Anal Bioanal Chem (2012) 402:833–842
DOI 10.1007/s00216-011-5479-0

ORIGINAL PAPER

Forensic analysis of *Salvia divinorum* using multivariate statistical procedures. Part I: discrimination from related *Salvia* species

Melissa A. Bodnar Willard · Victoria L. McGuffin · Ruth Waddell Smith

Received: 28 July 2011 / Revised: 29 September 2011 / Accepted: 4 October 2011 / Published online: 29 October 2011

Abstract *Salvia divinorum* is a hallucinogenic herb that is internationally regulated. In this study, salvinorin A, the active compound in *S. divinorum*, was extracted from *S. divinorum* plant leaves using a 5-min extraction with dichloromethane. Four additional *Salvia* species (*Salvia officinalis, Salvia guaranitica*, *Salvia splendens*, and *Salvia nemorosa*) were extracted using this procedure, and all extracts were analyzed by gas chromatography–mass spectrometry. Differentiation of *S. divinorum* from other *Salvia* species was successful based on visual assessment of the resulting chromatograms. To provide a more objective comparison, the total ion chromatograms (TICs) were subjected to principal components analysis (PCA). Prior to PCA, the TICs were subjected to a series of data pretreatment procedures to minimize non-chemical sources of variance in the data set. Successful discrimination of *S. divinorum* from the other four *Salvia* species was possible based on visual assessment of the PCA scores plot. To provide a numerical assessment of the discrimination, a series of statistical procedures such as Euclidean distance measurement, hierarchical cluster analysis, Student's *t* tests, Wilcoxon rank-sum tests, and Pearson product moment correlation were also applied to the PCA scores. The statistical procedures were then compared to determine the advantages and disadvantages for forensic applications.

Keywords Controlled substance identification · *Salvia divinorum* · Salvinorin A · *Salvia* species · Multivariate statistical procedures · Principal components analysis

Electronic supplementary material The online version of this article (doi:10.1007/s00216-011-5479-0) contains supplementary material, which is available to authorized users.

M. A. Bodnar Willard · V. L. McGuffin (✉)
Department of Chemistry, Michigan State University,
East Lansing, MI 48824, USA
e-mail: mcguffin@msu.edu

M. A. Bodnar Willard · R. Waddell Smith (✉)
Forensic Science Program, School of Criminal Justice,
Michigan State University,
East Lansing, MI 48824, USA
e-mail: rwsmith@msu.edu

Introduction

Salvia divinorum is a hallucinogenic perennial herb from the mint family (Lamiaceae or Labiatae), one of nearly a thousand species of *Salvia*. *S. divinorum* has recently become of interest in forensic science as the active ingredient, salvinorin A, is considered to be the most potent known hallucinogen of natural origin, rivaling even the semi-synthetic hallucinogen lysergic acid diethylamide [1, 2]. Due to its hallucinogenic effects, *S. divinorum* is regulated in many countries, including Australia, Belgium, Denmark, Estonia, Finland, Italy, Japan, Spain, and Sweden. While not federally regulated in the USA, as of 2011, 25 individual states have regulated either *S. divinorum* or salvinorin A, and several others have legislation pending [3, 4]. According to the National Forensic Laboratory Information System, submission of *S. divinorum* and salvinorin A to state and local forensic laboratories in the USA increased from one submission in 2004 to 68 in 2008, although this number decreased slightly to 55 in 2009 [5].

In terms of forensic identification, it is not possible to distinguish *S. divinorum* from other *Salvia* species based solely on visual inspection of the plant [6]. In forensic science laboratories, *S. divinorum* is currently identified by detection of salvinorin A in extracts of the plant material. As salvinorin A is only known to exist in *S. divinorum*, a

visual comparison of chromatograms of questioned *Salvia* samples may provide a means of differentiation, thereby allowing identification of the plant as *S. divinorum* [6].

A recent report by the US National Academy of Sciences' National Research Council, "Strengthening Forensic Science in the United States: A Path Forward", highlighted the need for statistical evaluation of evidence [7]. Hence, a more objective method for the differentiation of *S. divinorum* from other *Salvia* species is desirable. Multivariate statistical procedures, particularly those that are unsupervised, can be used for this purpose as there is no human intervention in the decision-making process. While such procedures have been applied for the association and discrimination of different types of forensic evidence, there have been no reports using such procedures for *S. divinorum* or related plant materials [8, 9].

Principal components analysis (PCA) is a widely used, unsupervised multivariate statistical procedure. PCA reduces a complex data set to a few principal components representing the greatest contributions to variance among the samples [10]. Samples that are chemically similar have similar scores and cluster together on the PCA scores plot, but chemically different samples are separated. Loadings plots are used to identify the chemical compounds that contribute most to the variance described by the principal components. PCA has been previously applied for association and discrimination of controlled substances. For example, it was used for chemical profiling of six target alkaloids in heroin [8] and for examining the resulting batch-to-batch variation in heroin cutting and distribution [9]. Despite this, PCA has not yet been investigated for discrimination of *S. divinorum* from other *Salvia* species.

As chromatograms of plant materials are very complex, containing hundreds of variables, data pretreatment steps may be necessary prior to PCA in order to minimize non-chemical sources of variance. Common pretreatment procedures for chromatographic data include background correction, smoothing, retention time alignment, and normalization [10, 11]. Background correction is often performed on chromatographic data to minimize low-frequency noise, primarily due to drift in the detector signal. Smoothing is performed to minimize high-frequency noise in chromatograms, primarily due to random fluctuations in the detector signal. Retention time alignment is performed to account for drift in retention time between analyses, and normalization is performed to account for variations in injection volume and instrument sensitivity between analyses.

A potential problem with PCA is that interpretation of the resulting scores plot is largely based on visual inspection, which can be subjective. This can be overcome by using statistical methods to provide a numerical evaluation of sample positioning on the scores plot. Examples include Euclidian distance measurements, hierarchical cluster analysis (HCA), Student's *t* tests, Wilcoxon rank-sum tests, and Pearson product moment correlation (PPMC) coefficients. The Euclidian distance represents the numerical distance between two points in multidimensional space [12]. In this research, Euclidian distances were calculated between selected pairs of the means of the PCA scores. Samples that are positioned closely in the PCA scores plot will have short Euclidian distances, whereas samples that are distinct will have longer Euclidian distances. HCA can be used to assess the similarity of a multivariate data set and, in this study, was calculated from the Euclidian distances in the PCA scores plot [12]. Additionally, HCA allows for a graphical representation of the association among samples in the data set, based on the similarity of the data. The Student's *t* test and the Wilcoxon rank-sum test can be used to test the hypothesis of equality of two samples at a given confidence level. The *t* test assumes that the data are normally distributed (Gaussian), whereas the Wilcoxon test makes no assumptions regarding the distribution of the data [12].

