%) and *I. pseudobuxus* (0.67%) (Filip *et al.*, 2001).



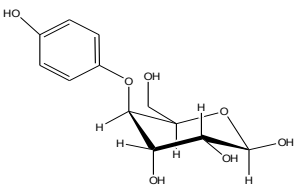
$R_1 = R_2 = R_3 = CH_3$, Caffeine
$R_1 = H, R_2 = R_3 = CH_3$, Theobromine
$R_1 = R_2 = CH_3, R_3 = H$, Theophylline



Ursolic acid



Chlorogenic acid



Arbutin(hydroquinone-β-glucopyranoside

Other chlorogenic acids that were found were ferulic acid, *p*-coumaric acid, caffeoylquinic acids, feruloylquinic acids, *p*-coumaroylquinic acids, and caffeoylferuloylquinic acids. Rutin, quercetin, and kaempferol were also found as main flanovoids. *Ilex pseudobuxus* showed the highest content of rutin (0.03%), which is quite below the value obtained for *I. paraguariensis* (0.06%), while similar quercitin and kaempferol contents were

found in the other *Ilex* species (Filip *et al.*, 2001). Ursolic acid was isolated as the main triterpene, together with amyrin. The presence of triterpenoidal saponins, such as metasaponins was reported (Kraemer *et al.*, 1996; Taketa *et al.*, 2000).

Table 8-1 Ilex species evaluated in this study (22).

Species	Region and year of seed-collection (Number of voucher specimen)
Ilex argentina Lillo	Cerro San Javier, Tucumán, Argentina, 1991 (109) Acheral, Tucumán, Argentina, 1991 (111) Conception, Tucumán, Argentina, 1991 (112) Quebrada de San Lorenzo, Salta, Argentina, 1995
Ilex brasiliensis (Spreng) Loes.	Rio Branco do Sul, Paraná, Brazil, 1990 (59) Pto. Esperanza, Misiones, Argentina, 1990 (221) Reserva Biologica de Limoy, Paraguay, 1997 (226) Nueva Esperanza, Paraguay, 1997 (230)
Ilex brevicuspis Reissek	San Pedro, Misiones, Argentina, 1987 (4) Canoinhas, Brazil, 1989 (15) Clevelandia, Paraná, Brazil, 1991 (94) Veranopolis, Río Grande do Sul, Brazil, 1992 (119)
Ilex dumosa var. *dumosa* Reissek	Campo Viera, Misiones, Argentina, 1989 (7) Canoinhas, Santa Catarina, Brazil, 1989 (13) Tijucas do Sul, Paraná, Brazil, 1990 (55) Campo Bom, Río Grande do Sul, Brazil, 1992 (113)
Ilex dumosa var. *guaranina* Loes.	Pto. Esperanza, Misiones, Argentina, 1996 (222) Reserva Biologica de Limoy, Paraguay, 1997 (227) Hernandarias, Paraguay, 1998 (235) Hernandarias, Paraguay, 1998 (243)
Ilex integerrima (Vellozo) Reissek	Tijucas do Sul, Paraná, Brazil, 1990 (56) San Mateo do sul, Paraná, Brazil, 1990 (69) Iratí, Paraná, Brazil, 1990 (72) Teixeira Soares, Paraná, Brazil, 1990 (73)
Ilex microdonta Reissek	Sao Francisco de Paula, Río Grande do Sul, Brazil, 1992 (121) Sao Francisco de Paula, Río Grande do Sul, Brazil, 1992 (121a) Sao Francisco de Paula, Río Grande do Sul, Brazil, 1992 (121b) Parque Nacional Aparados da Serra, Río Grande do Sul, Brazil,

Species	Region and year of seed-collection (Number of voucher specimen)
Ilex paraguariensis var. *paraguariensis* St. -Hill.	Chapecó, Santa Catarina, Brazil, 1989 (28) San Antonio, Misiones, Argentina, 1989 (45) Teixeira Soares, Paraná, Brazil, 1990 (74) Ijuí, Río Grande do Sul, Brazil, 1993 (133)
Ilex pseudobuxus Reissek	Pontal do Sul, Paraná, Brazil, 1990 (67) Campo Bom, Río Grande do Sul, Brazil, 1992 (114) Torres, Río Grande do Sul, Brazil, 1992 (131) Tramandaí, Río Grande do Sul, Brazil, 1992 (132)
Ilex taubertiana Loes.	Sao Francisco de Paula, Río Grande do Sul, Brazil, 1992 (124) Sao Francisco de Paula, Río Grande do Sul, Brazil, 1992 (124a) Sao Francisco de Paula, Río Grande do Sul, Brazil, 1992 (124b) Sao Francisco de Paula, Río Grande do Sul, Brazil, 1992 (124c)
Ilex theezans Reissek.	Major Viera, Río Grande do Sul, Brazil, 1989 (16) San Antonio, Misiones, Argentina, 1989 (46) Tijucas do Sul, Paraná, Brazil, 1990 (54) Veranopolis, Río Grande do Sul, Brazil, 1992 (118)

Ilex paraguariensis has many local congeneric substitutes that grow in the same habitat, the most common of which - *I. brevicuspis* as *I. argenina, I. brasiliensis, I. dumosa, I. integerrima, I. microdonta, I. pseudobuxus, I. taubertiana,* and *I. theezans* - were chosen for this study.

They are often used as substitutes or adulterants of *Ilex paraguariensis* (Yerba mate) (Graham, 1984; Giberti, 1989). Despite the extensive studies on the chemical composition of *Ilex* species, previous results are limited to the differentiation of particular metabolite analysis. For example, the main differences reported in previous references were a higher content of caffeoyl derivatives and flavonoids, and the presence of caffeine in *Ilex paraguariensis* and its absence, or presence in very low amounts, in other species. However, the classification or discrimination of each *Ilex* species is still practically impossible when the classical method of analyzing a single group of metabolites is applied and no method that could potentially contribute to their

detection in mixtures has been published. For the reliable differentiation of *Ilex* species, a systematic method involving a wide variety of metabolites (metabolomic profiling) could be a very useful contribution to these issues.

