FORENSIC ANALYSIS OF SALVIA DIVINORUM AND RELATED SALVIA SPECIES USING CHEMOMETRIC PROCEDURES

By

Melissa Anne Bodnar Willard

A THESIS

Submitted to Michigan State University in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Criminal Justice

ABSTRACT

FORENSIC ANALYSIS OF SALVIA DIVINORUM AND RELATED SALVIA SPECIES USING CHEMOMETRIC PROCEDURES

By

Melissa Anne Bodnar Willard

Salvia divinorum is a hallucinogenic herb that is internationally regulated. In this study, salvinorin A was extracted from S. divinorum using a 5-minute extraction with dichloromethane. This rapid and simple procedure provided an extraction efficiency of 97.6% and an interday precision of 9.6%. Five Salvia species were extracted and analyzed by gas chromatography-mass spectrometry (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 principal component analysis (PCA). Replicates of each species were closely positioned on the PCA scores plot, with clear distinction of S. divinorum from the other Salvia species. Four plant materials were then spiked with an extract of S. divinorum to simulate an adulterated sample that might be submitted to a forensic laboratory. The unadulterated and adulterated materials were extracted and analyzed by GC-MS. Again, by visual inspection of the chromatograms, the adulterated materials were associated to S. divinorum based on the presence of salvinorin A. Objective association was also demonstrated using PCA, where the adulterated plant materials were closely positioned to S. divinorum on the scores plot, but distinct from the native plant materials. 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.

ACKNOWLEDGMENTS

"The beginning of knowledge is the discovery of something we do not understand."

-Frank Herbert

I would like to thank those who walked beside me in this program, without your support I would not have been able to make it through: John McIlroy, Christy Hay, Jamie Baerncopf, Patty Joiner, Ruth Udey, Tiffany Van de Mark, and Monica Bugeja. Special thanks to my two principle investigators Ruth Waddell Smith and Victoria McGuffin who challenged me to continue to learn and open my mind to so much more than just the science I was studying.

In addition, I would like to thank my parents and the numerous teachers who have touched my life and initiated in me a deep desire to learn and appreciate the intricacies of our universe. I would also like to thank my husband, Tristan, without whom I would have lost perspective many times along this journey.

TABLE OF CONTENTS

List of Tables	vi
List of Figures	viii
List of Abbreviations	X
Chapter 1- Introduction	1
1.2 Legal Status of S divinorum and Salvinorin A	2
1.3 Review of Prior Research	
1.3.1 Salvinorin A Extraction Methods	9
1.3.2 Analysis of Salvinorin A	11
1.3.3 Chemometric Analysis of Plant Materials	
1.4 Requirements for Forensic Analysis and Identification of Salvinorin A	
1.5 Research Objectives	
1.6 References	27
Chapter 2- Theory of Instrumentation and Chemometric Procedures	31
2.1 Gas Chromatography	
2.1.1GC Mobile Phase	
2.1.2 Sample Introduction	
2.1.3 GC Column	
2.1.4 GC Oven and Temperature Program	
2.2 Mass Spectrometry	37
2.2.1 Ionization Chamber and Ionization	
2.2.2 Quadrupole Mass Analyzer	40
2.2.3 Continuous Dynode Electron Multiplier Detector	
2.3 Data Pretreatment Procedures.	
2.3.1 Background Subtraction.	
2.3.2 Smoothing.	
2.3.3 Retention Time Alignment	
2.3.4 Normalization	
2.4 Chemometric Procedures.	
2.4.1 Principal Components Analysis	
2.4.2 Pearson Product Moment Correlation	
2.4.5 Euclidian Distance.	
2.4.4 Student 1-test	
2.4.5 WIICOXON KANK-SUM Test.	
2.4.0 Comparison of Student 1-test and wilcoxon 1est	
2.4.7 mierarchical Cluster Analysis	

2.5 References	62
Chapter 3- Development of an Optimized Procedure for the Extraction of Salvin	orin A
from S. divinorum	64
3.1 Introduction	
3.2 Materials and Methods	65
3.2.1 Extraction Procedure	65
3.2.2 Extraction Efficiency	
3.2.3 GC-MS Analysis	67
3.3 Results and Discussion	
3.3.1 Preliminary Studies	68
3.3.2 Optimization of Extraction Solvent	
3.3.3 Optimization of Extraction Time	
3.3.4 Linear Range and Extraction Efficiency of Salvinorin A from S. divinorum	77
3.4 Conclusion.	
3.5 References	
Chapter 4- Differentiation of S. divinorum from other Salvia Species and Associa	tion of
Adulterated Plant Materials to S. divinorum	85
4.1 Introduction	85
4.2 Materials and Methods	89
4.2.1 Differentiation of S. divinorum from other Salvia Species	89
4.2.2 Adulteration of Plant Materials	89
4.2.3 GC-MS Analysis	90
4.2.4 Data Analysis	
4.3 Results and Discussion	
4.3.1 Differentiation of S. divinorum from other Salvia Species	
4.3.2 Adulteration of Plant Materials	107
4.4 Conclusions	123
4.5 References	125
('hanter 5 ('anclusions and Huture Wark	178

Chapter 5 Conclusions and Future work	128
5.1 Development of an Optimized Procedure for the Extraction of	
Salvinorin A from S. divinorum	
5.2 Differentiation of S. divinorum from other Salvia Species	129
5.3 Association of Adulterated Plant Materials to S. divinorum	
5.4 Future Work	
5.5 References	

LIST OF TABLES

Table 1.1 Legal status of salvinorin A and S. divinorum
Table 1.2 Previous extraction of S. divinorum leaves 10
Table 1.3. Prominent ions of Salvinorins A-D by EI GC-MS, ordered by relative abundance.
Table 1.4. Product ions observed in ESI-MS ⁿ , ordered by relative abundance17
Table 3.1 Total mass of S. divinorum extracted using rotary agitation in four solvents
Table 3.2. Mass of material extracted from S. divinorum with a one minute extraction time and different solvents. 72
Table 4.1 Tentative chemical identification of compounds of Salvia species in Figure 4.195
Table 4.2. Euclidian distance between Salvia species and S. divinorum relative to Euclidian distance between replicates of each species
Table 4.3. T-statistic (with associated degrees of freedom) for comparison of Salvia species. Cell color represents confidence level on heat map
Table 4.4. Wilcoxon rank-sum statistic for comparison of Salvia species. Cell color represents confidence level on heat map in Table 4.3
Table 4.5A. Pearson product moment correlation coefficients derived from principal component 1 for Salvia species. 105
Table 4.5B. Pearson product moment correlation coefficients derived from principal component 2 for Salvia species
Table 4.6 Tentative chemical identification of compounds in plant materials in Figure 4.5110
Table 4.7. Euclidian distance of (A) unadulterated samples to S. divinorum relative to Euclidian distance between replicates, (B) adulterated samples to S. divinorum relative to Euclidian distance between adulterated and unadulterated samples
Table 4.8. Unequal variance t-statistic (with associated degrees of freedom) for comparison of (A) unadulterated plant materials, (B) unadulterated plant materials with adulterated plant materials, and (C) adulterated plant materials. Cell color represents confidence level on heat map in Table 4.3

Table 4.9. Wilcoxon	rank-sum st	tatistic fo	or compa	rison of ((A) una	dulterate	ed plant	t materia	als, (B)
unadulterated plant r	naterials wit	th adulte	rated pla	nt materi	als, and	l (C) adu	lterated	l plant	
materials. Cell color	represents c	onfidenc	ce level o	n heat m	ap in T	able 4.3	• • • • • • • • • • • • • • • • • • • •	-	118
T 11 410 D	1 .		1	cc	. 1 .	1.0	• •	1	. 1

LIST OF FIGURES

Figure 1.1 Sampling of <i>Salvia</i> clades I and II [8]. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this thesis
Figure 1.2 . Chemical structure of salvinorin A, with the skeleton structure for an isoprene unit shown in bold
Figure 1.3. Mass spectrum of Salvinorin A by GC-MS13
Figure 1.4. Structures of salvinorin and divinatorin compounds14
Figure 2.1. Components of a gas chromatography system
Figure 2.2 A) Components of a mass spectrometry system B) Electron ionization source C) Quadrupole mass analyzer
Figure 2.3. A) Rise in the baseline of sample TIC B) Extracted ion profile of m/z 73, 193, 207, 221, 341, 355 C) Asymmetric sigmoid equation D) TIC after background subtraction E) TIC after data smoothing
Figure 2.4. First derivative zero crossing
Figure 2.5. A) Comparison of sample to target B) Peak matching window C) Alignment of sample to target
Figure 2.6. A) Nearest-neighbor linkage B) HCA dendrogram with highest similarity proportional to branch length
Figure 3.1. Extract of <i>S. divinorum</i> in dichloromethane using rotary agitation1) salvinorin B, 2) salvinorin A, 3) salvinorin C, 4) salvinorin D
Figure 3.2. Representative total ion chromatograms of <i>S. divinorum</i> extractions in (A) methanol, (B) acetonitrile, (C) acetone, (D) chloroform, (E) dichloromethane, (F) hexane IS = internal standard, 1) salvinorin B, 2) salvinorin A, 3) salvinorin C, 4) salvinorin D73
Figure 3.3. Semi-logarithmic graph of the mass of salvinorin A versus extraction time. Error bars represent \pm one standard deviation (n=9)
Figure 3.4. Representative total ion chromatogram of <i>S. divinorum</i> in dichloromethane extracted using the optimized extraction procedure. IS = internal standard, 1) salvinorin B, 2) salvinorin A, 3) salvinorin C, 4) salvinorin D

Figure 3.5. GC-MS calibration curve of salvinorin A, showing linearity ($y = 4 \times 10^7 x - 5 \times 10^6$,
$R^2 = 0.9951$) over the concentration range investigated of $0.10 - 12.00$ mg/mL salvinorin A81
Figure 4.1. Representative total ion chromatograms of (A) <i>S. officinalis</i> , (B) <i>S. splendens</i> , (C) <i>S. nemorosa</i> , and (D) <i>S. guaranitica</i>
Figure 4.2. Scores plot of principal components 1 and 2 for <i>Salvia</i> species. Representative example shown of Euclidian distance between <i>S. divinorum</i> and <i>S. guaranitica</i>
Figure 4.3. Loadings plots of (A) principal component 1 and (B) principal component 2 for scores plot of <i>Salvia</i> species
Figure 4.4 Hierarchical cluster analysis dendrogram showing similarity among <i>Salvia</i> species
Figure 4.5 Representative total ion chromatograms of unadulterated (A) <i>C. sativa</i> , (B) <i>N. tabacum</i>
Figure 4.6 Representative total ion chromatograms of adulterated plant materials (A) <i>S. officinalis</i> , (B) <i>S. divinorum</i> , (C) <i>C. sativa</i> , (D) <i>N. tabacum</i> 109
Figure 4.7. Scores plot of principal components 1 and 2 for unadulterated plant materials (closed interior). Scores for adulterated plant materials (open interior) are projected onto the plot (see text). Representative example shown of Euclidian distance between adulterated <i>C. sativa</i> and both unadulterated <i>S. divinorum</i> and unadulterated <i>C. sativa</i>
Figure 4.8. Loadings plots of (A) principal component 1 and (B) principal component 2 for scores plot of unadulterated and adulterated plant materials
Figure 4.9. Hierarchical cluster analysis dendrogram showing similarity among unadulterated and adulterated plant materials

LIST OF ABBREVIATIONS

0	degree
±	plus or minus
μ	micro
С.	Cannabis
С	celsius
DC	direct current
DEA	US Drug Enforcement Administration
EI	electron ionization
EIC	extracted ion chromatogram
EIP	extracted ion profile
ESI	electrospray ionization
eV	electron volts
g	gram
GC	gas chromatography
GCP	graphite carbon powder
HCA	hierarchal cluster analysis
KOR	kappa opioid receptor
L	liter
LC	liquid chromatography
m	meter
Μ	molarity
min	minute

MS	mass spectrometry
MW	molecular weight
m/z	mass-to-charge
n	nano
<i>N</i> .	Nicotiana
Ν	normal
Pa	Pascal
PC	principal component
PCA	principal component analysis
PPMC	Pearson product moment correlation
RF	radio frequency
RSD	relative standard deviation
<i>S</i> .	Salvia
8	second
SOP	standard operating procedure
TIC	total ion chromatogram
v	volume

CHAPTER 1

INTRODUCTION

Salvia divinorum is a hallucinogenic perennial herb from the mint family (Lamiaceae or Labiatae) that has recently become of great interest in the field of forensic science [1]. Thought to have originated in Oaxaca, Mexico, the plant was used for centuries by the Mazatec Indians for medicinal and spiritual practices [2]. The fresh leaves were chewed or brewed into tea or other infusions; however, in the present day the dry leaves are typically smoked [3]. According to the most recent figures released by the National Survey on Drug Use and Health, in 2006 more than 756,000 residents of the United States over the age of 12 smoked *S. divinorum* and 1.8 million had smoked the plant in their lifetime [4]. Spiking *S. divinorum* onto other plant materials, such as *Cannabis sativa* (marijuana), and then subsequently smoking the adulterated material, is another known route of administration. A recent case study reported the toxic psychosis of a patient whose marijuana cigarette had been adulterated with the leaves and leaf extract of *S. divinorum* [5].