PPMC coefficients provide a means to compare the similarity of the complete chromatogram. Coefficients are calculated pairwise, for example, a questioned sample and a reference standard, allowing comparison of the two chromatograms based on a single number. Coefficients range from +1 to −1, with positive coefficients indicating a positive correlation and negative coefficients indicating a negative correlation. Coefficients of ±0.80 or greater indicate strong correlation, coefficients ranging from ±0.50 to ±0.79 indicate moderate correlation, coefficients of ±0.49 or less indicate weak correlation, and coefficients close to zero indicate no correlation [12]. As such, PPMC coefficients can be applied to assess similarity between samples, especially those positioned closely in the PCA scores plot.

In this proof-of-concept study, a procedure for the extraction of salvinorin A from *S. divinorum* was firstly investigated, using different extraction solvents and extraction times. The extraction efficiency of the developed procedure was determined and the procedure was then used to extract four additional *Salvia* species (*Salvia officinalis, Salvia guaranitica, Salvia splendens*, and *Salvia nemorosa*). Extracts were analyzed by gas chromatography–mass spectrometry (GC-MS), and the resulting chromatograms were subjected to various data pretreatment procedures prior to PCA, to investigate discrimination of the five *Salvia* species. Euclidean distances, HCA, Student's *t* tests, Wilcoxon rank-sum tests, and PPMC coefficients were evaluated to provide an objective interpretation of the resulting PCA scores plot. The statistical procedures were compared to determine the advantages and disadvantages of each for the discrimination of *S. divinorum* from other *Salvia* species, particularly in forensic casework.

Materials and methods

Investigation of extraction procedure

Approximately 0.2 g of dried *S. divinorum* leaves (Ethnosupply, Vancouver, BC, Canada) were placed in separate acid-washed beakers with 15.0 mL of the appropriate solvent: methanol, acetonitrile, acetone, chloroform, dichloromethane, and hexane (99.9% purity; Honeywell Burdick and Jackson, Morristown, NJ). *S. divinorum* leaf material was extracted in each solvent for 1 min. All extractions were performed in triplicate. Extracts were filtered using a 0.45-μm nylon mesh (Small Parts Inc., Miami Lakes, FL) and then rinsed with 5.0 mL of the appropriate extraction solvent. The filtered solution was evaporated to dryness with nitrogen under gentle heating at 35 °C. Extracts were weighed and then stored at 4 °C until analysis. Extracts were reconstituted in a known volume of the appropriate solvent containing 0.0119 M progesterone (Sigma, St. Louis, MO) as an internal standard prior to GC-MS analysis. The most appropriate solvent was selected, based on the mass of salvinorin A extracted, and was used for all subsequent extractions. Using this solvent, extraction times of 1, 3, 6, 30, 100, and 300 min were then investigated in a similar manner.

Extraction efficiency of salvinorin A

Reference standards containing 0.1, 0.5, 2.0, 5.0, 7.5, 10.0, and 12.0 mg/mL salvinorin A (Chromodex, Irvine, CA) were prepared in dichloromethane containing 0.0119 M progesterone as an internal standard. After GC-MS analysis, a calibration curve over the range of 0.1 to 12.0 mg/mL salvinorin A was generated in Microsoft Excel (version 2007, Microsoft Corp., Redmond, WA) to determine the linear range of the detector.

The mass of salvinorin A present in *S. divinorum* had previously been found to have high variability [12]. Therefore, the extraction efficiency was determined using a similar *Salvia* species, *S. officinalis*, that did not contain salvinorin A. A 2.4-mg reference standard of salvinorin A was dissolved in 10.0 mL dichloromethane, then spiked onto 0.6 g of *S. officinalis* leaves and divided into three aliquots. Each 0.2-g aliquot was extracted using the developed procedure, evaporated to dryness, and reconstituted in 0.4 mL dichloromethane, resulting in an approximate mass of 2 mg/mL salvinorin A. An internal standard of 0.0119 M progesterone in dichloromethane was added prior to GC-MS analysis.

Extraction of other *Salvia* species

Dried leaves of *S. officinalis* were purchased from a commercial supplier (Penzeys Spices, Brookfield, WI). Samples of *S. guaranitica*, *S. splendens*, and *S. nemorosa* were acquired as fresh leaves (Department of Horticulture, Michigan State University, East Lansing, MI) and then dried using a food dehydrator (NESCO American Harvest, model ED-75PR, Two Rivers, WI) at ~35 °C for 24 h. The five *Salvia* species were then extracted in triplicate using the developed procedure, spiked with the internal standard, and analyzed by GC-MS.

GC-MS analysis

All extracts were analyzed by gas chromatography (Agilent Technologies, model 6890N, Santa Clara, CA) using a 5% phenyl/95% methylpolysiloxane stationary phase column (DB-5MS, 30-m length × 0.25-mm inner diameter × 0.25-μm film thickness, Agilent Technologies, Palo Alto, CA). Ultra-high purity helium was used as the carrier gas at a nominal flow rate of 1 mL/min. The inlet was maintained at 340 °C and 1 μL of sample extract was injected in split mode (50:1). The oven temperature program was as follows: 80 °C for 2 min, 10 °C/min to 340 °C, with a final hold at 340 °C for 4 min. The transfer line to the mass selective detector (Agilent Technologies, model 5973, Santa Clara, CA) was maintained at 340 °C. The detector was operated in electron ionization mode (70 eV) with a quadrupole mass analyzer in the full-scan mode (*m*/*z* 50–550) at a scan rate of 2.91 scans per second. All extracts were analyzed in triplicate.