The term "metabolome" has been used to describe the observable chemical profile or fingerprint of the metabolites in whole tissues (Ott *et al*., 2003). To obtain the most complete metabolomic profile it is fundamental to use a wide spectrum analytical technique that is rapid, reproducible, stable over time and requires only a very simple sample preparation. NMR is one of the techniques that meets those requirements. In the last decade, a number of techniques have been devised to develop NMR spectroscopy as a fingerprinting tool for the quality assessment of natural products such as food and medicinal plants. Multivariate or pattern recognition techniques such as the well-described principal component analysis (PCA) are important tools for the analysis of the data obtained by NMR. Recently, NMR in combination with PCA has been applied to the metabolomic profiling of several kinds of wines (Brescia *et al*., 2002), coffees (Charlton *et al*, 2002), juices (Vogels *et al*., 1996), beers (Duarte *et al*., 2002), and some plants (Forveille *et al*., 1996; Nord *et al*., 2001).

In this study we report a NMR spectroscopic method, coupled to principal component analysis (PCA) for the metabolomic analysis of ten *Ilex* species including two varieties. Based on this data, classification and discrimination is performed for the ten *Ilex* species. This leads to a clear differentiation of *Ilex* species based on a variety of metabolites.

8.2 Materials and methods

8.2.1 Materials:

Dried plant material of 11 *Ilex* species was provided by the Estación Experimental Agraria of Cerro Azul (INTA) (Misiones, Argentina). The samples were harvested 2 months prior to their use, dried for 3 minutes with a microwave (700 W), ground and preserved at -18C. Voucher specimens are preserved in the EEA Cerro Azul.

First grade chloroform and methanol were purchased from Merck Biosolve Ltd. (Valkenswaard, The Netherlands). $CDCl_3$ (99.96%) and D_2O (99.00%) were purchased from Cambridge Isotope Laboratories Inc. (Miami, FL, USA) and NaOD was purchased from Cortec (Paris, France). Arbutin was obtained from Sigma (St. Louis, MO, USA).

8.2.2 Methods

Extraction. Three hundred mg of ground material were placed in a centrifuge tube. Five millilitres of a 50% water/methanol mixture and 5 mL of chloroform were added to the tube followed by vortexing for 30 s and sonication for 1 min. The tube was then centrifuged at 3000 rpm for 20 min. This procedure was repeated twice. The organic fractions were transferred to a 25 ml round-bottom flask and taken to dryness with a rotary vacuum evaporator. The aqueous layer was diluted 10 times with deionized water and evaporated in SpeedVac.

NMR measurements. Organic fractions were dissolved in $CDCl_3$. Aqueous fractions were dissolved with D_2O to which KH_2PO_4 was added as a buffering agent. The pH of the D_2O for NMR measurements was adjusted to 6.0 using a 1 M NaOD solution. All spectra were recorded on a Bruker AV-400 NMR spectrometer operating at a proton NMR frequency of 400.13 MHz. For each sample, 128 scans were recorded with the following parameters: 0.126 Hz/point, pulse width (PW) = 4.0 ms (30°), and relaxation delay (RD) = 2.0 s. FIDs were Fourier transformed with LB = 0.3 Hz, GB = 0, and PC = 1.0. For quantitative analysis, peak area was used. The spectra were referenced to residual solvent $CDCl_3$ at 7.26 ppm for organic fraction and trimethyl silane propionic acid sodium salt (TSP) at 0.00 ppm for the aqueous fraction. Hexamethyl disilane (HMDS, 0.01%, v/v) for $CDCl_3$ and TSP (0.01%, w/v) for aqueous fractions were used as internal standards for scaling of all NMR signals. The total peak intensities in every 0.02 ppm in ^{1}H-NMR spectra in the range of d 0.0 – 10.0 were used as variables. The regions of **δ** 4.6-5.8 in aqueous fraction and **δ** 7.28-7.24 in organic fractions were excluded from the analysis because of the residual water and $CHCl_3$ signals respectively.

Data analysis. The ^{1}H-NMR spectra were automatically reduced to ASCII files using AMIX (v. 3.7, Bruker Biospin). Spectral intensities were scaled to HMDSO for the organic fraction and to TSP for aqueous fraction, and reduced to integrated regions of equal width (0.02 ppm) corresponding to the region of **δ** 0.40 –**δ** 10.00. PCA and discriminant analysis were performed with the SIMCA-P software (Umetrics, Umeå, Sweden).

HPLC analysis: A Waters HPLC system equipped with a 626 pump, a 2996 photodiode array detector fixed at 282 nm, and a 717 plus autosampler (Waters, Milford, USA was used for arbutin determination. Twenty microliter samples were injected onto a Inertsil ODS-3 (4.6 × 250 mm, s-3 mm) (GL Sciences Inc, Tokyo, Japan) column and eluted with a linear gradient starting at a proportion of 90:10 of AcOH/ H_2O/1% AcOH/ MeOH for

4.5 minutes and then changing to 70:30 in 26 min and to 90:10 in 20 min. The flow rate was 1.0 mL/min.

8.3 Results and discussion

Visual inspection of ^{1}H NMR spectra and assignments of the compounds

Among the *Ilex* species evaluated in this study, only *I. paraguariensis* was found to contain caffeine and theobromine, whereas theophylline was not detected. This finding is in accordance with reports of caffeine and theobromine in *I. paraguariensis* leaves dating back to the 19th century whereas the presence of theophylline, reported in very small quantities (Mazzafera P, 1994; Vázquez *et al*, 1986), is a matter of controversy, as other researchers did not detect this substance (Clifford & Ramírez-Martínez, 1990; Ashihara, H, 1993; Filip *et al*., 1998).

Characteristic signals due to one purine proton at δ 7.51 (1H, *s*), three *N*-methyls at δ 3.99, **δ** 3.59, and δ 3.41 (3H each, *s*) in the ^{1}H NMR spectrum of organic fractions of *I. paraguariensis* is in accordance with caffeine. In addition to these caffeine signals, δ 7.53 (1H, *s*), δ 3.97 (3H, *s*), and δ 3.54 (3H, *s*) were assigned to theobromine **(Figure 8.1).** Theophylline was not detected in *I. paraguariensis* leaves.

The main triterpenes of *Ilex* species were found to be ursolic acid analogues (Kraemer *et al*.,1996; Taketa *et al*., 2000). In the region of δ 5.0 - 5.5, characteristic olefinic protons (H-12) of triterpenes are detected at δ 5.12 (t, J = 3.3 Hz), 5.17 (t, J = 3.3 Hz), and 5.29 (t, J = 3.3 Hz) (Figure 8.2). Methyl signals of triterpenes in the range of δ 0.7 – d 1.2 were also found to be discriminating ^{1}H NMR signals of *Ilex* species.