Forensic identification of *S. divinorum* can be challenging because the plant is one of nearly a thousand *Salvia* species. Two main methods exist for classification of species in a plant genus, one based on physical morphology of leaves, stems, etc. and the other based on profiling genetic material in the plant. The method based on the physical characteristics of the plant is an older method and its ability to correlate morphology with the origin of the plants is in question, given the more recent findings of the genetic profiling method [6]. The *Salvia* genus is particularly difficult to characterize and it has been reported that differentiation of *S. divinorum* from other *Salvia* species is not possible through visual inspection of the plant, since the leaves

have no readily discernable characteristics [7]. Using the genetic method with the *rbcL* and *trnL-F* regions of chloroplast DNA, it was shown by Walker *et al.* that the *Salvia* genus does not descend from one ancestor but rather is polyphyletic [8]. The genus is comprised of three distinct lineages: clade I having predominantly Old World (Europe and Africa) with some New World lineage (the Americas), clade II having exclusively New World lineage, and clade III having an independent Asian lineage. A small sampling of clades I and II are shown in Figure 1.1 [8]. *S. divinorum* derives from clade II and is the subject of this dissertation.

Other *Salvia* species are also of ethnobotanical, horticultural, and culinary interest. For example, *S. officinalis* (clade I) is used as the common cooking sage, while *S. guaranitica* (clade II), *S. splendens* (clade II), and *S. nemorosa* (clade I) are used in landscaping and can customarily be found in commercial greenhouses [6, 8]. *Salvia* species of clade III are of rare lineage and not commonly available in North America. Traditionally, plant species related to sage are known to have mild sedative effects. However, *S. divinorum* is the only *Salvia* species known to have hallucinogenic properties [1-3,7]. There have been reports of *S. splendens* causing mild hallucinations as well, however these have not been scientifically substantiated [9].

1.1 Physiological Activity and Physical Effects of Salvinorin A

The active component in *S. divinorum*, salvinorin A, is considered to be the most potent known hallucinogen



Clade I

Clade II

Figure 1.1. Sampling of *Salvia* clades I and II [8]. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this thesis.

of natural origin, rivaling even the semi-synthetic hallucinogen lysergic acid diethylamide (LSD) [1-3]. Currently, salvinorin A is only known to exist in *S. divinorum* [6].

The chemical structure of salvinorin A (Figure 1.2) is classified as a neoclerodane diterpenoid. The general diterpenoid structure is comprised of four isoprene units (the skeleton structure for an isoprene unit is shown in bold in Figure 1.2). Terpenoids function as lipids that are one of the most abundant secondary metabolites in plants. Generally occurring in the essential oils of plants, terpenoids are heavily used for their aromatic qualities.

The hallucinogenic effects associated with salvinorin A are thought to be due to its strong agonist activity with the kappa opioid receptor (KOR). KORs are found in neurons associated with delivering pain signals to the brain, as well as in the spinal cord, and are heavily distributed throughout the brain. Salvinorin A has a high affinity for the KOR, indicated by the low dissociation constant of $K_i=1.0 \pm 0.1$ nM (cloned human KOR) [10]. KOR activation by the endogenous opioid receptor agonists (dynorphins) are known to produce symptoms of dysphoria, analgesia, sedation, and distortions in perceptions [11].

Effective doses of salvinorin A are reported to be extremely low, typically 200–1000 μ g, with a half-life of eight minutes in the brain [12]. In addition, recent research has shown that <10 μ g may account for its psychoactivity in the human brain [12]. The concentration (EC₅₀) at which 50% of the maximum response is observed is 3.1 ± 1.0 nM salvinorin A in humans, indicating an extremely high potency [13].

The effects of salvinorin A vary, although many users report calming sensations, hysterical laughter, hallucinations, out-of-body experiences, and loss of consciousness [3, 7, 14].



Figure 1.2 . Chemical structure of salvinorin A, with the skeleton structure for an isoprene unit shown in bold.

In a study conducted by Gonzalez *et al.*, subjects not only displayed the stimulant properties characteristic of psychedelic drugs, but also reported feelings of fatigue, weakness, and sluggishness that are more commonly associated with sedative drugs [14].

1.2 Legal Status of S. Divinorum and Salvinorin A

S. divinorum is internationally regulated in Australia, Denmark, Finland, Germany, Italy, Spain, and Sweden [15]. However, in the United States the US Drug Enforcement Administration (DEA) has listed *S. divinorum* under Drugs and Chemicals of Concern, but neither *S. divinorum* nor salvinorin A have been federally regulated under the Controlled Substances Act [15]. As of 2011, 26 individual states have regulated either the plant or salvinorin A (shown in Table 1.1) and several others have pending legislation [15, 16].

As shown in Table 1.1, Florida, Hawaii, Illinois, Kansas, Kentucky, Michigan, Missouri, Nebraska, North Dakota, Ohio, Oklahoma, and Virginia have classified *S. divinorum* and salvinorin A as Schedule I controlled substances. Schedule I controlled substances are considered to have the highest risk of abuse and have no acceptable medical usage. Delaware and Mississippi have classified only *S. divinorum* as a Schedule I controlled substance and do not regulate salvinorin A. In Minnesota the sale and possession of both salvinorin A and *S. divinorum* are considered misdemeanors. While, in South Dakota, possession of less than two ounces of either *S. divinorum* or salvinorin A is a misdemeanor and possession of greater than two ounces is a felony. In Tennessee, possession of salvinorin A is a misdemeanor, whereas in

Table 1.1. Legal status of salvinorin A and S. divinorum

State	Classification	Substance Regulated
Florida, Hawaii, Illinois, Kansas, Kentucky, Michigan, Missouri, Nebraska, North Dakota, Ohio, Oklahoma, and Virginia	Schedule I	Salvinorin A and S. divinorum
Delaware and Mississippi	Schedule I	S. divinorum
Minnesota and South Dakota	Misdemeanor/ Felony	Salvinorin A and S. divinorum
Tennessee and Wisconsin	Misdemeanor/ Felony	Salvinorin A
Georgia, Louisiana, North Carolina, and West Virginia	Misdemeanor	Salvinorin A and <i>S. divinorum</i> - (only if intended for human consumption)
California, Maine, and Maryland	Misdemeanor	Salvinorin A and <i>S. divinorum</i> -(only if sold to minors)

Wisconsin possession is not regulated, but manufacturing and sale are illegal. In Georgia, Louisiana, North Carolina, and West Virginia possession of *S. divinorum* and salvinorin A are legal if not intended for human consumption. Hence, growing *S. divinorum* is permitted for aesthetic, landscaping or decorative purposes. California, Maine, and Maryland have passed legislation against possession by minors or selling to minors, although possession by adults is permitted.

Although neither *S. divinorum* nor salvinorin A are federally regulated in the United States at this time, regulation occurring in the near future is a possibility given the increasing number of individual states that are passing their own legislation. States in which only *S. divinorum* is regulated, such as Delaware and Mississippi, may have problems in the forensic identification of the controlled substance. Identification may be challenging if salvinorin A was extracted from *S. divinorum* and then spiked onto a non-regulated plant material.

1.3 Review of Prior Research

There are inherent difficulties in the extraction and analysis of any chemical; however plant materials have several unique complications. The large number of components present in the plant can create difficulties in chromatographic analysis as compounds with similar boiling points may co-elute. There may also be difficulties in detection of the desired chemical analytes due to the matrix interferences of other components present in the plant. In addition, chemical reactions can occur between the individual compounds in the plant, resulting in a composition that is not representative of the original sample. An example of this has been reported for the chemical conversion of salvinorin A to salvinorin B in blood esterase [17]. The following summary of research details the progress in optimizing the extraction of salvinorin A from *S*.

divinorum and subsequent analysis. In addition, the use of chemometric procedures for simplification of data obtained from various plant materials is briefly reviewed.

1.3.1 Salvinorin A Extraction Methods

Salvinorin A was first isolated by Ortega *et al.* in the early 1980s, through extraction of dried, pulverized *S. divinorum* leaves with boiling chloroform [18]. However, neither the exact extraction method nor the time the leaves remained in contact with the chloroform were reported. Many procedures differing in solvent type, temperature and time of exposure to the leaves have since been used for extraction of salvinorin A from *S. divinorum*, as summarized in Table 1.2.

In a more recent study, Jermain *et al.* surveyed five forensic laboratories and each reported a different method for the extraction of salvinorin A from *S. divinorum* [23]. Methods included extracting *S. divinorum* leaves with methanol, chloroform, or acetone at ambient temperature, boiling the leaves in chloroform for 10 minutes, or a basic extraction with 1N sodium hydroxide partitioned to dichloromethane. All extracts were analyzed by gas chromatography-mass spectrometry (GC-MS) and the extraction methods were compared in terms of the abundance of the salvinorin A peak in the resulting chromatogram. The chromatogram for the one-minute extraction in chloroform resulted in the highest abundance of the salvinorin A peak. Although interesting, this study did not systematically evaluate extraction solvents or extraction times, but compared standard protocols from five laboratories.

Tsujikawa *et al.* reported another extraction method of *S. divinorum* using a repetition of shaking, ultrasonication, and centrifugation of the leaves in acetonitrile and a total extraction time of 20 minutes [24]. The solution was then treated with graphite carbon powder (GCP) and centrifuged to remove pigments in the extract, such as chlorophyll. This study found that the

Table 1.2. Previous extraction of *S. divinorum* leaves

Leaf Material	Extraction Procedure	Solvent	Temperature	Time	Ref.
Lyophilized and ground fresh leaf	Soxhlet	Ethyl ether	Boiling	24 hr	[19]
Fresh leaf then dried and powdered	Sequentially dipped in beakers	Chloroform	Ambient	30s, 4hr	[20]
Fresh leaf		Acetonitrile/water (50:50, v/v)	Ambient		[21]
Ground fresh leaf	Immersed in saturated aqueous ammonium buffer (pH 9.5)	Chloroform and 2- propanol (9:1, v/v)	Ambient	30 min	[22]
Dried and crushed leaf	Sonication	Methanol	Ambient		[22]

untreated extract, when stored in acetonitrile at 4 °C without light shielding, was subject to degradation of salvinorin A over a 48-hour period. Almost no degradation was observed with the GCP-treated extracts stored under the same conditions, indicating that the pigments may lead tothe degradation of salvinorin A. While this study involved a complicated and long extraction method, the results reinforce the hypothesis that shorter extraction times are potentially more desirable as demonstrated by Jermain *et al.*. A shorter extraction time, without pulverization of the *S. divinorum* leaves, would avoid the extraction of pigments and hence degradation of salvinorin A.