Data pretreatment and analysis

Chromatograms of all extracts were subjected to pretreatment prior to data analysis. For all pretreatment procedures, the user-defined conditions were varied and selected to be most appropriate for this data set. Firstly, background correction was necessary due to the increased background in the chromatograms arising from the high final oven temperature. For this pretreatment, extracted ion chromatograms (EICs) representative of prominent background ions (*m*/*z* 73, 193, 207, 221, 341, and 355) were generated in the ChemStation software (version 01.02.16, Agilent Technologies, Santa Clara, CA) for each sample replicate. The EICs were compiled to form an extracted ion profile (EIP) of the background in Microsoft Excel. The EIP for each chromatogram was fitted to an asymmetric sigmoid equation using TableCurve 2D (version 1.0, Jandel Scientific, San Rafael, CA). The fitted EIP equation was regenerated in Microsoft Excel and subtracted from the corresponding total ion chromatogram (TIC), resulting in a background-corrected TIC. Each corrected TIC was then smoothed using a Fourier-transform smoothing algorithm with a window size of 2 data points (OriginPro, version 7.5853, Origin Lab Corp., Northampton, MA).

The pretreated TICs were retention time aligned using a peak-matching algorithm [13] with a window size of 4 data points. The target chromatogram used for alignment was created by summing one replicate chromatogram of each *Salvia* species and dividing by the total number of species (five). This ensured that the target contained all chemical compounds in all *Salvia* species.

Normalization of the chromatograms was problematic due to differences in the number and abundance of compounds extracted from each *Salvia* species. Several normalization procedures were investigated and maximum peak normalization was determined to be the most appropriate [14]. Maximum peak normalization was performed (Microsoft Excel) by dividing each data point in the chromatogram by the abundance of the maximum peak in that chromatogram, then multiplying by the average maximum abundance of all chromatograms in the data set. In this way, each chromatogram was scaled such that the compound with the highest abundance was equal to that of all other chromatograms.

The pretreated data set was then subjected to PCA using Pirouette (version 2.02, Infometrix, Bothell, WA). To generate plots of publication quality, the eigenvectors and scores from Pirouette were plotted using Origin (OriginPro, version 7.5853, Origin Lab Corp., Northampton, MA). Scores for the *Salvia* species were plotted to generate the scores plot, whereas the eigenvectors were plotted against retention time to generate the loadings plots. Chemical compounds in the loadings plots were identified by comparison of mass spectra with the National Institute of Standards and Technology database (version 2.0, NIST, Gaithersburg, MD).

Several procedures were evaluated to statistically assess the PCA scores plot. Mean scores on principal components 1 and 2 (PC1 and PC2, respectively) were calculated for all *Salvia* species. Euclidian distances were calculated (Microsoft Excel) between the mean score of each species and *S. divinorum* in PC1 and PC2. HCA was performed in Pirouette using Euclidian distance with single-cluster linkage to assess similarity of the PCA scores. Finally, the unequal-variance Student's *t* test and Wilcoxon rank-sum test were performed (Microsoft Excel) on the mean scores for each *Salvia* species in PC1 and PC2.

PPMC coefficients derived from PCA loadings were calculated in Microsoft Excel from the full chromatogram [12]. The chromatograms that had been pretreated and mean centered were multiplied by the eigenvector (PC1 or PC2) that provided greatest differentiation of the *Salvia* species. This method enabled a pairwise comparison of the chromatograms based on the variance described by the associated PC.

Results and discussion

Investigation of salvinorin A extraction from *S. divinorum*

Although a variety of methods have been reported for the extraction of salvinorin A from *S. divinorum* [2, 15–17], a systematic study of extraction solvents or extraction times has not yet been conducted. In this research, six solvents with varying polarity were selected (methanol, acetonitrile, acetone, chloroform, dichloromethane, and hexane) and the corresponding solubility parameters are reported in Table 1. Methanol, acetone, and chloroform are currently used for extraction of salvinorin A from *S. divinorum* in forensic laboratories [16], whereas acetonitrile has also been reported [17]. Dichloromethane and hexane were chosen to investigate the extraction ability of less polar solvents. Salvinorin A was extracted from dried *S. divinorum* leaves in each of the six solvents for 1 min. The results of the solvent study are summarized in Table 1.

Chloroform extracted the greatest total mass from *S. divinorum* leaves (28.47±6.79 mg/g), with methanol and dichloromethane extracting the next greatest masses. Dichloromethane extracted the greatest mass of salvinorin A (0.0241±0.0022 mg/g), with acetone and chloroform extracting lesser but comparable masses. Both methanol and acetonitrile extracted a very low mass of salvinorin A (<0.0059 mg/g), whereas hexane extracted no detectable mass.

The lower mass of salvinorin A extracted in methanol may be due to the chemical instability of salvinorin A in protic solvents. Salvinorin A can be converted to a free acid through cleavage of one or more of the ester groups, thereby decreasing the volatility for GC analysis. Moreover, the background in chromatograms of the methanol extracts was high and variable. Alcohols and water are known to react with, and sometimes cleave, siloxane groups from the stationary phase [18]. The background spectra in this study contained various siloxane compounds, such as $(Si(CH_3)_2O–)_n$. Given the variability of the background in the methanol extracts, as well as the low mass of salvinorin A extracted, methanol was discounted as a viable solvent for the extraction of salvinorin A.

As noted previously, hexane extracted no detectable mass of salvinorin A under the standard GC-MS conditions. To increase the mass of sample introduced on the column, hexane extracts were also analyzed using splitless injection. In addition, to increase the selectivity of the analysis, the hexane extract was re-analyzed using selected ion monitoring for *m*/*z* 432 and 94 (corresponding to the molecular ion and base peak of salvinorin A, respectively). However, salvinorin A was not observed under any of these conditions. Hexane was therefore determined to be an

Table 1 Mass of material extracted from *Salvia divinorum* with a 1-min extraction time and different solvents

Solvent	Solubility parameter [20] ($MPa^{1/2}$)	Average total mass extracted (mg/g)	Average mass salvinorin A extracted (mg/g)	Interday RSD (%)
Methanol	29.7	10.78±5.79	0.0026±0.0008	24.72
Acetonitrile	24.3	2.36±3.24	0.0059±0.0006	40.62
Acetone	20.3	6.09±4.01	0.0192±0.0050	71.26
Dichloromethane	19.8	9.55±1.80	0.0241±0.0022	9.62
Chloroform	19.0	28.47±6.79	0.0170±0.0079	17.31
Hexane	14.9	1.13±0.42	ND	ND

mg/g milligrams of material per gram of leaf extracted (mean± standard deviation (n=9)), *RSD* relative standard deviation (n=6), *ND* not detected

unsuitable solvent for the extraction of salvinorin A and was not considered further.