^{1}H NMR spectra of the aqueous fractions for the Ilex species are shown in Figure 8-3. The patterns in the aromatic region (δ 6.0 – δ 8.0) are quite different from each other. Previous reports have shown that some phenylpropanoids such as caffeic acid, chlorogenic acid, 3,4-dicaffeoylquinic acid, 3,5-dicaffeoylquinic acid, and 4,5-dicaffeoylquinic acid are the main characteristic metabolites of *Ilex* species (Filip *et al., 2001*). In accordance with this paper, the signals of the main differentiating aromatic compounds in the extracts were assigned to phenylpropanoids **(Figure 8.3).**

^{1}H NMR spectrum of *I. paraguariensis* **(Figure 8-4)** is in accordance with the spectrum of phenylpropanoids, showing the typical signals due to two *trans* olefinic protons (*J* = 15-16 Hz) in the region of δ 6.1 - **δ** 6.5 (H-8') and **δ** 7.6 – 7.7 (H-7'). The complex pattern of the ^{1}H NMR spectrum in **δ** 6.9 - **δ** 7.3 shows that this plant contains several phenylpropanoids such as caffeic acid, *p*-coumaric acid, and their glycosides.

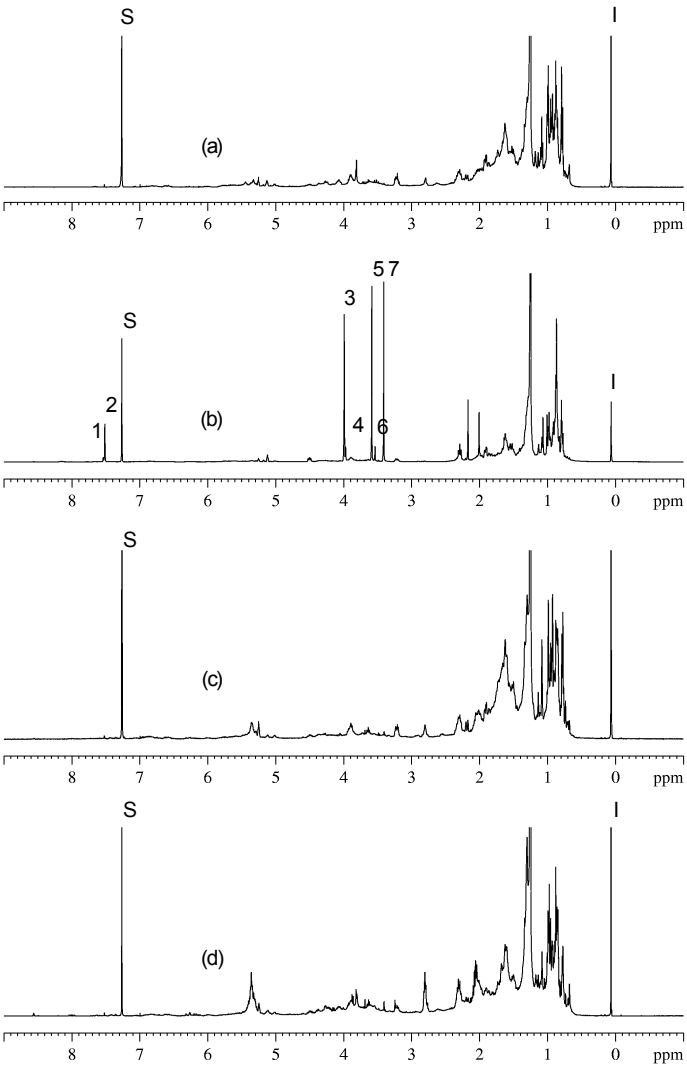


Fig. 8-1 ^{1}H NMR spectra for organic fractions of I. argentina leaves (a), I. paraguariensis leaves (b), I. pseudobuxus leaves (c), and I. tauberiana leaves (d). Peaks: 1, H-8 of theobromine; 2, H-8 of caffeine; 3, 7-methyl of caffeine; 4, 7-methyl of theobromine; 5, 3-methyl of caffeine; 6, 3-methyl of theobromine; 7, 1-methyl of caffeine; S, residual solvent signal of CDCl3; I, internal standard of HMDSO.

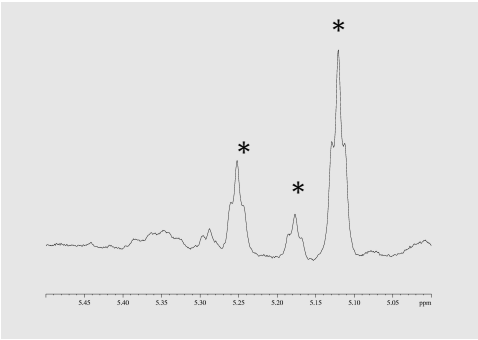


*Fig. 8- 2 ^{1}HNMR spectra for organic fraction of I. paraguariensis leaves in the range of 5.0–5.5. *, possible signals of H-12 protons of ursolic acid analogues.*

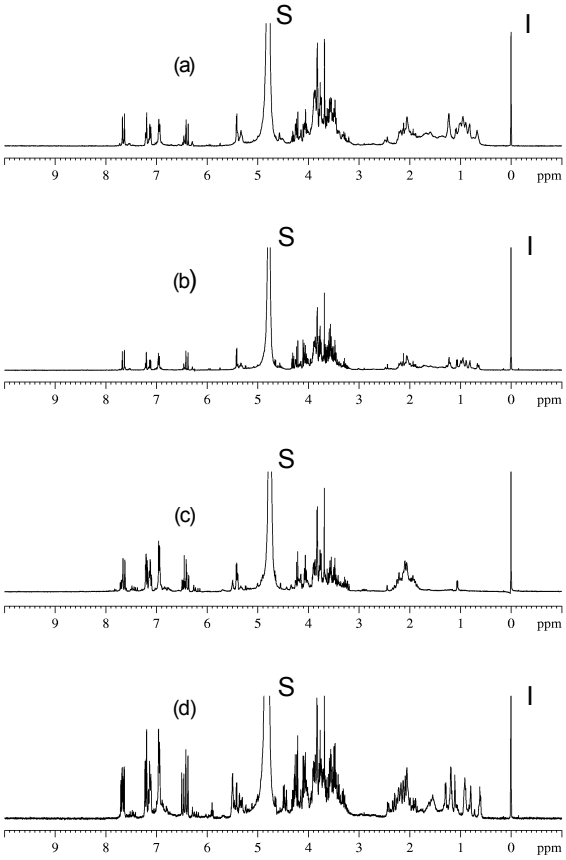


Fig. 8-3 1H NMR spectra for aqueous fractions of I. dumosa var. dumosa leaves (a), I. dumosa var. guaranina leaves (b), I. paraguariensis leaves (c), and I. pseudobuxus leaves (d). Peaks: S, residual solvent signal of HDO; I, internal standard of TSP.