1.3.2 Analysis of Salvinorin A

Following the initial isolation of salvinorin A, Ortega *et al.* determined the structure of salvinorin A by infrared (IR) and nuclear magnetic resonance (NMR) spectroscopies and confirmed it by x-ray crystallography [18]. A variety of analytical techniques have since been used to analyze *S. divinorum* for salvinorin A content including liquid chromatography (LC) with UV absorbance detection [25], GC-MS [22], reversed-phase LC-MS with atmospheric pressure chemical ionization (APCI) in the negative ion mode [26], and LC multistage MS^n (n = 1 – 6) with electrospray ionization [21, 26].

In GC, a gaseous mobile phase is used to transport the sample components through a column containing either a liquid or solid stationary phase. Separation is based on the volatility of the components as well as their rate of partitioning between the two phases. However, in LC a liquid mobile phase is used to transport the sample through a solid stationary phase and separation is based on the polarity of the compound and the rate of partition between the two phases. Both GC and LC instruments can be coupled to a mass spectrometer (MS), which acts as

the detector. Fragmentation of the separated sample components occurs in the MS, allowing for the definitive identification of each component. Both GC-MS and LC-MS can be considered to be complimentary technique; GC-MS allows for the analysis of small, nonpolar volatile compounds while LC-MS is used for the analysis of larger, non-volatile compounds, with higher polarity. For example, GC-MS is useful for the analysis of terpenes and flavonoids in plant materials and LC-MS is useful for the analysis of lipids, phenols and acids.

Giroud *et al.* were the first to identify a seized specimen of *S. divinorum*. This was accomplished through a methanol extraction of dried leaf material and subsequent detection of salvinorin A by GC-MS, as shown in Figure 1.3 [20]. The extraction time for this study was not reported. The molecular ion (mass-to-charge ratio (m/z) 432), base peak (m/z 94), and other significant fragment ions were used to confirm the identity of salvinorin A, in the absence of an authentic standard. In addition, Giroud *et al.* developed a method involving acetylation of the carboxylic acid groups in an extract of the fresh *S. divinorum* leaf prior to GC-MS analysis. The acetylation process increased the volatility of the extract and decreased the likelihood of sample degradation in the injection port of the GC.

S. divinorum also contains other salvinorin and divinatorin compounds Figure 1.4 that are closely related to salvinorin A in structure, and also appear to be unique to *S. divinorum* [7, 25]. Prominent ions of salvinorins A-D analyzed using GC-MS with electron ionization (EI) are shown in Table 1.3 [23]. In a separate study, salvinorin A, five closely related salvinorins (B-E), and three divinatorins (A-C) identified in *S. divinorum* were separated



Figure 1.3. Mass spectrum of Salvinorin A by GC-MS [22]



Salvinorins		Divinatorins	
A $R = COCH_3$	$\mathbf{C} \mathbf{R}_1 = \mathbf{COCH}_3 \mathbf{R}_2 = \mathbf{COCH}_3$	$\mathbf{A} \mathbf{R}_1 = \mathbf{C}\mathbf{H}_3$	$R_2 = OH$
$\mathbf{B} \mathbf{R} = \mathbf{O}\mathbf{H}$	$\mathbf{D} \mathbf{R}_1 = \mathbf{OH} \qquad \mathbf{R}_2 = \mathbf{COCH}_3$	B $R_1 = CH_2OH$	$R_2 = OH$
	$\mathbf{E} \mathbf{R}_1 = \mathrm{COCH}_3 \mathbf{R}_2 = \mathrm{OH}$	$\mathbf{C} \mathbf{R}_1 = \mathbf{CH}_2\mathbf{O}\mathbf{C}\mathbf{O}\mathbf{C}\mathbf{H}_3$	$R_2 = H$

Figure 1.4 Structures of salvinorin and divinatorin compounds

Compound	Base Peak (m/z)	Other Prominent Ions (m/z)	MW
Salvinorin A	94	43, 273, 55, 121, 81, 107	432
Salvinorin B	43	94, 107, 291, 55, 81, 121	390
Salvinorin C	43	94, 121, 81, 372, 399	474
Salvinorin D	43	94, 81, 121, 55, 400, 163	432

Table 1.3. Prominent ions of Salvinorins A-D by EI GC-MS, ordered by relative abundance [23]

by LC-MSⁿ (n = 1–6) with electrospray ionization [21]. Salvinorins A, D, and E have the same molecular weight (MW) and, hence, are isobaric in GC-MS and single-stage LC-MS, presenting problems for forensic identification. However, salvinorins A, D, and E have subsequent dissociation patterns that are unique in MS^2 , MS^3 , and MS^4 , allowing for a distinction between the salvinorins [21]. Tentative fragmentation pathways for salvinorins A, C, and F have also been proposed and are useful for identification [21]. Prominent fragments for the salvinorins and divinatorin B are listed in Table 1.4 [21].

Salvinorin A has been detected in spiked samples of human and rhesus monkey urine and plasma by reversed-phase LC-MS with APCI in the negative ion mode [26]. In a similar study, reversed-phase LC-MS/MS with electrospray ionization has been reported for detection of salvinorin A in rat and dog serum [17]. In these studies, standards of salvinorin A were spiked into the respective bodily fluids and then analyzed to determine if detection was possible. At a concentration of 0.1 μ M, salvinorin A was detectable in the serum [26]. When 4.4 μ M salvinorin A was spiked into both human and monkey plasma and observed over a time period of 70 minutes, salvinorin A concentration decreased at seemingly the same rate that the salvinorin B concentration increased [17]. The researchers hypothesized that esterase enzymes in the blood may be converting salvinorin A to salvinorin B. This theory could explain why salvinorin A has a short half-life of 8 minutes in the human brain [12].

Compound	Precursor ion m/z [M+H] ⁺	MS ²	MS ³	MS ⁴
Salvinorin A	433	373	355	337, 323, 309, 295
			341	313
			313	295, 285, 267
Salvinorin B	391	373	355	337, 323, 309, 295
			341	313
			313	295, 285, 267
Salvinorin C	475	415	355	323, 337
			261	
		457	415	
Salvinorin D	433	415		
		373	341	
			313	267, 285, 295
Salvinorin E	433	415	397	
			295	277
Salvinorin F	375	357	325	251, 297
Divinatorin B	363	345	327	267, 295

Table 1.4. Product ions observed in ESI-MSⁿ, ordered by relative abundance [21]

Although the previous methods show promise for identification of *S. divinorum*, the practicality of using LC-MSⁿ in a controlled substance section of a crime laboratory is low. These sections currently have limited access to LC-MS instrumentation. Hence, these methods may not be the most practical for routine forensic identification of salvinorin A in *S. divinorum*. For the purpose of this dissertation, GC-MS will be the primary technique used to analyze *S. divinorum* samples. A more detailed description of the GC-MS system will be given in Chapter 2 of this dissertation.

1.3.3 Chemometric Analysis of Plant Materials

Chemometric procedures have been used in analytical and physical chemistry for a number of years and are growing in popularity among forensic analysts [27-30]. Chemometric procedures allow for a non-subjective means of comparing forensic evidence. For example, rather than visually comparing chromatograms of controlled substances, chemometric procedures can be used for an objective comparison. These procedures are especially useful in cases involving chromatograms of plant materials, which, due to the complexity of the chromatograms, are difficult to visually discriminate. In GC-MS or LC-MS analysis, a large number of (time, intensity) data is collected where each data point is a variable (the GC-MS chromatograms collected for this dissertation have approximately 5,000 data points each). Chemometric procedures can simplify these complex data sets and present the information of a large number of variables in only two or three dimensional plots, making comparison simpler.

For chemometric procedures to be successful however, variance must arise only from chemical differences in the samples. Therefore, data pretreatment steps are necessary to remove non-chemical sources of variance prior to performing chemometric procedures [27]. Depending on the type of data and the chemometric procedure being performed, different pretreatment steps may be necessary. Specifically for GC data, procedures such as background subtraction, retention time alignment, and normalization are commonly used [27]. A more detailed description of data pretreatment steps will be given in Chapter 2 of this dissertation.

Principal component analysis (PCA) is a chemometric procedure used for the association and discrimination of complex samples [27-30]. Using PCA, a complex data set (such as GC-MS data) is reduced to a few principal components that represent the greatest contributions to variance among the samples [28]. The PCA scores plot (e.g. graph of principal component 1 versus principal component 2) can then be used to examine the association or discrimination of the samples. Samples that are chemically similar have similar values for the principal components and cluster together, but chemically different samples do not. A more detailed description of PCA is included in Chapter 2 of this dissertation.

PCA has been used analytically and forensically as a means of identifying the chemical constituents that provide the greatest differentiation among samples. It has been used for the chemical profiling of 3371 seized heroin samples [29]. GC analysis of six target alkaloids in the heroin samples was performed and PCA applied to identify samples of similar chemical profiles (approximately 20 chemical classes were found). A correlation value for each sample was calculated and the combination of correlation values and PCA was used for development of an artificial neural network (ANN). The ANN was then used to test 468 known samples, resulting in a successful classification rate of 96%.

In addition, PCA has been used to examine batch-to-batch variation in heroin distribution to differentiate various sources of heroin [30]. Three batches of seized illicit heroin samples were

divided into sub-groups. Cutting substances (such as sucrose, caffeine, procaine, and glucose) were added to each sub-group to simulate a dealer-user network and determine if dilution of the sample affected the differentiation. After analysis by GC-MS, PCA was performed and clustering of each of the three original batches of seized tablets resulted despite different proportions of cutting agents.

A potential problem with PCA is that interpretation of the scores plot is based on visual assessment which can be subjective. This can be overcome using additional statistical procedures to investigate association and discrimination of the samples based on the PCA scores. Examples include Euclidian distance measurement, student t-test, Wilcoxon rank-sum, and hierarchal cluster analysis (HCA). The Euclidian distance is the numerical distance between the means of two samples scores. Samples that are close to one another on the PCA scores plot have short Euclidian distances, whereas samples that are distinct have longer Euclidian distances. The student t-test and the Wilcoxon rank-sum test can be used to assess discrimination of samples at given confidence levels. The t-test assumes the data are normally distributed, whereas the Wilcoxon test makes no assumptions regarding the distribution of the data. The visual representation of association between samples in HCA can be useful in demonstrating the similarities between samples and is complimentary in confirming the association observed in the PCA scores plot.

Pearson product moment correlation (PPMC) coefficients provide an additional means to compare samples that might be particularly useful in a forensic setting. Coefficients are calculated for pair-wise comparisons of chromatograms, for example, a questioned sample and a reference standard, allowing comparison of 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. A coefficient of ± 0.80 or greater indicates 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 [28]. Furthermore, as PPMC coefficients assess similarities among samples and PCA identifies differences among samples, the two procedures can be considered complementary. As such, PPMC coefficients may be beneficial for demonstrating the correlation between pairs of chromatograms when associating or discriminating samples.

1.4 Requirements for Forensic Analysis and Identification of Salvinorin A

Two analytical techniques are needed for the definitive forensic identification of a controlled substance [32]. Identification is generally accomplished through either a presumptive test (identifies the class of substance) or a selective test (tentative identification of substance), combined with a confirmatory test (definitive identification). Gas chromatography is classified as a selective test and mass spectrometry as a confirmatory test; hence, the combination of GC-MS is suitable for the identification of most controlled substances that are sufficiently volatile to be analyzed with this technique.

In forensic laboratories in states where *S. divinorum* or its active component are regulated, extraction methods for salvinorin A are widely varied [23]. An extraction and analysis procedure for a forensic laboratory should be rapid, reproducible, and simple to perform. In the event of further state regulation and the potential of federal regulation, an extraction method and analysis technique meeting the above criteria would be necessary to develop a standard operating procedure (SOP). Currently, no studies have systematically compared extraction solvents,

procedures, or times, which would make adopting a SOP for the extraction of salvinorin A from *S. divinorum* difficult.