The complexity of each chromatogram was also compared for the remaining solvents. A solvent that efficiently extracts salvinorin A but does not extract other compounds from the plant material is desirable to simplify identification of salvinorin A. Acetone and dichloromethane extracted the fewest compounds, with salvinorins A, B, C, and D predominating. Methanol, acetonitrile, and chloroform extracted a greater number of compounds, as would be expected based on the ability of these solvents to disrupt cell walls and extract polar compounds, such as plant pigments, from inside the plant cells.

Previous research demonstrated that salvinorin A is unstable when stored in solution over 24 to 48 h [17]. The instability was attributed to degradation of salvinorin A, potentially by plant pigments. Therefore, the stability of the reconstituted salvinorin A extracts was investigated after storage at 4 °C for 24 h. The extracts were analyzed in triplicate on two consecutive days and the results of the interday stability study are summarized in Table 1. A decrease in salvinorin A abundance over the 2-day period indicated that degradation was occurring and led to poor precision of salvinorin A peak areas in all solvents. The variability in salvinorin A abundance was the highest in acetone (71.26% relative standard deviation (RSD)) and lowest in dichloromethane (9.62% RSD), suggesting that salvinorin A is more stable in dichloromethane than the other solvents. The polarity of dichloromethane may be insufficient for extraction of plant pigments, as confirmed by the pale color of the extracts, but sufficient for extraction of salvinorin A. Consequently, dichloromethane was chosen as the extraction solvent, as it extracted the greatest mass of salvinorin A, with the least extraction of other plant compounds, and the highest interday precision.

The extraction time was then investigated using dichloromethane as the solvent. Extraction times of 1, 3, 6, 30, 100, and 300 min were considered. Since salvinorin A is found primarily in the trichomes on *S. divinorum* leaves [2], shorter extraction times may be favorable to reduce breakdown of cell walls or extraction of other plant materials, such as plant pigments. Therefore, a 1-min extraction was chosen to represent the shortest extraction time that can be reasonably reproduced. The 300-min extraction time represented a more exhaustive extraction of the *S. divinorum* leaves. Extracts were analyzed by GC-MS and the salvinorin A abundance in each extract was plotted on a semi-logarithmic graph as a function of extraction time, as shown in Fig. 1.

The mass of salvinorin A extracted increased with exposure time to dichloromethane, reaching a maximum between 3 and 6 min (0.0796±0.0153 and 0.0882± 0.0202 mg/g), then remained relatively constant from 30 to 300 min (Fig. 1). There was no statistically significant difference in the mass of salvinorin A extracted using 3- and 6-min extraction times. Therefore, an extraction time of 5 min was chosen for convenience. These results confirm that shorter extraction times result in higher abundance of salvinorin A [2]. However, the standard deviations for triplicate extractions are generally higher for the shorter extraction times. Variations inherent in manual extraction could have a greater impact on shorter extraction times than longer extraction times, resulting in higher standard deviations. However, shorter extraction times are more practical in forensic laboratories. The decrease in abundance of

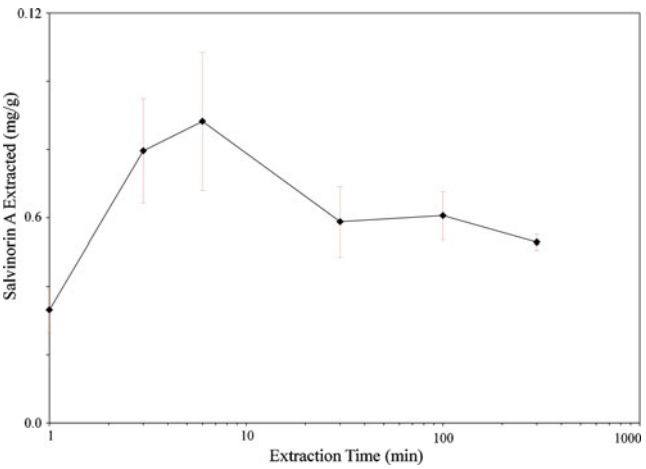

Fig. 1 Semi-logarithmic graph of the mass of salvinorin A versus extraction time. *Error bars* represent±one standard deviation (n=9)

salvinorin A extracted from 6 to 300 min may have been due to the instability of salvinorin A in solution, as demonstrated by Tsujikawa et al. [17]. The reactivity of salvinorin A with other chemical compounds, such as plant pigments, would increase with exposure time to the solvent.

A representative TIC of a dichloromethane extract of *S. divinorum* using the 5-min extraction time is shown in Fig. 2. Although salvinorins B, C, and D were also observed, salvinorin A was the most abundant compound, resulting in a simple chromatogram that allows easy identification of salvinorin A (retention time=26.62 min, molecular ion=*m*/*z* 432).

The extraction efficiency of the optimized procedure was determined using *S. officinalis* as the matrix due to the inherent variability of salvinorin A observed among leaf samples of *S. divinorum* [15]. An extraction efficiency of 97.6% was determined with 1.95±0.20 mg/mL recovered from the 2.00 mg/mL spike. It was also noteworthy that the analytical method was linear over the concentration range investigated of 0.1–12.0 mg/mL salvinorin A.

Differentiation of *S. divinorum* from other *Salvia* species

The developed extraction procedure was used to extract *S. divinorum* and four other *Salvia* species: *S. officinalis, S. guaranitica, S. splendens*, and *S. nemorosa.* The extracts were then analyzed by GC-MS. Representative TICs of the four *Salvia* species are shown in Fig. 3 with chemical identities of selected compounds defined in Table 2. Many of the chemical compounds are terpenes or alkanes commonly found in plant materials.

The chromatograms of *S. divinorum* (Fig. 2) and the other *Salvia* species (Fig. 3) were visually compared. As none of the other *Salvia* species contained salvinorin A (26.62-min retention time), subjective differentiation of *S. divinorum* from all *Salvia* species was possible.