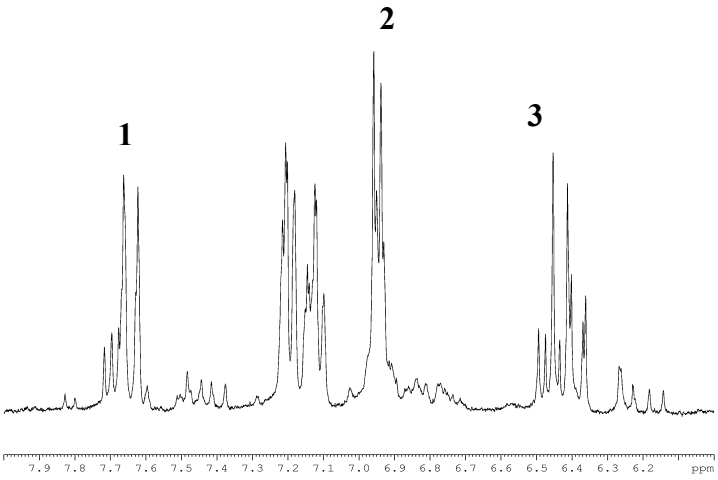


Fig 8-4. 1HNMR spectra for an aqueous fraction of I.paraguariensis leaves in the range of δ 6.0–8.0. Peaks: 1, H-7′ region of phenylpropanoids; 2, aromatic region of phenylpropanoids; 3, H-8′ region of phenylpropanoid.

Intriguingly, intense signals at δ 7.06 (1H, *d*, *J* = 9.0 Hz), δ 6.88 (1H, *d*, *J* = 9.0 Hz), and δ 4.89 (1H, *d*, *J* = 7.6 Hz) appeared in *I. argentina, I. brasiliensis, I. integerrima, I. microdonta, I. taubertiana,* and *I. theezans* **(Figure 8-5).** These signals were assigned to H-2, H-3, and the anomeric proton of glucose in arbutin, respectively.

This was confirmed by 2D-NMR spectra such as ^{1}H-^{1}H-COSY, HMQC, and HMBC and comparison with the reference compound arbutin. There have been no previous reports of arbutin in *Ilex* species. In this study, however, it was found that some *Ilex* species contain arbutin as a major metabolite.

For further confirmation and quantitative analysis of arbutin in *Ilex* species evaluated in this study, HPLC analysis was performed. Among the species studied, *I. pseudobuxus* showed the highest content of arbutin (106.0 mg/g) **(Figure 8-6).**

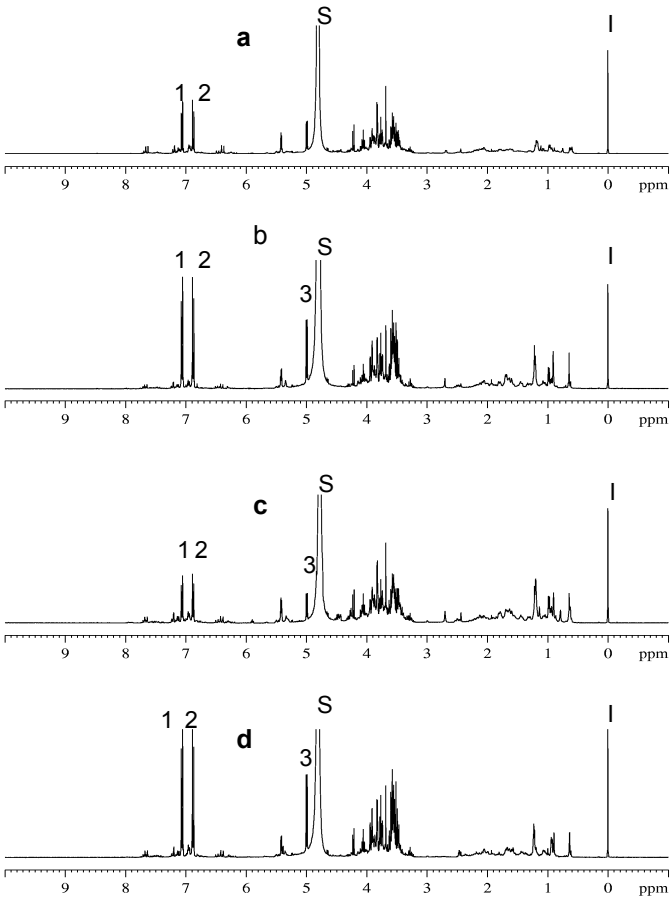


Fig.8-5 ^{1}H NMR spectra for aqueous fractions of I. argentina leaves (a), I. brasiliensis leaves (b), I. brevicuspis leaves (c), and I. theezans leaves (d). Peaks: S, residual solvent signal of HDO; I, internal standard of TSP; 1, H-2′ and H-6′ of arbutin; 2, H-3′ and H-5′ of arbutin; 3, H-1 of arbutin