None of the previous methods for extracting salvinorin A from S. divinorum, (summarized in Table 1.2), give an indication of the extraction efficiency of the method, only that salvinorin A was present in the solution after extraction. Methods that used long extraction times or complicated procedures, such as the Soxhlet [19] or chloroform extractions [20], may not be practical for a forensic laboratory due to their case loads and time restrictions. In addition, methods that involve pulverizing the S. divinorum leaves or long extraction times appear to extract a greater number of other plant components along with salvinorin A. Previous research by Seibert showed that the preponderance of salvinorin A is located in the trichomes on the outside of the S. divinorum leaves [20]. If a component is on the outside, less time would be needed for diffusion into the extraction solvent. This would suggest that short extraction times (in keeping with studies by Jermain *et al.* and Tsujikawa *et al.*) would be sufficient for extraction of salvinorin A. With short extraction times, there is little to no breakdown of cell walls or extraction of other plant materials (such as plant pigments). If forensic identification of salvinorin A is the goal of the extraction, complexity of the extract should be avoided to allow for a simple identification.

Previous research by Jermain *et al.* compared the procedures for extraction of salvinorin A from *S. divinorum* of five forensic laboratories [23]. However, no internal standard or other techniques allowing for quantification of the abundance of salvinorin A were reported, nor were data pretreatment steps taken to account for the variations in the instrument or injection volumes between analyses.

The research performed by Giroud *et al.* used an extraction technique for the fresh *S. divinorum* leaf that was complicated and time consuming, and may not be practical for routine analysis [22]. For the extraction of the dried *S. divinorum* leaf, the chemical stability of salvinorin A in a protic solvent, such as methanol, could be an issue. Potentially, methanol could convert salvinorin A to a free acid through cleavage of one or more of the ester groups. Alcohols and water are known to react and cleave siloxane groups from the stationary phase of GC-MS columns [33] which would influence the reproducibility of the GC-MS analysis of methanol extracts. Both chromatograms shown by Giroud *et al.* [22] indicated that additional components of the plant, rather than just salvinorin A, were extracted. This could lead to the problems mentioned above, such as co-elution of components, matrix interferences, and chemical interchangeability when analyzed by GC-MS.

A means to differentiate *S. divinorum* from other *Salvia* species would also be necessary for forensic identification of *S. divinorum* if further state or federal regulation were to occur. Salvinorin A is currently only known to exist in *S. divinorum* [7]; therefore, a visual comparison of chromatograms of questioned *Salvia* samples may provide a means of differentiation. In addition, although salvinorin A is not known to occur in any other plants materials, the extract is commercially available and can be easily added to any plant material that is smoked or ingested. In cases where salvinorin A is spiked onto other regulated or non-regulated plant materials, a method to associate extracts of the adulterated material to *S. divinorum* would also be useful.

The publication of the National Research Council report, *"Strengthening Forensic Science in the United States: a Path Forward"*, highlighted the need for statistical evaluation of evidence [34]. As such, placing a statistical confidence on the differentiation or association of evidence is becoming increasingly important in forensic science. A subjective evaluation, such

as visual comparison of chromatograms, might be better replaced with an objective method of differentiation. Chemometric procedures may offer a means for objective differentiation of *S*. *divinorum* from other *Salvia* species, as well as an objective association of adulterated material to *S. divinorum*, but have not yet been investigated for these purposes.

1.5 Research Objectives

The objectives of this research are 1) to optimize an extraction method for salvinorin A from *S. divinorum* based on extraction solvent and extraction time, 2) to use the optimized extraction method to investigate the subjective and objective differentiation of *S. divinorum* from other *Salvia* species, based on the presence or absence of salvinorin A, and 3) to use chemometric procedures to associate extracts of plant materials adulterated with *S. divinorum* to *S. divinorum*.

The first objective of this work then, is to optimize an extraction method for salvinorin A from *S. divinorum* based on extraction solvent and extraction time. It is hypothesized that the most efficient extraction could occur with a solvent of low to medium polarity, as indicated by the low polarity of the salvinorin A chemical structure. Short extraction times are also anticipated to be beneficial in order to decrease breakdown of the cell walls and, hence, reduce the extraction of other undesired components from the plant material. The presence of only salvinorins in the extract would then allow for a uncomplicated GC chromatogram and potentially simpler identification of salvinorin A. Six solvents of varying polarity will be used to extract salvinorin A from dried *S. divinorum* leaves. The extracts will then be analyzed by GC-MS and the optimal solvent chosen on the basis of the extraction efficiency and precision, as well as the stability of the extract over a 24 hour period. Using the optimal solvent, extraction

times from 1-300 minutes will be investigated to determine the extraction time that allows for extraction of the highest yield of salvinorin A.

The second research objective is to investigate the differentiation of *S. divinorum* from other *Salvia* species. When the extracts of the *Salvia* species are analyzed by GC-MS, it is hypothesized that their corresponding chromatograms will be dissimilar and not contain salvinorin A, allowing for visual differentiation. To test this, the optimized extraction procedure will be used for the extraction of *S. divinorum* and four other *Salvia* species (*S. officinalis* (clade I), *S. guaranitica* (clade II), *S. splendens* (clade II) and *S. nemorosa* (clade I)) [6, 8]. The extracts will then be analyzed by GC-MS and the chromatograms visually compared to determine if differentiation is possible. It is also anticipated that chemometric procedures would allow for a non-subjective method of differentiation. Given this hypothesis, PCA, Euclidian distances, HCA, student t-test, Wilcoxon test, and PPMC coefficients, will be performed on the GC-MS data of five *Salvia* species to determine if clustering of the *S. divinorum* extracts are distinguishable from the extracts of the other *Salvia* species.

The final objective of this research is to use PCA to associate extracts of plant materials adulterated with an extract of *S. divinorum* leaves to the *S. divinorum*. It is hypothesized that all samples containing salvinorin A will cluster closely together regardless of the plant matrix, thereby allowing an objective method for the association of adulterated material to *S. divinorum*. To assess this theory, *S. divinorum* leaves will be extracted and the resulting extract will be spiked onto four plant matrices (*S. divinorum, S. officinalis, C. sativa* and *N. tabacum*) to simulate an adulterated plant material that a forensic laboratory might encounter. The adulterated plant matrices will then be extracted using the optimized extraction method and analyzed by GC-MS. PCA, Euclidian distances, student t-test, Wilcoxon test, HCA and PPMC coefficients, will
be performed on the GC-MS data to investigate association of the adulterated material to *S*. *divinorum*.

The knowledge obtained from this work will be useful to forensic analysts currently in the states where *S. divinorum* or salvinorin A are regulated. In addition, in the event of further state or federal regulation, results from this research will be invaluable in developing a nationwide SOP for providing a statistical method for the association and discrimination of adulterated plant materials. The chemometric procedures described herein may also prove to be applicable to other controlled substances.

REFERENCES

References

- 1. Epling, C.; Játiva-M, C. D., A New Species of *Salvia* from Mexico. *Botanical Museum Leaflets* 1962, 20.
- 2. Valdes, L. J.; Diaz, J. L.; Paul, A. G., Ethnopharmacology of Ska-Maria-Pastora (*Salvia divinorum* Epling and Jativa-M). *J. Ethnopharmacol.* 1983, 7, 287-312.
- 3. Siebert, D. J., *Salvia divinorum* and Salvinorin A: New Pharmacologic Findings. *J. Ethnopharmacol.* 1994, 43, 53-56.
- 4. The NSDUH Report. *Use of Specific Hallucinogens: 2006*; Substance Abuse and Mental Health Services Administration: Rockville, MD., February 14, 2008.
- 5. Paulzen, M.; Grunder, G., Toxic Psychosis after Intake of the Hallucinogen Salvinorin A. *J. Clin. Psychiatry*. 2008, 69, 1501-1502.
- 6. Tychonievich, J. Salvia Clades. Personal Communication. 2009.
- 7. Munro, T. The Chemistry of *Salvia divinorum*. Ph.D. Dissertation, 2006, University of Melbourne, Melbourne, Australia, p 283.
- 8. Walker, J.; Sytsma, K.; Treutlein, J.; Wink, M., *Salvia* (Lamiaceae) is not Monophyletic: Implications for the Systematics, Radiation, and Ecological Specializations of *Salvia* and Tribe Mentheae. *Am. J. Bot.* 2004, 91, 1115-1125.
- 9. The Corroboree. Active *Salvia* database. <u>http://www.shaman-australis.com/forum/index.php?showtopic=17236</u> (Accessed March 2010).
- Lee, D.; Ma, Z.; Liu-Chen, L.; Wang, Y.; Chen, Y.; Carlezon, W.; Cohen, B., New Neoclerodane diterpenoids isolated from the leaves of *Salvia divinorum* and their binding affinities for human K opioid receptors. *Bioorg. Med. Chem.* 2005, 13, 5635-5639.
- 11. Prisinzano, T.E.; Tidgewell, K.; Harding, W.W., Kappa opioids as potential treatments for stimulant dependence. *AAPS. J.* 2005, 7, E592-9.
- 12. Hooker, J. M.; Xua, Y.; Schiffer, W.; Shea, C.; Carter, P.; Fowler, J. S., Pharmacokinetics of the potent hallucinogen, salvinorin A in primates parallels the rapid onset and short duration effect in humans. *NeuroImage* 2008, In Press.

- Lee, D.; Ma, Z.; Liu-Chen, L.; Wang, Y.; Chen, Y.; Carlezon, W.; Cohen, B., New Neoclerodane diterpenoids isolated from the leaves of *Salvia divinorum* and their binding affinities for human K opioid receptors. *Bioorg. Med. Chem.* 2005, 13, 5635-5639.
- Gonzalez, D.; Riba, J.; Bouso, J.; Gomez-Jarabo, G.; Barbanoj, M., Pattern of use and subjective effects of *Salvia divinorum* among recreational users. *Drug Alcohol Depen*. 2006, 85, 157-162.
- 15. US Department of Justice, D. E. A. *Drugs and Chemicals of Concern: Salvia divinorum and Salvinorin A.* December 2010.
- 16. House Bill 6038, Ninety-fifth Legislature State of Michigan, October 2010.
- Schmidt, M. S.; Prisinzano, T. E.; Tidgewell, K.; Harding, W.; Kreek, M. J.; Murry, D. J., Determination of Salvinorin A in Body Fluids by High Performance Liquid Chromatography - Atmospheric Pressure Chemical Ionization. *J. Chromatogr. B* 2005, 818, 221-225.
- Ortega, A.; Blount, J. F.; Manchand, P. S., Salvinorin, a New Trans-Neoclerodane Diterpene from *Salvia-Divinorum* (Labiatae). J. Chem. Soc. Perk. T. 1 1982, 2505-2508.
- Valdes, L.; Butler, W.; Hatfield, G.; Paul, A.; Koreeda, M., Divinorin A, a Psychotropic Terpenoid, and Divinorin B from the Hallucinogenic Mexican Mint, *Salvia divinorum. J. Org. Chem.*, 1984, 49, 4716.
- 20. Siebert, D., Localization of Salvinorin A and Related Compounds in Glandular Trichomes of the Psychoactive Sage, *Salvia divinorum*. *Ann. Bot. London*, 2004, 93, 763-771.
- 21. Medana, C.; Massolino, C.; Pazzi, M.; Baiocchi, C., Determination of Salvinorins and Divinatorins in *Salvia divinorum* Leaves by Liquid Chromatography/Multistage Mass Spectrometry. *Rapid Commun. Mass Sp.* 2006, 20, 131-136.
- 22. Giroud, C.; Felber, F.; Augsburger, M.; Horisberger, B.; Rivier, L.; Mangin, P., *Salvia divinorum*: a Hallucinogenic Mint which might become a New Recreational Drug in Switzerland. *Forensic Sci. Int.* 2000, 112, 143-150.
- 23. Jermain, J.; Evans, H., Analyzing *Salvia divinorum* and its Active Ingredient Salvinorin A Utilizing Thin Layer Chromatography and Gas Chromatography/ Mass Spectrometry. *J. Forensic Sci.* 2009, 54, 1-5.