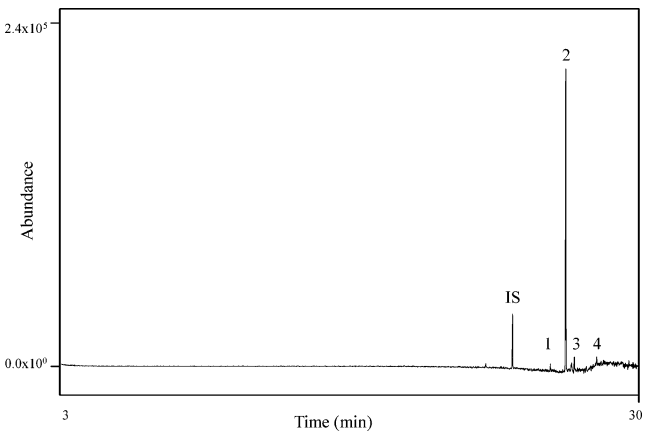

Fig. 2 Representative TIC of *S. divinorum* in dichloromethane extracted using the optimized extraction procedure. *IS* internal standard; *1*, salvinorin B; *2*, salvinorin A; *3*, salvinorin C; *4*, salvinorin D

Principal components analysis was then investigated as an objective method for differentiation and the resulting scores plot is shown in Fig. 4. The first two principal components account for 64% of the total variance among the five *Salvia* species. Replicates of each individual extract of each *Salvia* species are clustered, demonstrating the acceptable reproducibility of the analytical method. However, there is some spread among the three extracts of each species, due to the inherent variability in plant material, as well as variability in the extraction procedure. Nevertheless, each of the *Salvia* species is distinct from *S. divinorum* by visual assessment of the scores plot. *S. splendens* and *S. nemorosa* have overlapping scores and cannot be fully distinguished in PC1 and PC2. Even considering PC3, which accounts for an additional 19% of the variance in the data set, these two species cannot be distinguished through visual assessment of the scores plot.

The positioning of each *Salvia* species on the scores plot can be explained with reference to the loadings plots for PC1 and PC2 (Fig. 5a, b). Compounds loading negatively on PC1 include eucalyptol, thujone, camphor, viridiflorol, carnosol, and totarol. These volatile compounds are present in *S. officinalis* and, hence, this species is positioned negatively on PC1 in the scores plot. In contrast, compounds loading positively on PC1 are higher boiling compounds such as octacosane, eicosane, squalene, falcarinol, columbin, and the insecticide pyridaben. These compounds are present in *S. guaranitica, S. splendens*, and *S. nemorosa*, resulting in the positive position of these species in the scores plot. As salvinorin A does not contribute to the loadings plot for PC1, *S. divinorum* is positioned near zero on the scores plot. Hence, *Salvia* species are distinguished in PC1 based on the volatility of the chemical compounds present. *S. officinalis* is often used for culinary purposes due to the large number of volatile, fragrant compounds, whereas the other *Salvia* species do not contain similar compounds.

Positioning of the species on PC2 can be explained in a similar manner. Compounds loading negatively on PC2 include falcarinol, columbin, and salvinorin A. Accordingly, *S. divinorum* and *S. guaranitica* are positioned negatively in the scores plot. In contrast, the other compounds listed above load positively on PC2. As these compounds are found in *S. officinalis*, *S. splendens*, and *S. nemorosa*, these species are positioned positively on the scores plot. It is important to emphasize that these chemical compounds and, hence, the corresponding PCA scores were identified in an unsupervised manner, thereby allowing an objective assessment of the *Salvia* species.

While a visual examination of the scores plot can give an indication of the association and discrimination of samples, such visual interpretation can still be somewhat subjective.

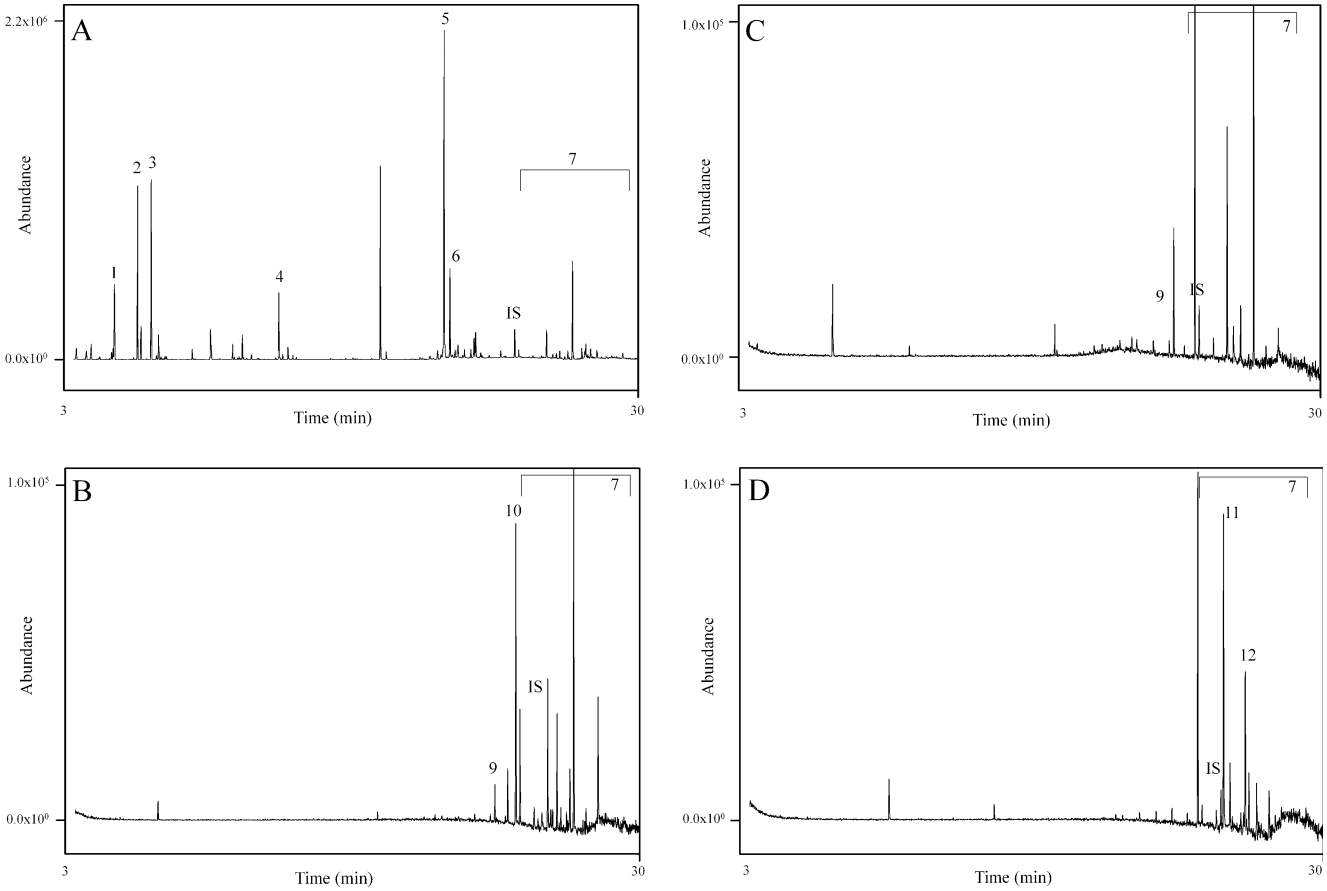

Fig. 3 Representative TICs of **A** *S. officinalis*, **B** *S. splendens*, **C** *S. nemorosa*, and **D** *S. guaranitica*. The chemical identities of numbered compounds are defined in Table 2