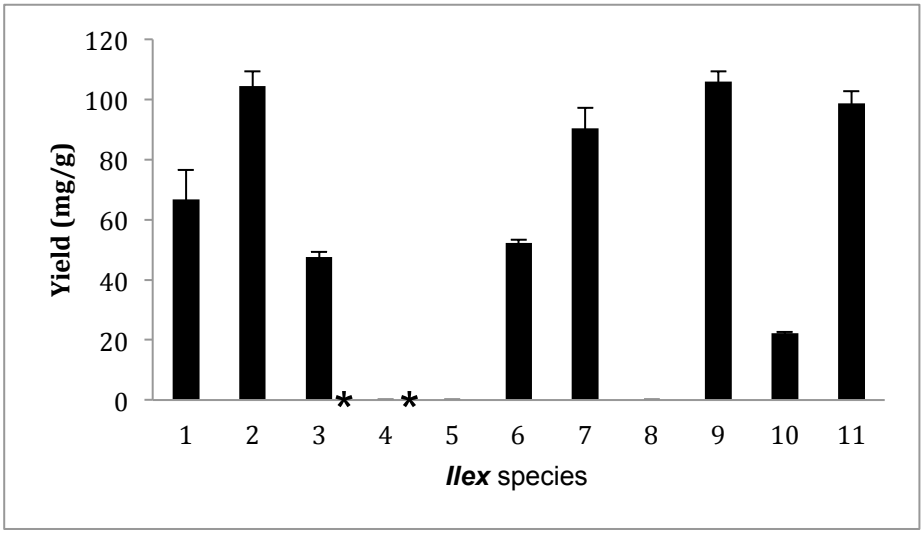
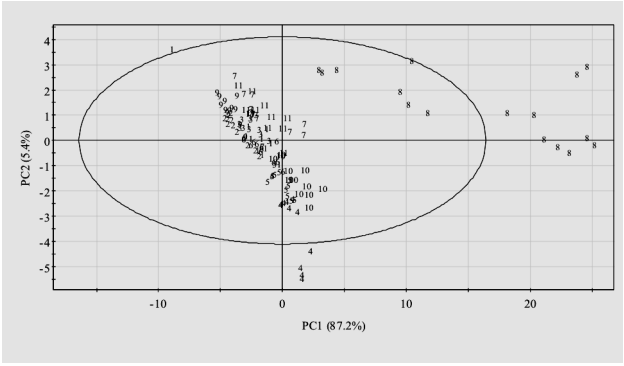
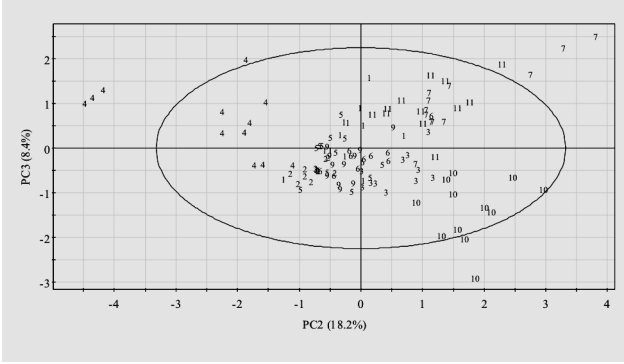


*Fig. 8-6. Yield (mg/g) of arbutin in Ilex species obtained from HPLC analysis: 1, I. argentina; 2, I. brasiliensis; 3, I. brevicuspis; 4, I. dumosa var. dumosa; 5, I. dumosa var. guaranina; 6, I. integerrima; 7, I. microdonta; 8, I. paraguariensis var. paraguariensis; 9, I. pseudobuxus; 10, I. taubertiana; 11, I. theezans; *, not detected. Results are based on triplicate analysis.*

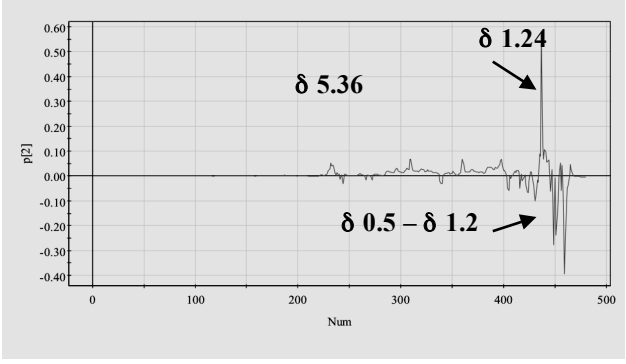
PCA. Principal component analysis (PCA) is an unsupervised clustering method requiring no knowledge of the data set and acts to reduce the dimensionality of multivariate data while preserving most of the variance within it (Eriksson *et al*, 2001). The data for PCA can be scaled in different ways. If the data is mean-centered with no scaling then a covariance matrix is produced, but if the data is mean-centered and the columns of the data matrix scaled to unit variance, a correlation matrix is produced (Ward *et al*, 2003). Both of the methods were applied to the ^{1}H-NMR data set of *Ilex* species, the covariance method showing a better separation of the species. Application of PCA to the organic fractions resulted in a good separation of *I. paraguariensis* from other species by PC1 **(Figure 8–7a).** This separation is due to the signals of caffeine, theobromine, and triterpenoids. For the detailed inspection of other *Ilex* species, PCA was done for all the species excluding *I. paraguariensis*.



(a)



(b)



(c)

Fig. 8-7. Score plot of discriminating PC scores of organic fraction for Ilex species with I. paraguariensis (a) and without I. paraguariensis (b) following PCA analysis. 1: I. argentina, 2: I. brasiliensis, 3: I. brevicuspis, 4: I. dumosa var. dumosa, 5: I. dumosa var. guaranina, 6: I. integerrima, 7: I. microdonta, 8: I. paraguariensis var. paraguariensis, 9: I. pseudobuxus, 10: I. taubertiana, 11: I. theezans.

As a next step, the focus was placed on the results obtained from the aqueous fraction. For a clear separation, however, PC3 is needed apart from PC1 and PC2.