- Tsjikawa, K; Kuwayama, K.; Miyaguchi, H.; Kanamori, T.; Iwata, Y.; Yoshida, T.; Inoue, H., Determination of salvinorin A and salvinorin B in *Salvia divinorum*-related products circulated in Japan, *Forensic Sci. Int.* 2008, 180, 105-109.
- Gruber, J. W.; Siebert, D. J.; Der Marderosian, A. H.; Hock, R. S., High Performance Liquid Chromatographic Quantification of Salvinorin A from Tissues of *Salvia divinorum* Epling & Jativa-M. *Phytochem. Analysis* 1999, 10, 22-25.
- 26. Barnes, S.; Prasain, J. K.; Wang, C.-C.; Moore, D. R., Applications of LC-MS in the Study of the Uptake, Distribution, Metabolism and Excretion of Bioactive Polyphenols from Dietary Supplements. *Life Sci.* 2006, 78, 2054-2059.
- Hupp, A. M.; Marshall, L. J.; Campbell, D. I.; Smith, R. W.; McGuffin, V. L., Chemometric Analysis of Diesel Fuel for Forensic and Environmental Applications. *Anal. Chim. Acta* 2008, 606, 159-171.
- 28. Brereton, R. Applied Chemometrics for Students. England: John Wiley and Sons, 2007.
- Esseiva, P.; Anglada, F.; Dujourdy, L.; Taroni, F.; Margot, P.; Pasquier, E.; Dawson, M.; Roux, C.; Doble, P., Chemical Profiling and Classification of Illicit Heroin by Principal Component Analysis, Calculation of Inter-Sample Correlation and Artificial Neural Networks. *Talanta* 2005, 67, 360-367.
- 30. Klemenc, S., In Common Batch Searching of Illicit Heroin Samples-Evaluation of Data by Chemometric Methods. *Forensic Sci. Int.* 2001, 115, 43-52.
- 31. Skoog, D.; Holler, F.; Nieman, T., Principles of instrumental analysis. 6th ed. Belmont, CA: Thomson, 2007.
- 32. Scientific Working Group for the Analysis of Seized Drugs, Recommendations, January 2011.
- Leger, C.; Lira, H.; Paterson, R., Preparation and properties of surface modified ceramic membranes. Part II. Gas and liquid permeability's of 5 nm alumina membranes modified by a monolayer of bound polydimethylsiloxane (PDMS) silicone oil. *J. Membrane Sci.* 1996, 120, 135-146.
- Strengthening Forensic Science in the United States: A Path Forward. National Research Council of the National Academies. The National Academies Press, Washington D.C. 2009.

CHAPTER 2

THEORY OF INSTRUMENTATION AND CHEMOMETRIC PROCEDURES

2.1 Gas Chromatography

Gas chromatography-mass spectrometry (GC-MS) is the predominate means of analysis for controlled substances identification in forensic chemistry. Separation of sample compounds is performed in the GC, while MS provides a fragmentation pattern that is unique to the sample compounds. In this work, GC-MS was used for the analysis and identification of *S. divinorum* and other plant materials. The following sections give an overview of the GC system, as well as the MS detector, used in this research.

A schematic design of a typical GC system is shown in Figure 2.1. The system generally consists of an injection port into which the sample is introduced, the carrier gas, which acts as the mobile phase, and a column containing the stationary phase. The column is then contained inside an oven.

Separation of a sample mixture in GC is based on the volatility of the chemical compounds in the sample, as well as their distribution coefficients between the mobile phase and the stationary phase. The mobile phase, used to transport the sample compounds through a column containing the stationary phase, is generally an inert gas. Both packed columns and open tubular capillary columns can be used in GC; however capillary columns are generally preferred due to higher efficiency. In gas-liquid chromatography, the stationary phase consists of a viscous liquid (often a silicone-based polymer) immobilized onto an inert porous solid for a packed column or immobilized



Carrier Gas

Figure 2.1. Components of a gas chromatography system

onto the wall for open tubular capillary columns. Separation is achieved through partitioning of the sample between the mobile and liquid stationary phases. In gas-solid chromatography, the stationary phase consists of a high-surface-area solid (inorganic or polymer). Separation occurs through adsorption of the sample onto the stationary phase [1]. Gas-liquid chromatography with an open tubular capillary column was used in this research; therefore the following discussion on typical GC parameters refer to capillary columns.

2.1.1 GC Mobile Phase

The inert gas used as the mobile phase in GC is generally stored in pressurized cylinders and is introduced into the injection port by way of copper tubing and a flow regulator. The carrier gas flows through the column at a set flow rate, generally 1 mL/min. Helium, nitrogen, or hydrogen gas can be used as the carrier gas in GC. Helium is generally preferred, however, as it is non-flammable and comparable to nitrogen and hydrogen in separation efficiency.

2.1.2 Sample Introduction

A schematic design of an injection port is shown in the insert in Figure 2.1. The needle of the syringe, containing the sample, passes through a silicone rubber septum into a heated chamber (typically 250-300 °C) containing a glass inlet liner. Vaporization of the sample occurs in the inlet liner. The carrier gas is introduced into the chamber and the vaporized sample then passes through the inlet liner onto the head of the column. The carrier gas will move either the entire sample (known as splitless injection) or only a portion of the sample (known as split injection) onto the column. A split valve is present for the removal of sample if a split injection is performed. Splitless injections are preferred when dealing with trace amounts of sample compounds, while split injections (generally split ratios of 50:1 or 100:1) are preferred with

concentrated sample compounds to avoid overloading the column which could result in fronting of the peaks.

To be analyzed by GC, a sample must consist of chemical compounds that are volatile and thermally stable within the operating temperatures of the instrument (generally 20 - 300 °C), such as small, non-polar analytes. If the compounds are not volatile they remain behind in the inlet liner and, consequently, contaminate the liner. If thermally labile, they could decompose when exposed to the high injection port temperatures inhibiting the original compound from entering the column.

Generally 1 μ L of sample is introduced into the injection port using either an autosampler coupled to the GC or through manual injection using a syringe. Ideally, the sample should be introduced into the injection port without bias due to volatility or concentration, in a narrow band, and in a reproducible manner. An autosampler generally allows for the highest precision of injection, however, no autosampler was available for this research; therefore, manual injection was used. In manual injection, the type of syringe and method used to inject the sample can substantially affect the volume and composition of the sample entering the column, as well as the precision of the injection. A recent study found that a 1 μ L gas-tight syringe with a two inch spacer allowed for optimal precision of replicate injections using a manual injection procedure [2].

The sample should be injected rapidly in order to create a narrow band of vapor entering the head of the column. A slow injection can cause broadening of the peaks and hence a loss in resolution. Immediately following release of the sample, a partial vaporization of the compounds in the needle of the syringe is induced by the high temperature of the injection port. Compounds from the sample that have low molecular weights will vaporize in the needle first,

with the heavier compounds taking longer. The needle should remain in the injection port long enough for the entire sample to vaporize; otherwise greater concentrations of the more volatile compounds will enter the column. As a result, the sample analyzed will not be truly representative of the composition of the original sample [3]. Vaporization of the remaining compounds occurs in the inlet liner. However, polar compounds can adsorb onto the glass surface of the inlet liner which can cause peak tailing and a loss of sensitivity in the subsequent chromatogram. To prevent adsorption, most liners are chemically deactivated using a surface coating of trimethylchlorosilane or hexamethyldisiloxane.

2.1.3 GC Column

After vaporization of the sample in the injection port, the compounds of the sample are carried through the column in the flow of carrier gas. Compounds with low boiling points (and hence more volatile) will elute from the column before the compounds with higher boiling points. The time of elution, known as the retention time, is the time taken for the compound to travel through the column and reach the detector.

Partitioning of the sample into the stationary phase of the column allows for further separation of the compounds. Columns varying in stationary phase composition are commercially available for gas-liquid chromatography analyses. The optimal stationary phase composition for a given analysis is dependent on the polarity of the sample compounds. A nonpolar polymer (such as polydimethyl siloxane) provides separation based almost exclusively on boiling point or volatility. Slightly polar groups (such as phenyl) can be polymerized with the polydimethyl siloxane to increase the polarity of the stationary phase, while remaining overall non-polar. The addition of the phenyl groups increases the interaction of the aromatic sample components with the stationary phase and, hence, increases their retention time. These stationary

phases are generally coated onto the walls of the column and are often cross-linked in order to prevent thermal degradation, which is known as column bleed.

The flow rate of the carrier gas can also affect the retention time of compounds. If a high flow rate is used the samples will move through the column faster, thereby decreasing the retention time. However, an increased flow rate will decrease the amount of time the compounds can remain in the stationary phase and hence the separation between the compounds in the sample. Therefore, a compromise is necessary between flow rates and time efficient analysis without compromising the separation of the compounds.

2.1.4 GC Oven and Temperature Program

The GC column is located inside an oven to allow strict temperature control during the analysis. In addition to the flow rate, the temperature of the column also affects the retention and separation of the compounds in a sample. The greatest separation of compounds is achieved at low temperatures to reduce volatility and allow maximum interaction of the sample compounds with the stationary phase. However, this results in long elution times that may not be practical for analysis. Therefore, a compromise is sought to allow for the desired separation over a reasonable time period. The oven can be held at a constant temperature (known as isothermal) or ramped in a linear or step-wise manner (known as a temperature gradient). If the sample is a mixture of compounds would not separate. The temperature gradient can be used in these cases to increase the speed of the analysis and still allow sufficient interaction with the stationary phase for separation. Slower ramp rates provide higher resolution, which may be useful for complex samples; however there is increased band broadening. Faster ramp rates result in a more rapid analysis but with decreased resolution.

2.2 Mass Spectrometry

A mass spectrometer (MS) can be coupled to the GC for detection of sample compounds. The MS generally consists of an ionization chamber, a mass analyzer, and a detector as shown in Figure 2.2A. After the sample components are separated in the GC, they pass through a transfer line that is heated (typically 250 - 300 °C) to prevent condensation. As a result, the separated compounds are eluted from the column directly into the ionization chamber and ionized. The ions pass to a mass analyzer where separation occurs based on their mass-to-charge (m/z) ratio. The separated ions then enter a detector where amplification and detection of the signal occurs. The MS components are contained inside a vacuum with pressures of $P = 10^{-4} - 10^{-6}$ torr [1]. The vacuum is necessary to reduce collisions of the ions formed in the ionization chamber with other molecules, which could result in fragmentation or neutralization of the ions. In addition, the vacuum increases the mean free path of the molecules and prevents corrosion of instrument parts. The flow rate from capillary GC (1 to 2 mL/min) is sufficiently low that the MS vacuum system (typically with a pumping speed of 300 L/s or 1.8×10^7 mL/min) is adequate to accommodate the pressure. This allows for direct coupling of the GC column to the MS ionization chamber.



Figure 2.2 A) Components of a mass spectrometry system B) Electron ionization source C) Quadrupole mass analyzer

2.2.1 Ionization Chamber and Ionization

In order for a compound to be detected by MS, it must contain a charge. The first step in MS then is to ionize the compounds separated by the GC, which occurs in the ionization chamber. Many types of ionization methods exist, but the most common in benchtop instruments used in forensic laboratories, as well as the one used in this research, is called electron ionization (EI). A schematic diagram of EI is shown in Figure 2.2B.

As the vaporized sample compounds enter the ionization chamber and react with electrons that have been released from a heated wire filament, generally tungsten. The electrons are accelerated toward an anode by a voltage. The electron beam and sample interaction result in the loss of an electron (e^-) from the sample compound (M), to form the molecular ion (M^{+}) as shown in Equation 2.1.

$$M + e^{-} \rightarrow M^{+} + 2e^{-} \tag{2.1}$$

EI is considered a "hard" ionization technique that results in transferring a large amount of energy to the molecule. This causes extensive fragmentation of the molecular ion, sometimes to the extent that it may not be observed. The extent of fragmentation will depend both on the chemical structure of the sample compound and on the energy of the electron beam. An ionization energy of 70 eV is generally chosen because this allows for sufficient energy transfer to the sample compounds to produce ionization and fragmentation that is useful for structural analysis.