Several statistical procedures were used to provide a quantitative assessment of the scores plot. Euclidian distances were calculated between replicate scores of the same species, as well as between the mean scores of *S. divinorum* and the other *Salvia* species. The distances are shown in Table 3 with one example also demonstrated on the PCA scores plot (Fig. 4). In each case, the distance between the mean score of *S. divinorum* and each other *Salvia* species was 4.7–17.5 times greater than that between replicates of those species. Therefore, *S. divinorum* is distinct from the other *Salvia* species as the distance is clearly more than three times the standard deviation among replicates (representing 99.7% confidence limit for a normal distribution [12]).

HCA was then performed on the scores of each *Salvia* species. The branch lengths of the dendrogram, shown in Fig. 6, represent the similarity index among the plant materials [12]. The similarity index between replicates from all extractions ranged from 0.983 for *S. divinorum* to 0.776 for *S. officinalis*, indicating that replicates of *S. divinorum* showed the highest precision and replicates of *S. officinalis* the lowest. These results are expected, as *S. divinorum* contains the smallest number and *S. officinalis* contains the largest number of volatile compounds. The similarity index between *S. divinorum* and the other *Salvia* species ranged from 0.528 for *S. guaranitica* to 0.000 for *S. officinalis*, indicating that *S. guaranitica* was the most similar and *S. officinalis* was the least similar to *S. divinorum*. It is noteworthy that *S. divinorum* and *S. guaranitica* are members of the same clade, whereas *S. officinalis* belongs to a different clade [19].

The Student's *t* test and the Wilcoxon rank-sum test were performed, comparing the mean score of *S. divinorum* with each other *Salvia* species, and the results are reported in Tables S1 and S2, respectively, in the Electronic supplementary material. Using the Student's *t* test, each of the *Salvia* species was statistically distinguishable from *S. divinorum* at the 99.9% confidence level on both PC1 and PC2. Even *S. splendens* and *S. nemorosa*, which were overlapping in the scores plot, were statistically distinguishable at confidence levels of 98% on PC1 and 99% on PC2. Analogous differentiation was provided by the Wilcoxon test at similar confidence levels.

The PPMC coefficients were calculated for all pairwise combinations of the *Salvia* species using the product of the mean-centered chromatogram and the eigenvector of each

Table 2 Tentative chemical identification of compounds of *Salvia* species in Figs. 2 and 3

Salvia species	Peak	Retention time (min)	Chemical identity
Salvia divinorum	8	26.620	Salvinorin A
Salvia officinalis	1	5.024	Eucalyptol
	2	6.129	Thujone
	3	6.770	Camphor
	4	12.858	Veridiflorol
	5	20.737	Carnosol
	6	21.017	Totarol
	7	25.623	Octacosane
		26.858	Eicosane
Salvia guaranitica	11	25.322	Falcarinol
	12	26.334	Columbin
	7	24.302	Hexacosane
		25.623	Octacosane
		26.858	Eicosane
Salvia splendens	9	23.100	Pyridaben
	10	23.707	Squalene
	7	24.302	Hexacosane
		25.623	Octacosane
		26.858	Eicosane
Salvia nemorosa	9	23.100	Pyridaben
	7	25.623	Octacosane
		26.858	Eicosane

PC to provide a quantitative assessment of the scores plot. The mean PPMC coefficients (R_{mean}) for PC1 are summarized in Table 4. Strong correlations were observed among replicate extractions for each species (R_{mean}=0.9477–0.9981). When compared with the theoretical value of 1.0000, slightly lower coefficients were observed due to inherent variations in the plant material as well as in the extraction procedure. Weak correlations were observed between *S. divinorum* and each of the other *Salvia* species (R_{mean}=0.2906–0.4428), in agreement with their position on the PCA scores plot in Fig. 4. In addition, there was strong correlation between *S. splendens* and *S. nemorosa*,

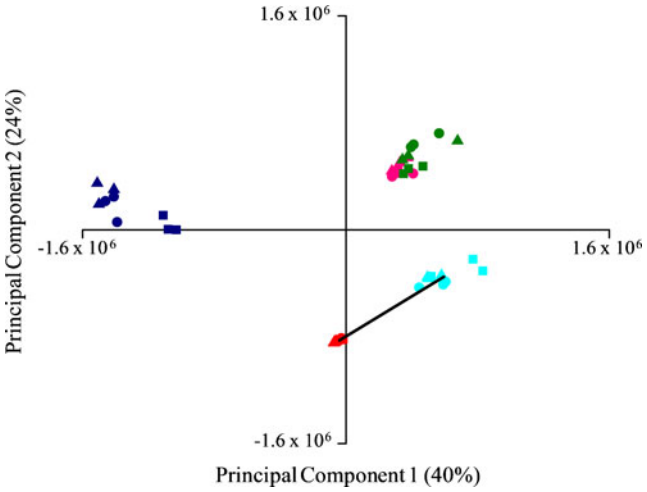

Fig. 4 Scores plot of principal components 1 and 2 for *Salvia* species. Each species is represented by a different color: *S. divinorum* ■, *S. officinalis* ▲, *S. guaranitica* ▼, *S. splendens* ●, and *S. nemorosa* ◆. Shapes represent the three different extractions for each species. Representative example shown of Euclidian distance between *S. divinorum* and *S. guaranitica*