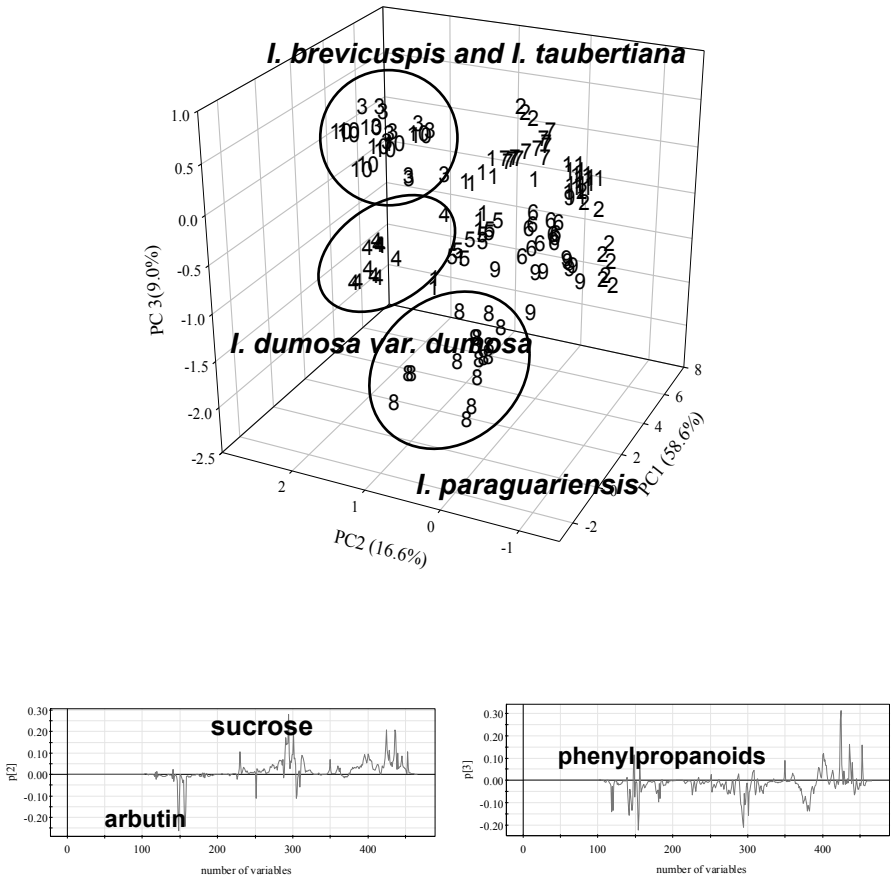


Fig. 8-8 Score plot of PC1, PC2, and PC3 scores of aqueous fraction for Ilex species following PCA (a) and loading plots for PC2 (b) and PC3 (c): 1, I. argentina; 2, I. brasiliensis; 3, I. brevicuspis; 4, I. dumosa var. dumosa; 5, I. dumosa var. guaranina; 6, I. integerrima; 7, I. microdonta; 8, I. paraguariensis var. paraguariensis; 9, I. pseudobuxus; 10, I. taubertiana; 11, I. theezans.

As seen in Figure 8-8a, there is a clear discrimination between *I. paraguariensis* and other *Ilex* species. Notably, aqueous fractions of *I. dumosa* var. *dumosa*, *I. breviscuspus*,

and *I. taubertiana* show unique metabolomic fingerprints. This separation took place in the first three principal components which cumulatively accounted for 84.2% of the variation.

The separation between *I. paraguariensis* and other species was easily achieved with the PC2 and PC3 values. *Ilex paraguariensis* has a lower PC2 and PC3. PC2 value is affected by the amount of arbutin and sucrose and a higher PC2 value means lower phenylpropanoids and sucrose content **(Figure 8-8b).** PC3 was largely influenced by the quantity of phenylpropanoids and its lower value indicated a higher amount of phenylpropanoids **(Figure 8-8c).**

8.4 Classification of *Ilex* species based on PCA

Pattern recognition is often described as a procedure for formulating rules of classification (Albano *et al.*, 1978). The most encountered goal of a pattern recognition application is classification (Kowalski *et al.*, 1972). Using a collection of knowns and a classification rule, a set of unknowns is classified. On the basis of given classes, each of which contains a number of observations mapped by a multitude of variables, guidelines and rules are developed to make it possible to classify new observations as similar or dissimilar to the members of the existing classes (Eriksson *et al.*, 2001). Data that was observed in a class of one *Ilex* species was classified by soft independent modeling of class analogy (SIMCA) (Wold *et al.*, 1984). Using this method, each class of *Ilex* species was modeled separately by disjoint PC-models. Based on the residual variation of each class, the distance to the model (DModX) of each observation was computed. Prior to the classification, PCA was carried out for each species to obtain optimum PCs, which were necessary to build the model of each species. After each particular model based on optimum PCs (e.g., the first seven PCs were used for the metabolites obtained from organic fractions of *I. paraguariensis*) had been constructed, all species were compared using the distance to each species model. In the classification based on the metabolites obtained from organic fractions, most of the species do not overlap and showed unique metabolomic profiles as an example of *I. paraguariensis* (**Figure 8- 9a**) with the exception of *I. pseudobuxus*, *I. brasiliensis*, and *I. theezans* (**Figure 8- 9b**). In the classification of metabolites obtained from aqueous fractions, no species overlapped. As shown in the case of *I. paraguariensis* in **Figure 8-10a**, even the species *I. pseudobuxus*, *I. brasiliensis*, and *I. theezans*, the metabolite profiles of which could not completely be resolved for the organic fractions, were clearly distinguished from each other with the metabolomic profile of the aqueous fractions **(Figure 10b).**