The number of ions (I) produced per unit of time can be given by Equation 2.2.

$$I = NPiV$$
(2.2)

where N is a constant proportionality coefficient, P is the pressure of the sample, i is the electrical current, and V is the volume of sample. A sample compound that is analyzed at 70 eV

electron energy will produce approximately one ion for every 1000 molecules, which is equivalent to 0.1% ionization efficiency [1]. At lower energy fewer ions are produced resulting in a lower sensitivity.

Following ionization, an extraction plate is held at a negative charge in relation to a positive repeller plate, which causes the positive ions formed in the chamber to be accelerated towards additional accelerating plates. These are held at high potentials to attract the ions and focus them into the mass analyzer by means of a focusing slit.

2.2.2 Quadrupole Mass Analyzer

The molecular ions and their fragments exit the ionization chamber, and are accelerated into a mass analyzer, where separation occurs based on their mass and charge. Many different mass analyzers are commercially available; however the quadrupole mass analyzer is most commonly used for routine forensic analyses. The quadrupole consists of four circular parallel rods with an alternating radio frequency (RF) potential and a constant direct current (DC) potential applied to all four of the rods, as shown in Figure 2.2C.

As the RF potential oscillates between the positive and negative phases of its cycle, the ions from the sample compounds will focus and defocus depending on the phase and the charge on the ion. For example, a positively charged ion will be repelled from the rods in the positive phase and attracted to the rods in the negative phase. The ions are separated based on the stability of their trajectories in the electric fields.

Smaller ions are affected more by the RF potential; due to their higher velocities, they focus and defocus faster than larger ions. Controlling the RF potential will act as a high pass filter, allowing high m/z ions to have stable trajectories (resonant ions) and pass through the quadrupole to reach the detector. The low m/z ions will collide with the rods and be neutralized

(non-resonant ions) and, therefore, not be detected. Control of the DC potential will act as a low pass filter, allowing low m/z ions to have stable trajectories and pass through the quadrupole to reach the detector while high m/z ions will collide with the rods and be neutralized. Each m/z ratio has a corresponding RF/DC voltage that allows the ion to successfully pass through the quadrupoles and on to the detector. In order to detect the full mass range (generally m/z 50 - 500) of all the sample compounds, the RF/DC ratio is kept constant and the potential across the quadrupole rods is scanned.

2.2.3 Continuous Dynode Electron Multiplier Detector

After an ion has successfully passed through the quadrupole analyzer, it enters a detector. Many types of detectors are commercially available, but the continuous dynode electron multiplier is the most common. This detector has a curved, horn shaped, continuous dynode that allows for repeated collisions with the surface of the dynode. The surface is coated with a thin film of semi-conducting material (such as aluminum or lead(II) oxide) that readily emits electrons. A potential is applied across the detector, with a high negative potential at the entrance and the exit referenced to ground. The ions that pass through the mass analyzer strike the surface of the dynode, causing emission of secondary electrons. The secondary electrons are then attracted to the less negative potential of the surface and the process is repeated. The total signal amplification is 2^n where n is the number of collisions with the dynode surface, typically resulting in $10^5 - 10^6$ increase in the initial signal [1]. The current is sent to an amplifier and then to a data system to generate the spectrum.

The results of a GC separation of sample compounds are displayed in a graph of the ion abundance versus the retention time of the compounds of the sample. In GC-MS this is known as a total ion chromatogram (TIC). The chromatogram for one m/z can also be compiled and is

known as an extracted ion chromatogram (EIC). When multiple EICs are summed together, this is known as an extracted ion profile (EIP).

The results of the process of ionization, fragmentation, and detection are displayed as a graph of the ion abundance versus the m/z of each ion, known as a mass spectrum. A sample mass spectrum of salvinorin A was shown in Figure 1.2. The molecular ion of salvinorin A is m/z 432 and the ion with the highest abundance (known as the base peak) is m/z 94. Each of the separated sample compounds will give a unique fragmentation pattern, thereby allowing for the definitive identification of each compound.

2.3 Data Pretreatment Procedures

Non-chemical sources of variation can cause differences in background, noise, and retention times between multiple GC analyses. Such variation can result from differences in injection volume (manual injection is especially susceptible), flow, temperature, or degradation of the column with time. Data pretreatment steps are necessary to remove these non-chemical sources of variance prior to performing chemometric procedures, such as principal components analysis (PCA), so that any variance is only due to chemical differences in the samples. Depending on the type of data and the chemometric procedure being performed, different pretreatment steps may be necessary. Specifically for GC data, procedures such as background subtraction, peak smoothing, retention time alignment, and normalization are commonly used [4].

2.3.1 Background Subtraction

The background in a chromatogram can vary between analyses due to impurities from the mobile and stationary phases (i.e. column bleed). Impurities in the carrier gas, such as hydrocarbons, can cause increased noise and background current in the chromatogram not due to

the chemical compounds in the sample. In addition, when a high temperature program is used, the liquid coating from the stationary phase can undergo degradation, resulting in a rise in the baseline of the chromatogram. A representative rise in the baseline of a chromatogram is shown in Figure 2.3A. Background subtraction is performed to eliminate the variations due to these impurities. The method used in this research for background subtraction involves the generation of an EIC that corresponds to common ions observed from column degradation [2]. The EIC can be acquired for the ions corresponding to a mobile phase impurity or the stationary phase degradation, by plotting the signal of a single m/z over time. Several common EICs correlating to siloxane compounds from stationary phase degradation are at m/z 73, 193, 207, 221, 341, and 355. These EICs can then be summed together to create an extracted ion profile (EIP) that accounts for the majority of the background signal, shown in Figure 2.3B.

The EIP is then fit by nonlinear regression to an appropriate equation. For example, the rise in the baseline observed at higher temperatures is generally representative of an asymmetric sigmoid equation:

$$y = a + \frac{b}{\left(1 + \exp\left(-\frac{x - d\ln(2^{1/e} - 1) - c)}{d}\right)\right)^{e}}$$
(2.3)

where a and b are the mean abundances at the beginning and end of the chromatogram, c is the retention time at the point of inflection, and d and e are dependent parameters that define the shapes of the corresponding two regions of the curve, as shown in Figure 2.3C. This equation can be regenerated and subtracted from the TIC to eliminate or minimize background sources of non-chemical variance in the chromatogram. An example of a background subtracted TIC is shown in Figure 2.3D.



Figure 2.3. A) Original TIC B) Extracted ion profile of m/z 73, 193, 207, 221, 341, and 355 showing rise in the background, C) Asymmetric sigmoid equation, D) TIC after background subtraction, E) TIC after FFT smoothing

2.3.2 Smoothing

Variations in the detector (i.e. continuous dynode electron multiplier for MS) can create non-chemical sources of perturbation in the chromatogram, known as noise. Smoothing is performed to minimize this noise in the chromatogram and many types of smoothing algorithms exist, such as Savitzky-Golay, adjacent averaging, percentile filter, and fast Fourier transform filter [1]. In this research, a fast Fourier transform (FFT) smoothing algorithm was applied to the chromatograms. In Fourier transformations, signals that have been collected in the time-domain (such as a chromatogram) are converted to a frequency-domain signal. The frequency domain signal is then multiplied by a digital low-pass filter function [1]. A FFT smooth removes compounds with frequencies higher than a specified frequency (F_{cutoff}), which is determined by

$$F_{\text{cutoff}} = \frac{1}{\text{nDt}}$$
(2.3)

where n is the number of data points and Δt is the time spacing between adjacent data points [1]. Larger values of n correspond to a lower cutoff frequency, resulting in a higher degree of smoothing, while smaller values of n correspond to higher cutoff frequency and less smoothing. An inverse Fourier transform then returns the filtered frequency domain data back to the time domain.

Visual inspection of the degree of smoothing is necessary regardless of the type of algorithm, as over-smoothing can cause a significant loss in resolution. However, an appropriate level of smoothing can be beneficial to remove non-chemical sources of variance as well as to improve the signal-to-noise ratio and shapes of the peaks in the chromatogram, as shown in Figure 2.3E.

2.3.3 Retention Time Alignment

Small differences in carrier gas flow rate or temperature can cause the same analyte to elute at slightly different retention times, thereby introducing a non-chemical source of variation. Retention time alignment is performed to account for these drifts in retention time between GC analyses [4]. A peak matching alignment algorithm was used to align the chromatograms in this research [5].

In order to align multiple chromatograms, a target must be chosen. The target chromatogram used in this research was compiled from the mean abundance at each retention time in the chromatograms under investigation. The alignment algorithm first attempts to identify the peaks in the target chromatogram. To do so, the first derivative, or difference between each pair of points in the target chromatogram, is calculated. A threshold value is set by the user and is generally defined as five times the standard deviation of the baseline noise. The leading edge of a peak is identified when the difference in abundance between two points exceeds the threshold value. The alignment algorithm then calculates the zero crossing through interpolation and this is defined as the peak maximum; the next zero crossing is then the tailing edge of the peak (Figure 2.4A). The identified peaks in the target chromatogram.

Peaks in the sample chromatograms are then compared to the peaks in the target chromatogram (Figure 2.5A). A window size is set by the user to define the number of data points that the algorithm can use to match peaks (Figure 2.5B). If the peak in a sample chromatogram is within the window size (e.g. five points) of the peak in the target chromatogram, then that peak is considered a match, and the time points on either side of the zero crossing of the first derivative are aligned through interpolation (Figure 2.5C). The process



Figure 2.4. First derivative zero crossing



Figure 2.5. A) Comparison of sample to target, B) Peak matching window, C) Alignment of sample to target

is repeated for each peak in each of the sample chromatograms to align them to the target chromatogram.

Several factors can affect the quality of the alignment. For instance, if a large window size is used, peaks in the sample chromatograms that have similar retention time but are not due to the same compound may be aligned together and may misrepresent the data. However, if a small window size is used, normal drifts in the retention time may not be accommodated and peaks corresponding to the same compound may not be aligned. In addition, if a peak in either the target or sample chromatograms is below the threshold value, it will not be identified and, consequently, will not be aligned. The mass spectrometer scan rate (typically set at 2.91 seconds) also influences the alignment. A fast scan rate leads to a longer time between points, thereby increasing the possibility of error in the interpolation and alignment of the zero-crossing of the peaks. Optimization of the alignment parameters can result in a powerful tool with the ability to decrease the contribution of non-chemical variance in the data set.

2.3.4 Normalization

Normalization is performed to minimize non-chemical variations in the abundances of chromatographic data. These variations can be caused, for example, by differences in injection volume and/ or instrument sensitivity between analyses. Several types of normalization procedures, such as total area, peak maximum, or single peak normalization, can be used depending on the type of data. Maximum peak normalization was used in this research due to large differences in the amount of material extracted from different plant species. Using this normalization procedure, each variable (retention time) in the chromatogram is scaled relative to the maximum abundance in that chromatogram. The largest peaks in the data set then have an

equal abundance, hence minimizing variation due to differences in peak abundances among samples.

2.4 Chemometric Procedures

2.4.1 Principal Components Analysis

Principal components analysis (PCA) is a widely used multivariate statistical procedure that reduces a complex data set (such as GC-MS data) to a few principal components that represent the greatest contributions to variance among the samples [6]. PCA highlights the relationships among samples that may otherwise be difficult to observe due to the complexity of the data. For GC data, each retention time is a variable, meaning that each data point in a chromatogram is a dimension. An advantage of PCA is the ability to reduce dimensionality, with no loss of chemical information, such that underlying patterns in the data can be observed.