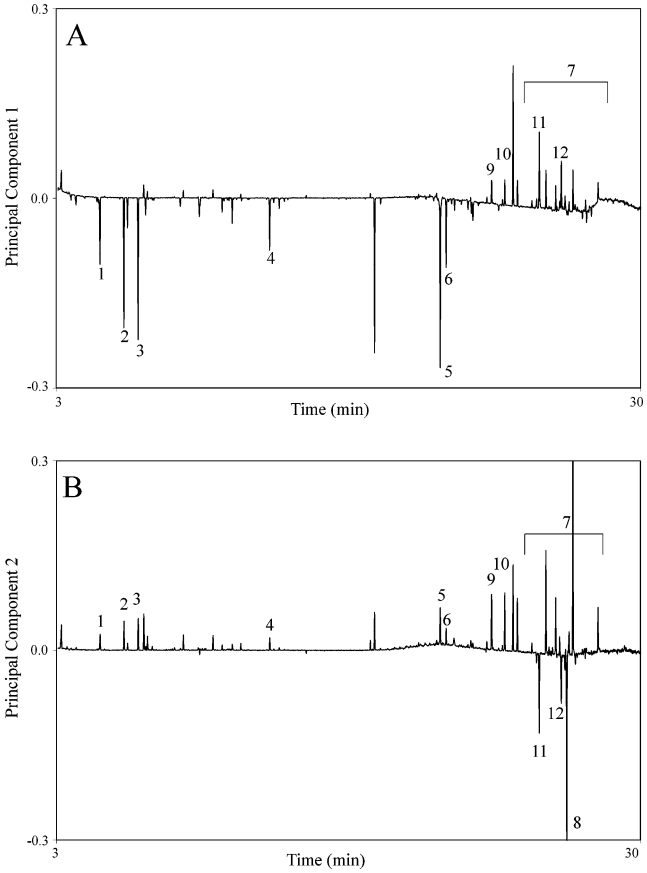

Fig. 5 Loadings plot of principal components **A** PC1 and **B** PC2 for scores plot of *Salvia* species. The chemical identities of numbered compounds are defined in Table 2

Table 3 Euclidian distance between *Salvia divinorum* and other *Salvia* species relative to Euclidian distance between replicates of each species

	Euclidian distance between replicates[a]	Euclidian distance to *S. divinorum*	Euclidian distance ratio
Salvia divinorum	$3.03\pm1.87\times10^4$		
Salvia officinalis	$2.81\pm1.65\times10^5$	1.62×10^6	5.8
Salvia guaranitica	$1.70\pm1.14\times10^5$	7.92×10^5	4.7
Salvia splendens	$1.80\pm1.05\times10^5$	1.48×10^6	8.2
Salvia nemorosa	$7.57\pm4.50\times10^4$	1.33×10^6	17.5

Euclidian distance calculated based on scores from principal components analysis

[a] Mean±standard deviation (n=36)

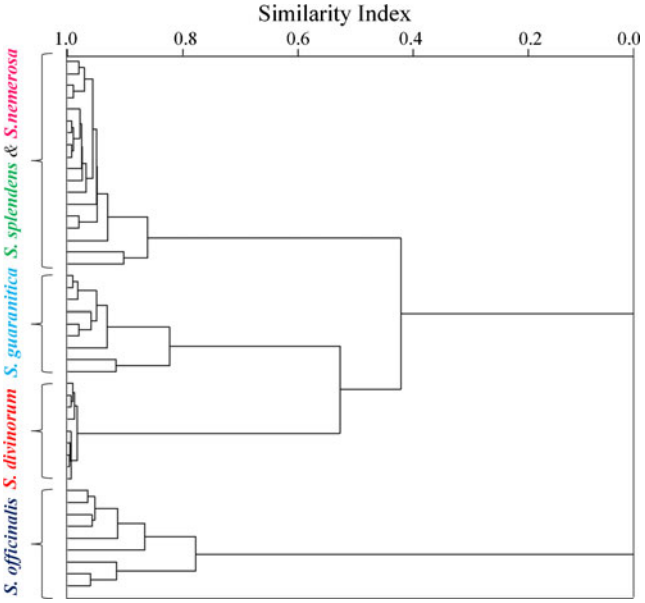

Fig. 6 HCA dendrogram showing similarity among *Salvia* species

which overlap on the scores plot. However, *S. splendens* and *S. nemorosa* are more closely correlated with their corresponding species than with each other, but differentiation is still not possible considering the standard deviations. The negative correlation coefficients derive from the process of mean centering and, hence, are not chemically relevant.

In view of these results, statistical discrimination of *S. divinorum* from the other *Salvia* species was possible using PCA in combination with Euclidian distance, HCA, the Student's *t* test, the Wilcoxon rank-sum test, or PPMC coefficients. Euclidian distances were useful in quantifying the association or discrimination between samples on the PCA scores plot; however, the values were only meaningful relative to the standard deviation. HCA provided a useful visualization tool for the Euclidian distances; however, the similarity index was relative to the population being tested. The statistical methods for hypothesis testing assigned a confidence level to the association or discrimination observed in the PCA scores plot [12]. The Student's *t* test is acknowledged as one of the best tests for minimizing β (type II error) for a fixed α (type I error). However, this test assumes the data are normally distributed with uniform variance, which may not be true for all types of data [12]. In contrast, the Wilcoxon rank-sum test is non-parametric and does not assume a normal distribution. This test is considered to be as effective as the Student's *t* test for normally distributed data and, in cases where the data are not normally distributed, more effective than the *t* test [12]. Similar conclusions were obtained for the Student's *t* test and the Wilcoxon rank-sum test in this work. The PPMC coefficient assesses the absolute similarity between each pair of samples, which results in a single value representing the multivariate data. However, the number of PPMC coefficients increases with the square of the number of samples, making the comparison difficult and time-consuming. Despite these advantages and limitations, each of these statistical procedures was able to provide an objective, quantitative assessment of the association or discrimination of samples by PCA. Among these options, HCA may be most appropriate for forensic purposes as the quantitative information can be presented to the jury in a simple and convenient graphical format.

Conclusions

In this study, salvinorin A was extracted from *S. divinorum* using a 5-min extraction with dichloromethane. This simple and rapid procedure provided an extraction efficiency of 97.6% and an interday precision of 9.6%.

Five *Salvia* species were extracted and analyzed by GC-MS. By visual inspection of chromatograms, *S. divinorum* was differentiated from the other species based on the presence of salvinorin A. Objective differentiation was also demonstrated using the multivariate statistical procedure of PCA. Replicates of each species were closely positioned on the PCA scores plot, with clear distinction of *S. divinorum* from the other *Salvia* species.