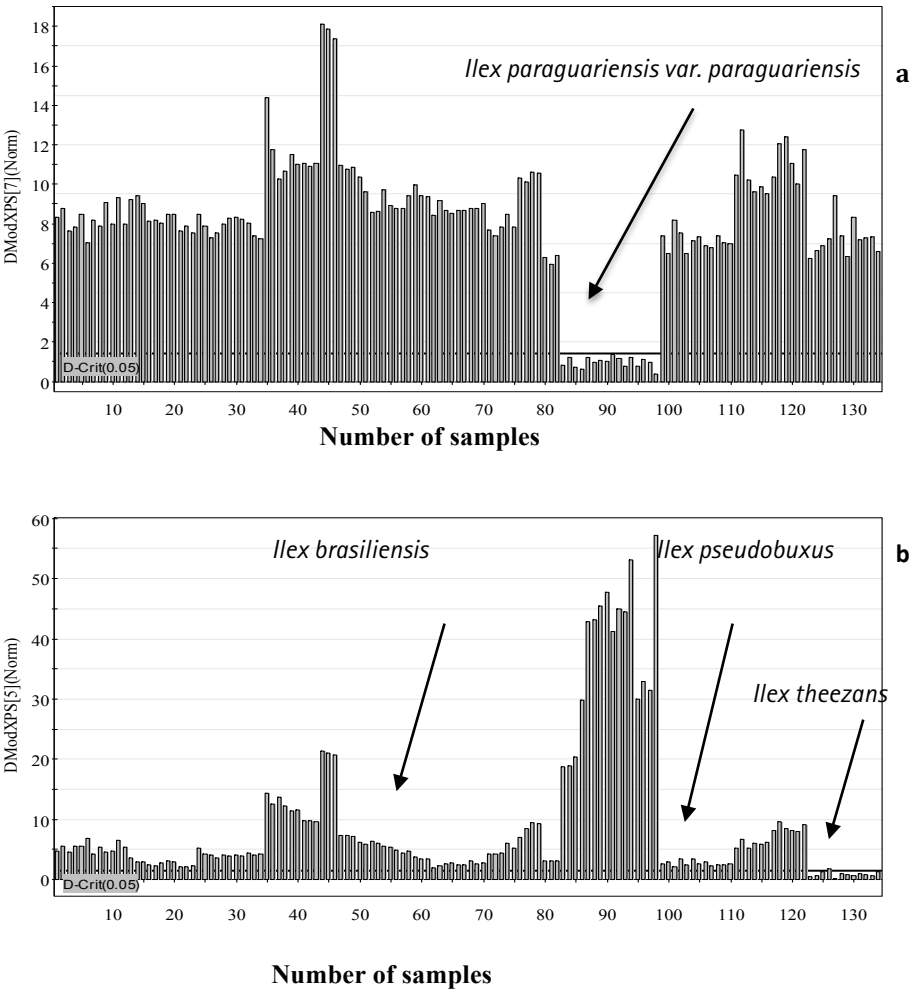


Fig. 8-9 DModX plot for Ilex species prediction set of organic fractions of I. paraguariensis var. paraguariensis using the first seven PCs (a) and of I. theezans using the first five PCs (b): 1–12, I. argentina; 13–23, I. brasiliensis; 24–34, I. brevicuspis; 35–46, I. dumosa var. dumosa; 47–58, I. dumosa var. guaranina; 59–70, I. integerrima; 71–82, I. microdonta; 83–98, I. paraguariensis var. paraguariensis; 99–110, I. pseudobuxus; 111–122, I. taubertiana; 123–134, I. theezans.

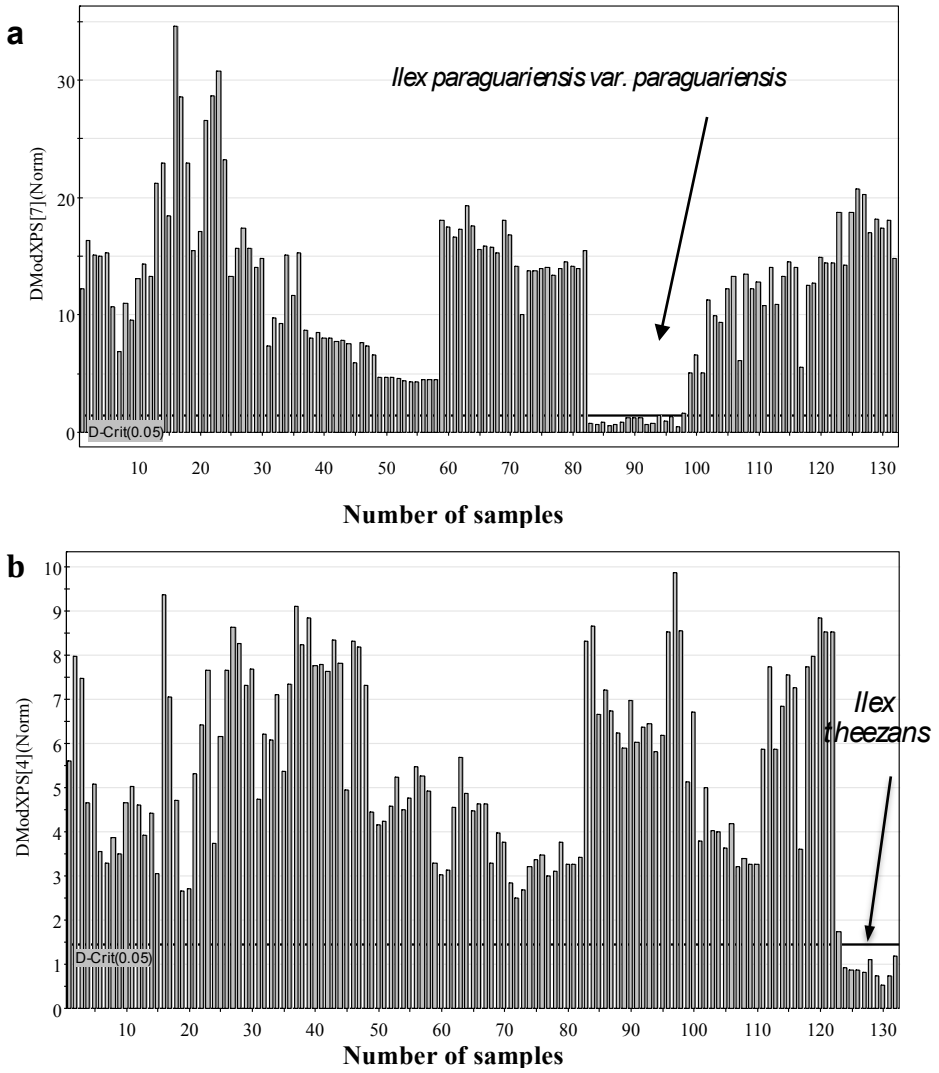


Figure 8-10. DModX plot for Ilex species prediction set of aqueous fractions of I. paraguariensis var. paraguariensiss using the first seven PCs (a) and of I. theezans using the first four PCs (b): 1–12, I. argentina; 13–24, I. brasiliensis; 25–36, I. brevicuspis; 37–48, I. dumosa var. dumosa; 49–58, I. dumosa var. guaranina; 59–70, I. integerrima; 71–82, I. microdonta; 83–98, I. paraguariensis var. paraguariensis; 99–110, I. pseudobuxus; 111–122, I. taubertiana; 123–132, I. theezans

This study proves that it is possible to discriminate the 11 *Ilex* species, *I. argentina, I. brasiliensis, I. brevicuspis, I. dumosa* var. *dumosa, I. dumosa* var. *guaranina, I. integerrima, I. microdonta, I. paraguariensis* var. *paraguariensis, I. pseudobuxus, I. taubertiana*, and *I. theezans* by multivariate analysis of their metabolite fingerprints obtained by ^{1}H-NMR spectra of crude extracts of the plant materials. The major compounds contributing to the discrimination were found to be several phenylpropanoids and arbutin. In particular, arbutin, which has not been reported yet as a component of *Ilex* species, was found to be a discriminating metabolite (biomarker) for *I. argentina, I. brasiliensis, I. brevicuspis, I. integerrima, I. microdonta, I. pseudobuxus, I. taubertiana*, and *I. theezans*. However, arbutin was not present in a detectable amount in the samples of *I. dumosa* var. *dumosa, I. dumosa* var. *guaranina,* and *I. paraguariensis* var. *paraguariensis* which were analized. Besides these major phenolic metabolites, a number of putative minor metabolites also play a role to differentiate *Ilex* species.

Alongside simple principal component analysis (PCA) used for reduction of the NMR data set obtained from the metabolites, classification based on each species model was done and obviously complements the differentiation of *Ilex* species based on the metabolite profiles.

The method using ^{1}H-NMR and multivariate analysis may afford consistent discrimination of *Ilex* species based on metabolomic profiling as a tool for chemotaxonomical studies.