Principal components are linear combinations of the original variables, but are uncorrelated. The data are represented in n-dimensional space, where n is the number of variables. New latent axes are positioned to maximize variance in the data set. Principal component 1 (PC1) accounts for the greatest variance, and subsequent PCs are positioned orthogonally and account for the next greatest variance. For n variables or dimensions, n PCs are determined.

For PCA to produce chemically meaningful results, variance must arise only from chemical differences in the samples. Therefore, the data pretreatment steps described previously are necessary to remove non-chemical sources of variance prior to performing PCA. Following data pretreatment, the covariance of data points in the chromatograms, measuring the deviation from the mean in each dimension, is calculated as

$$cov(x,y) = \frac{\sum_{i=1}^{n} (x_i - \mu_x) (y_i - \mu_y)}{n - 1}$$
(2.4)

where x_i and y_i are data points and μ_x and μ_y are respective means in the x and y dimensions of n number of dimensions (or variables) in the data set [6]. Note that the data are mean-centered during the process of calculating the covariance. A covariance matrix can then be assembled. For example, a data set with *x*, *y*, and *z* dimensions would result in the following covariance matrix:

$$\begin{bmatrix} \operatorname{cov}(x, x) & \operatorname{cov}(x, y) & \operatorname{cov}(x, z) \\ \operatorname{cov}(y, x) & \operatorname{cov}(y, y) & \operatorname{cov}(y, z) \\ \operatorname{cov}(z, x) & \operatorname{cov}(z, y) & \operatorname{cov}(z, z) \end{bmatrix}$$

The covariance matrix shown above, is symmetrical about the main diagonal as cov(x, y) = cov(y, x). If the data set to which PCA is applied is not square, additional zeros are added until a square matrix can be constructed.

The eigenvectors of the covariance matrix for the total data set are calculated next and these are the principal components. An eigenvector is a unit vector (v) of a matrix (A) for which

$$Av = cv \tag{2.5}$$

is true, where c is a scalar [6]. Therefore, when the eigenvector is multiplied by the data matrix, a multiple of the original vector is obtained. The scalar c is the multiple of the original vector and is referred to as the eigenvalue of the matrix. For n dimensions, there are n eigenvectors that describe the variance in the data set. Eigenvectors correspond to PCs and the eigenvalue corresponds to the variance in the data set described by that PC.

The two most common outputs from PCA are scores and loadings plots. A score is obtained by multiplying the eigenvector (or PC) by the mean-centered data to give a single number for each sample. A scores plot can then be constructed by plotting the scores for the sample for PC1against PC2 (or any other combinations of PCs). A graph of PC1 versus PC2 will

show the majority of the variance in the data set, PC2 versus PC3 the next greatest amount of variance, and so on. Samples that are chemically similar will have similar scores and, therefore, will be clustered together on the PCA scores plot. In contrast, chemically different samples will have dissimilar scores and will not be clustered.

An individual score cannot provide direct comparison on its own and, is therefore, only meaningful relative to other scores. Statistical testing can, however, be used to compare the association or discrimination of scores, and hence, of the samples they represent. Further discussion of statistical testing for the evaluation of PCA data is presented in 2.4.2 -2.4.7.

In addition, the PCA loadings plots can be useful for identifying the chemical compounds that contribute most to the variance among the samples. For chromatographic data, loadings plots are constructed by plotting the PC of interest against retention time. The chemical compounds positioned the furthest from zero on the ordinate of the loadings plots are responsible for the most variance in the data set. Based on retention time (abscissa of the loadings plot), the chemical identity of such compounds can be determined. Hence, the positioning of samples in the scores plot can be explained based on the chemical compounds that contribute most to the variance described by each PC. Both PCA scores and loadings plots are powerful chemometric tools to assess association and discrimination among samples in complex data sets.

2.4.2 Pearson Product Moment Correlation Coefficients

Pearson product moment correlation (PPMC) coefficients are a statistical measure of the association, or similarity, between variables. In this research, PPMC coefficients derived from PCA loadings were used to assess similarity between pairs of chromatograms. These PPMC coefficients (r_{AB}) were calculated by

$$r_{AB} = \frac{\sum_{t=1}^{n} \left(\left(PC \times A_{t} \right) - \mu_{A} \right) \left(\left(PC \times B_{t} \right) - \mu_{B} \right)}{\sqrt{\sum_{t=1}^{n} \left(\left(PC \times A_{t} \right) - \mu_{A} \right)^{2} \left(\left(PC \times B_{t} \right) - \mu_{B} \right)^{2}}}$$
(2.6)

where PC corresponds to the loadings responsible for the greatest differentiation of the samples, A_t and B_t correspond to abundances at retention time t in two chromatograms, A and B, with means of μ_A and μ_B .

A PPMC coefficient of +1 indicates a positive linear relationship, whereas -1 indicates a negative linear relationship between two loadings plots. A coefficient of 0.80 or greater indicates high correlation, 0.50 - 0.79 indicates moderate correlation, 0.49 or less indicates little correlation, and coefficients close to zero indicate no correlation [6].

Theoretically, the PPMC coefficient between replicates should be close to one. Deviation from one is a measure of the lack of precision of the analysis, *i.e.* greater deviation indicates less precision. Conversely, loadings plots from different samples, such as different plant materials, should have lower PPMC values. PPMC coefficients derived from PCA loadings allow a comparison of samples based on the variance the associated PC is describing. In this manner, PPMC coefficients can be used as a measure of the similarity or dissimilarity of samples.

2.4.3 Euclidian Distance

The Euclidian distance is the numerical distance between the means of two populations. In this research, Euclidian distances were calculated to provide a statistical measure of the association and discrimination of samples in the PCA scores plot. The Euclidian distance between two samples (A, B) is calculated by

$$d_{AB} = \sqrt{\left(PC1_{A} - PC1_{B}\right)^{2} + \left(PC2_{A} - PC2_{B}\right)^{2}}$$
(2.7)

where $PC1_A$ and $PC2_A$ are the mean scores of sample A on PC 1 and 2, respectively, and $PC1_B$ and $PC2_B$ are the mean scores of sample B on each PC. Samples that are close to one another on the PCA scores plot have short Euclidian distances, whereas samples that are distinct have longer Euclidian distances. In this manner, Euclidian distances are a numerical value used to describe the visual association or discrimination seen in a PCA scores plot. Similar to scores, Euclidian distances, are meaningful only when compared to other distances between samples in a data set. 2.4.4 Student T-test

A Student t-test is used to statistically determine if two samples are from the same population. If the two samples can be distinguished from each other a null hypothesis, H_0 , is stated

$$H_{O}:\left|\mu_{A}-\mu_{B}\right|\neq0$$
(2.8)

where the means of the two sample sets, μ_A and μ_B respectively, are not the same. If the two samples cannot be distinguished from each other, then the alternative hypothesis, H_a , is stated

$$\mathbf{H}_{\mathbf{a}} : \left| \boldsymbol{\mu}_{\mathbf{A}} - \boldsymbol{\mu}_{\mathbf{B}} \right| = 0 \tag{2.9}$$

where the means of the two sample sets, μ_A and μ_B respectively, are the same.

Two types of errors can be made in hypothesis testing, type I and type II errors, respectively known as false positive and false negative results in forensic science. Type I errors arise if H_0 is rejected when it is true and type II errors arise if H_0 is not rejected when it is false. The probabilities of these errors occurring are denoted as α and β , respectively. Although both types of errors should ideally be minimized, generally a fixed α is chosen and β is minimized as much as possible [8].

To determine which hypothesis is verified, H_0 or H_a , a t-test can be used. The manner in which the t-statistic is calculated depends on the homoscedasticity of the data; therefore, an F-test is initially used to test the variances of the two populations and is calculated by

$$F_{\text{calc}} = \frac{\sigma_{\text{A}}^2}{\sigma_{\text{B}}^2}$$
(2.10)

where σ_A is the standard deviation of population A and is the larger variation and σ_B is the standard deviation of population B. The degrees of freedom, df, for each population are calculated by

$$df = n-1$$
 (2.11)

where n is the number of samples in the population. The F_{calc} with the given df is then compared to a table of F critical values, F_{crit} , at a given confidence level. When F_{calc} is greater than F_{crit} , the variances are statistically equivalent. If, however, F_{crit} is greater than or equal to F_{calc} , then the variances are not statistically equivalent.

If the variances are determined to be statistically equivalent then a pooled standard deviation, σ_{pooled} , is calculated

$$\sigma_{\text{pooled}} = \sqrt{\frac{\sigma_{A}^{2} \left(n_{A}^{-1}\right) + \sigma_{B}^{2} \left(n_{B}^{-1}\right)}{n_{A}^{+n} B^{-2}}}$$
(2.12)

where n_A and n_B are the number of samples in population A and B, respectively. The σ_{pooled} is then used to calculate the equal variance t-statistic

$$t_{calc} = \frac{\left(\mu_{A} - \mu_{B}\right)}{\sigma_{pooled}\sqrt{\frac{1}{n_{A}} - \frac{1}{n_{B}}}}$$
(2.13)

where μ_A and μ_B are the means in population A and B, respectively.

However, if the variances are not equivalent, the unequal variance t-test, or Welch t-test, is performed, which is calculated as

$$t_{calc} = \frac{\left| \mu_A - \mu_B \right|}{\sqrt{\frac{\sigma_A^2}{n_A} - \frac{\sigma_B^2}{n_B}}}$$
(2.14)

An approximation of the degrees of freedom, df, is calculated by

$$df = \frac{\left(\frac{\sigma_A^2}{n_A} - \frac{\sigma_B^2}{n_B}\right)}{\frac{1}{n_A^{-1}} \left(\frac{\sigma_A^2}{n_A}\right)^2 + \frac{1}{n_B^{-1}} \left(\frac{\sigma_B^2}{n_B}\right)^2}$$
(2.15)

This equation generally leads to a degree of freedom that is not an integer and, therefore, to be conservative, should be rounded down to the nearest integer.

To determine if H_0 is true at a given confidence level, the value of t_{calc} is compared to a table of critical values, t_{crit} , at the desired level of statistical significance. When t_{calc} is less than or equal to t_{crit} , the H_0 is verified and the scores being compared are considered statistically

distinct. Alternatively, H_a is verified when t_{calc} is greater than t_{crit} and the scored being compared are considered statistically associated. In this manner, a confidence level is given to the association or discrimination of the sample scores on a PCA scores plot.

2.4.5 Wilcoxon Rank-Sum Test

The Wilcoxon Rank-Sum Test is a non-parametric test that is used to statistically determine if two samples are from the same population. If the two cannot be statistically distinguished, a null hypothesis, H_0 , is stated according to Equation 2.8, whereas if the two populations can be distinguished from each other, then the alternative hypothesis, H_a , is stated according to Equation 2.9. In order to determine which hypothesis is verified, the Wilcoxon test statistic, W, is calculated and used to determine the Z_{calc} at the desired level of statistical significance.

For two populations, A and B, containing n_A and n_B number of observations, respectively, the independent observations are merged and then ranked from lowest to highest value. For example, with PCA data, observations in populations A and B would refer to the scores of a sample on PC1 and PC2. If two observations have the same numerical value, a mean rank is used for both. The Wilcoxon test statistic, W, is calculated by

$$W = R - \frac{n(n+1)}{2}$$
(2.16)

where R is the sum of the ranks from either of the populations and n is the number of observations in that respective population. The mean, μ_W , and standard deviation, σ_W , of W can then be calculated by

$$\mu_{\rm W} = \frac{n_{\rm A} \left(n_{\rm A} + n_{\rm B} + 1 \right)}{2} \tag{2.17}$$

$$\sigma_{\rm W} = \sqrt{\frac{n_{\rm A} n_{\rm B} (n_{\rm A} + n_{\rm B} + 1)}{12}} \tag{2.18}$$

The W-statistic can be calculated for either population and used to find the Z-statistic, Z_{calc} by

$$Z_{calc} = \frac{\left| W - \mu_{W} \right|}{\sigma_{W}}$$
(2.19)

The value of Z_{calc} is the same regardless of which population is used to calculate W.