Several statistical methods were investigated to provide a numerical evaluation of the positioning in the PCA scores plot. These methods were used for discrimination of S. *divinorum* from four other *Salvia* species and for association of replicates of the same species. Among these

Table 4 Pearson product moment correlation coefficients derived from principal component 1 for *Salvia* species

	Salvia divinorum	*Salvia officinalis*	*Salvia guaranitica*	*Salvia splendens*	*Salvia nemorosa*
S. divinorum	0.9981±0.0008				
S. officinalis	−0.4428±0.0598	0.9834±0.0062			
S. guaranitica	−0.1257±0.0382	−0.6742±0.0286	0.9477±0.0313		
S. splendens	0.1567±0.0922	−0.8432±0.0361	0.5343±0.0556	0.9616±0.0194	
S. nemorosa	0.2906±0.1233	−0.8673±0.0192	0.4825±0.0734	0.9235±0.0416	0.9675±0.0195

Mean±standard deviation (n=81)

methods, HCA and PPMC coefficients provide a bounded range of values (0–1) that indicate the level of discrimination and association. HCA evaluates all of the data simultaneously, but the similarity indices are relative to that data. In contrast, PPMC evaluates the data in a pairwise manner, but the correlation coefficients are absolute. Both of these methods are useful for forensic purposes. HCA can be used to demonstrate association of a submitted sample to a reference collection while PPMC can be used to demonstrate association of a submitted sample and a standard. Among the other methods investigated, the Student's *t* test and Wilcoxon rank-sum test provide for hypothesis testing with an ascribed level of statistical confidence. The null hypothesis, if true, confirms association and, if false, confirms discrimination. The Student's *t* test assumes the data are normally distributed, but the Wilcoxon test has no similar restrictions. Both tests yield the same conclusions at similar confidence levels in these studies.

The knowledge gained from this work will be directly useful to forensic analysts in countries and states where *S. divinorum* or salvinorin A are currently regulated. In addition, the multivariate statistical procedures used for objective association and discrimination in this proof-of-concept study may be more broadly applicable to other controlled substances and to other analytical techniques. Such statistical assessment meets the recommendation for the objective evaluation of forensic evidence highlighted in the report by the US National Academy of Sciences' National Research Council [7].

Acknowledgments The authors acknowledge Dr. Joseph Tychonievich and Prof. Ryan Warner (Department of Horticulture, Michigan State University) for samples of *S. guaranitica*, *S. splendens*, and *S. nemorosa*. Dr. Kathryn Severin (Department of Chemistry, Michigan State University) is also acknowledged for training and use of the GC-MS. This project was partially funded by the Midwest Forensics Resource Center (grant no. SC-10-335) and by a Forensic Sciences Foundation/ National Institute of Justice Student Research Grant. The opinions, findings and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect those of the Department of Justice.

References

1. Epling C, Játiva C (1962) A new species of *Salvia* from Mexico. Bot Mus Leafl 20
2. Siebert D (2004) Localization of salvinorin A and related compounds in glandular trichomes of the psychoactive sage, *Salvia divinorum*. Ann Bot Lond 93:763–771
3. US Department of Justice (2010) D.E.A. Drugs and Chemicals of Concern: *Salvia divinorum* and salvinorin A. Office of Diversion Control, US Drug Enforcement Administration (DEA), Arlington
4. House Bill 6038, Ninety-fifth Legislature State of Michigan, October 2010
5. Salvia divinorum and salvinorin A. U.S. Department of Justice, Drug Enforcement Administration. 02/17/10. Available at www.justice.gov/dea/concern/Salvia%20divinorum%20and%20salvinorin%20A_Feb%202010%20wo%20image.pdf. Accessed September 2011
6. Munro T (2006) The chemistry of *Salvia divinorum*. Dissertation, University of Melbourne, Melbourne, Australia
7. National Research Council of the National Academies (2009) Strengthening forensic science in the United States: a path forward. The National Academies Press, Washington
8. Esseiva P, Anglada F, Dujourdy L, Taroni F, Margot P, Pasquier E, Dawson M, Roux C, Doble P (2005) Chemical profiling and classification of illicit heroin by principal component analysis, calculation of inter-sample correlation and artificial neural networks. Talanta 67:360–367
9. Klemenc S (2001) In common batch searching of illicit heroin samples—evaluation of data by chemometric methods. Forensic Sci Int 115:43–52
10. Brereton R (2007) Applied chemometrics for students. Wiley, West Sussex, England
11. Hupp A, Marshall L, Campbell D, Smith R, McGuffin V (2008) Chemometric analysis of diesel fuel for forensic and environmental applications. Anal Chim Acta 606:159–171
12. Devore J (1990) Probability and statistics for engineering and the sciences. Duxbury Press, Belmont, CA
13. Johnson K, Wright B, Jarman K, Synovec R (2003) High-speed peak matching algorithm for retention time alignment of gas chromatographic data for chemometric analysis. J Chromatogr A 996:141–155
14. Willard M (2011) Forensic analysis of *Salvia divinorum* and related *Salvia* species using multivariate statistical procedures. Dissertation, Michigan State University, East Lansing, MI, USA
15. Gruber J, Siebert D, Der Marderosian A, Hock R (1999) High performance liquid chromatographic quantification of salvinorin A from tissues of *Salvia divinorum* Epling & Jativa-M. Phytochem Anal 10:22–25
16. Jermain J, Evans H (2009) Analyzing *Salvia divinorum* and its active ingredient salvinorin A utilizing thin layer chromatography and gas chromatography/mass spectrometry. J Forensic Sci 54:1–5
17. Tsujikawa K, Kuwayama K, Miyaguchi H, Kanamori T, Iwata Y, Yoshida T, Inoue H (2008) Determination of salvinorin A and salvinorin B in *Salvia divinorum*-related products circulated in Japan. Forensic Sci Int 180:105–109
18. Leger C, Lira H, Paterson R (1996) Preparation and properties of surface modified ceramic membranes. Part II. Gas and liquid permeabilities of 5 nm alumina membranes modified by a monolayer of bound polydimethylsiloxane (PDMS) silicone oil. J Membr Sci 120:135–146
19. Walker J, Sytsma K, Treutlein J, Wink M (2004) *Salvia* (Lamiaceae) is not monophyletic: implications for the systematics, radiation, and ecological specializations of *Salvia* and tribe Mentheae. Am J Bot 91:1115–1125
20. Barton A (1975) Solubility parameters. Chem Rev 75:731