This method could also be useful for quality control issues which are critical to the Yerba mate producing industry, such as detection of adulterants, leaf/stalk content, detection of inadequate drying or conservation of plants, and standardisation of blends among others for which there is not an adequate analytical solution at this time.

.....

REFERENCES

- **Albano C, Dunn III WJ, Edlund U, Johansson E, Nordén B, Sjöström M, Wold S** (1978). Four levels of pattern recognition, Anal Chim Acta, 103, 429-443.
- **Ashihara H** (1993). Purine metabolism and the biosynthesis of caffeine in maté leaves. Phytochemistry 33, 1427-1430.
- **Bertoni MH, Prat Kircun SD, Kanzig RG, Cattaneo YP** (1992). Hojas frescas de *Ilex paraguariensi*s S. Hill. III. An Asoc Quím Argent; 80: 493-501.

- **Blumenthal M, Busse W, Goldberg A, Gruenwald J, Hall T, Klein S, Riggins C, Rister R, Eds.** (1998) The complete German Comission E monographs. Austin: American Botanical Council.
- **Brescia MA, Caldarola V, De Giglio A, Benedetti D, Fanizzi FP, Sacco A** (2002). Characterization of the geographical origin of Italian red wine based on traditional and nuclear magnetic resonance spectrometric determinations. Anal Chim Acta, 458, 117-186.
- **Cañigueral S, Vila R, Wichtl M** Eds (1998). Plantas medicinales y drogas vegetales para infusión y tisana. Manual para farmacéuticos y médicos. Milan: OEMF International SRL.
- **Charlton AJ, Farrington WHH, Brereton P** (2002). Application of ^{1}H NMR and multivariate statistics for screening complex mixtures: quality control and authenticity of instant coffee. J Agric Food Chem, 50, 3098-3103.
- **Clifford MN, Ramirez-Martinez JR** (1990) Chlorogenic acids and purine alkaloids contents of Maté (*Ilex paraguariensis*) leaf and beverage. Food Chem 35, 13-21.
- **Duarte I, Barros A. Belton PS, Righelato R, Spraul M, Humpfer E, Gil AM** (2002) High-resolution nuclear magnetic resonance spectroscopy and multivariate analysis for the characterization of beer. J Agric Food Chem, 50, 2475-2481.
- **Eriksson L, Johansson E, Kettaneh-Wold N, Wold S** (2001). Multi- and Megavariate Data Analysis. Umetrics Academy, Umeå, Sweden; pp. 43-70.
- **Filip R, Lopez PG, Coussio JD, Ferraro GE** (1998). Mate substitutes or adulterants: study of xanthine content. Phytother Res. 12, 129-131.
- **Filip R, Lopez PG, Giberti G, Coussio JD, Ferraro GE** (2001). Phenolic compounds in seven South American *Ilex* species. Fitoterapia, 72, 774-778.
- **Forveille L, Vercauteren J, Rutledge DN** (1996) D. N. Multivariate statistical analysis of two-dimensional NMR data to differentiate grapevine cultivars and clones. Food Chem, 57, 441-450.
- **Giberti G** (1989). Los parientes silvestres de la yerba mate y el problema de su adulteración. Dominguezia 1989; 7:1-22.
- **Gorzalczany S, Filip R, Alonso MR, Miño J, Ferraro GE, Acevedo C** (2001). Choleretic effect and intestinal propulsion of "mate" (*Ilex paraguariensis)* and its substitutes or adulterants. J.Ethnopharmacol, 75, 291–294.
- **Graham, H** (1984) Mate. In The methylxanthine beverages and foods: Chemistry, consumption and health effects, Spiller, G. A. Ed.; Alan R. Liss, New York, pp. 179-183.
- **Kraemer KH, Taketa ATC, Schenkel EP, Gosmann G, Guillaume D** (1996). Matesaponin 5, a highly polar saponin from *Ilex paraguariensis*. Phytochemistry, 42, 1119-1122.
- **Kowalski BR, Bender CF** (1972). Pattern recognition. A powerful approach to interpreting chemical data. J Am Chemical Soc, 94, 5632-5639.
- **Mazzafera P** (1994) Caffeine, theobromine and theophylline distribution in *Ilex paraguariensis*. Revista Brasileira de Fisiologia Vegetal, 6, 49-151.

- **Newall C, Anderson LA, Phillipson D** (1996). Herbal Medicines. A Guide for Health-Care Professionals. London: The Pharmaceutical Press, 1996.
- **Nord LH, Kenne L, Jacobsson SP** (2001). Multivariate analysis of ^{1}H NMR spectra for saponins from *Quilaja saponaria* Molina. Anal Chim Acta 446, 199-209.
- **Ott KH, Araníbar N, Singh B, Stockton GW** (2003). Metabonomics classifies pathways affected by bioactive compounds. Artificial neural network classification of NMR spectra. Phytochemistry, 62, 971-985.
- **Reginatto FH, Athayde ML, Gosmann G, Schenkel EP** (1999) Methylxanthines accumulation in *Ilex* species-caffeine and theobromine in erva- mate *(Ilex paraguariensis*) and other *Ilex* species. J Brazilian Chemical Soc, 10, 443-446
- **Schmalko ME, Alzadora SM** (2001) Color, chlorophyll, caffeine and water content variation during yerba mate processing. Drying-Technol ,19, 599-610.
- **Taketa ATC, Schmittmann-Schlager T, Guilaume D, Gosmann G, Schenkel EP** (2000) Triterpenoid glycosides and a triterpene from *Ilex brevicuspis*. Phytochemistry, 53, 901-904.
- **Vázquez A, Moyna P** (1986). Studies on mate drinking. J Ethnopharmacol ; 18: 267-272.
- **Vogels, JTWE, Terwel L, Tas AC, Van den Berg F, Dukel F, Van der Greef J** (1996). 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, 44, 175-180.
- **Ward J, Harris C, Lewis J, Beale M H** (2003). Assessment of ^{1}H NMR spectroscopy and multivariate analysis as a technique for metabolite fingerprinting of *Arabidopsis thaliana*. Phytochemistry, 62, 949-957.
- **Wold S, Albano C, Dunn III W J, Edlund U, Esbensen K, Geladi P, Helberg S, Johansson E, Lindberg W, Sjöström M** (1984). Multivariate data analysis in chemistry, In Chemometrics: Mathematics and statistics in chemistry, Kowalski B. R., Ed.; D. Reidel Publishing Company, Dordrecht, The Netherlands.