To determine if H_0 is true at a given confidence level, the value of Z_{calc} is compared to a table of critical values, Z_{crit} , at the desired level of statistical significance. When the absolute value of Z_{calc} is less than or equal to Z_{crit} , the H_0 is verified and the scored being compared are considered statistically distinct. Alternatively, H_a is verified when Z_{calc} is greater than Z_{crit} and the scored being compared are considered statistically associated. When H_0 is validated, the ranks of the observations should be interspersed between the two populations, while the ranks of the observations should be distinct between the two populations if H_a is validated. In this manner, a confidence level is given to the association or discrimination of the sample scores on a PCA scores plot.

2.4.6 Comparison of Student T-test and Wilcoxon Test

Both the student t-test and the Wilcoxon rank-sum test can be used to statistically determine if two samples are from the same population. The student t-test is recognized as the best test for minimizing β for a fixed α [8]. However, the student t-test assumes the data are

normally distributed with uniform variance, which may not be true of all data [8]. The Wilcoxon rank-sum test is a non-parametric test that does not make the assumption that data are normally distributed.

The asymptotic relative efficiency (ARE) is a means of calculating the effectiveness of statistical tests and has been used to compare the student t-test and the Wilcoxon rank-sum test [8]. For large sample sets (n >100) with normal distribution, the Wilcoxon test is approximately 95% as efficient as the t-test [8]. For any distribution, the Wilcoxon test is at least 86% as efficient as the t-test and, in cases where the distribution is not normal, is more efficient than the t-test [8]. Therefore, the Wilcoxon test is considered to be relatively efficient in comparison to the t-test for normal distribution and, in cases where the distribution is not normal, more efficient than the t-test.

2.4.7 Hierarchical Cluster Analysis

Hierarchical cluster analysis (HCA) allows graphical representation of the association among samples in a multivariate data set, based on the similarity of the data. Two general types of clustering are used in HCA, agglomerative (each sample is defined as its own cluster initially and then grouped until a single cluster is formed) and divisive (all samples are considered as a single cluster initially and then divided until each sample is its own cluster). In this work, agglomerative clustering is used. Distances between samples are calculated and compared to determine which samples are closest to one another, thus forming the respective clusters. There are a variety of methods to calculate both the distances between samples as well as the linkage methods. In this work, a Euclidian distance with nearest-neighbor linkage is used and is discussed below.

In agglomerative HCA, the Euclidian distance, d_{AB}, between a pair of sample scores,

 PC_{A} and $PC_{B},$ can be calculated using a univariate version of Equation 2.7

$$d_{AB} = \sqrt{\left(PC_A - PC_B\right)^2}$$
(2.20)

The similarity, SAB, between the two samples is then calculated according to

$$S_{AB} = 1 - \frac{d_{AB}}{d_{max}}$$
(2.21)

where d_{max} is the distance between the two farthest samples in the data set. In this manner, the samples that are farthest from one another are assigned a similarity of zero, while the samples that are identical are assigned a similarity value of one.

For the nearest-neighbor method linkage, the distance is calculated as the distance between the two closest samples in the two clusters. Distances are calculated between all pairs of samples initially. The two closest (most similar) samples are linked to form a cluster (Figure 2.6 A1). The samples or cluster with the next shortest distance are then linked to either add to an existing cluster (Figure 2.6 A2 and 3) or form a new cluster (Figure 2.6 A4). This process continues until all the samples form a single cluster (Figure 2.6 A5). A tree branch graph, called a dendrogram, is used to display the relationships between the samples, based on their calculated similarities (Figure 2.6 B). The visual representation of association between samples can be useful in demonstrating the similarities between scores generated by PCA and can be complimentary in confirming the association seen in the PCA scores plot.



B



Figure 2.6. A) Nearest-neighbor linkage, B) HCA dendrogram with highest similarity proportional to branch length
REFERENCES

References

- 1. Skoog, D.; Holler, F.; Nieman, T., Principles of instrumental analysis. 6th ed. Belmont, CA: Thomson, 2007.
- 2. McIlroy, J., Injection study. Personal Communication. 2009.
- 3. Hinshaw, J., Syringes for gas-chromatography. *LC GC Eur.*, 2006, 19.
- 4. Hupp, A. M.; Marshall, L. J.; Campbell, D. I.; Smith, R. W.; McGuffin, V. L., Chemometric analysis of diesel fuel for forensic and environmental applications. *Anal. Chim. Acta* 2008, 606, 159-171.
- 5. Johnson K.; Wright B.; Jarman K.; Synovec R.; High-speed peak matching algorithm for retention time alignment of gas chromatographic data for chemometric analysis. *J. Chromatogr. A* 2003, 996,141-55.
- 6. Brereton, R., Applied chemometrics for students. England: John Wiley and Sons, 2007.
- 7. Smith, L., A tutorial on principal component analysis. University of Otago, New Zealand, 2002.
- 8. Devore, J., Probability and statistics for engineering and the sciences. Duxbury Press, Belmont, CA, 1990.

CHAPTER 3

DEVELOPMENT OF AN OPTIMIZED PROCEDURE FOR THE EXTRACTION OF SALVINORIN A FROM S. DIVINORUM

3.1 Introduction

An increasing number of individual states have regulated either *S. divinorum* or salvinorin A and several others have pending legislation [1, 2]. In the event of further state regulation and the potential of federal regulation, a fast, reproducible, and simple procedure for extraction of salvinorin A from *S. divinorum* will be necessary for development of a standard operating procedure (SOP). In forensic laboratories in states where *S. divinorum* or its active component are regulated, extraction methods for salvinorin A are widely varied [3]. Currently, no systematic comparison of extraction solvents or times has been performed, which will make developing a SOP for the extraction of salvinorin A from *S. divinorum* A from *S. divinorum* difficult.

The goal of this work is to develop a procedure for the optimal extraction of salvinorin A from *S. divinorum*. To be practical in a forensic science laboratory, the procedure must also be time-efficient. Both extraction solvents and times will be systematically investigated. Six solvents of varying polarity (methanol, acetonitrile, acetone, chloroform, dichloromethane, and hexane) will be used to extract salvinorin A from the dried *S. divinorum* leaves. Solvents with a wide range of polarities are desirable to determine the solvent that extracts the greatest mass of salvinorin A. The most efficient extraction is likely to occur with a solvent of low to medium polarity, due to the polarity of salvinorin A, as demonstrated by the chemical structure in Chapter 1 (1). The extracts will then be analyzed by gas chromatography-mass spectrometry (GC-MS)

and the optimal solvent will be chosen based on the salvinorin A mass extracted, the complexity of the GC chromatogram, and the interday stability of the extracts.

Previous research by Seibert has shown that the majority of salvinorin A is located in trichomes on the leaf surface of the *S. divinorum* plant [4]. This suggests that the most efficient extraction of salvinorin A will occur with short extraction times that are insufficient for breakdown of the cell walls, therefore minimizing the extraction of other plant materials. The presence of only salvinorins in the extract will result in an uncomplicated GC chromatogram, therefore making identification of salvinorin A simpler.

The optimal solvent will then be used to investigate six extraction times (1, 3, 6, 30, 100, 300 minutes). The extracts will be analyzed by GC-MS and the optimal extraction time will be chosen based on the mass of salvinorin A extracted. Finally, the optimized procedure will be used to determine the extraction efficiency of salvinorin A from *S. divinorum*.

3.2 Materials and Methods

3.2.1 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 (Honeywell Burdick and Jackson, 99.9% purity, Morristown, NJ). *S. divinorum* leaf material was extracted in each solvent for 1 minute. 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 optimal solvent was determined based on the mass of salvinorin A extracted. Using the optimized extraction solvent, extraction times of 1, 3, 6, 30, 100, and 300 minutes were then investigated in a similar manner.

3.2.2 Extraction Efficiency

Reference standards containing 0.1, 0.5, 2, 5, 7.5, 10, and 12 mg/mL of 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 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 [5], 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 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 optimized extraction method, evaporated to dryness, and reconstituted in 0.4 mL of 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.

3.2.3 GC-MS Analysis

All extracts were analyzed by gas chromatography (Agilent Technologies, model 6890N, Santa Clara, CA) equipped with a 5%-phenyl/95%-methylpolysiloxane stationary phase column (DB-5MS, 30 m length x 0.25 mm inner diameter x 0.25 µm film thickness, Agilent Technologies, Palo Alto, CA). Polydimethyl siloxane is a non-polar polymer that is coated to the walls of the column (it is often cross-linked in order to prevent degradation known as column bleed). Phenyl groups are polymerized with the polydimethyl siloxane to increase the polarity. The addition of the aromatic groups increases the interaction of the non-polar components with the stationary phase and, hence, increases the retention time of the non-polar components.

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 μ L of sample extract was injected in split mode (50:1). An injection method was developed for the analysis of the extracts to ensure a reproducible injection. After 1 μ L of the sample was drawn up into a 1 μ L gas-tight syringe (Hamilton Company, Reno, NV) with a two inch spacer, the needle was placed in the GC injection port, and held for 2 seconds prior to depressing the plunger. After the sample was injected the needle remained in the port for another 2 seconds before removal. This allowed the entire sample to vaporize prior to removal of the syringe.

The oven temperature program was as follows: 80 °C for 2 minutes, 10 °C/minute to 340 °C, with a final hold at 340 °C for 4 minutes. 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.

3.3 Results and Discussion

3.3.1 Preliminary Studies

In preliminary work, salvinorin A was extracted from *S. divinorum* by Soxhlet extraction for 24 hours at the boiling point of the four solvents investigated: methanol (64.5 °C), acetone (56.5 °C), dichloromethane (39.8 °C), and hexane (68.0 °C) [6-9]. Each extract was analyzed by GC-MS; however, no salvinorin A was observed in the resulting chromatograms. Given the stability of salvinorin A in solution, it was determined that degradation had occurred at the higher temperatures used for the extraction.

Rotary agitation extraction with the same four solvents was also performed for 24 hours at ambient temperature and the extracts analyzed by GC-MS. The recoveries of *S. divinorum* in milligrams per gram of dried leaf material for each of the solvents are shown in Table 3.1. Methanol extracted the greatest mass of *S. divinorum* followed by dichloromethane, acetone, and hexane. Nevertheless, the abundance of the salvinorin A peak was highest in the dichloromethane extracts, followed by methanol and acetone. The total ion chromatogram of a dichloromethane extract of *S. divinorum* is shown in Figure 3.1. Although salvinorin A was also present in the hexane extracts, the peak was below the background signal. As can be seen in Figure 3.1, other components (various straight and branched alkanes and alkenes) from the plant material were also extracted using the rotary agitation procedure. Table 3.1 Total mass of S. divinorum extracted using rotary agitation in four solvents

Total Mass S. divinorum
Extracted (mg/g)*
86.3
32.1
44.4
13.0

* mg of material per g of leaf extracted



Time (min)

Figure 3.1. Extract of *S. divinorum* in dichloromethane using rotary agitation. 1) salvinorin B, 2) salvinorin A, 3) salvinorin C, 4) salvinorin D.

The Soxhlet extraction indicated that increased temperatures for extraction of salvinorin A from *S. divinorum* could potentially result in degradation of salvinorin A. The rotary agitation extraction demonstrated that salvinorin A was detected using ambient temperatures. However, given the long extraction times for the rotary agitation, other plant components were also observed in the chromatograms. This indicated that shorter extraction times may be beneficial for the simplicity of identification as well as prove more reasonable for use in forensic laboratories. Based on these preliminary findings, shorter extraction times at ambient temperatures were subsequently investigated.

3.3.2 Optimization of Extraction Solvent

Six solvents were selected with varying polarity (methanol, acetonitrile, acetone, chloroform, dichloromethane, and hexane) and the corresponding solubility parameters are reported in Table 3.2. Representative chromatograms of *S. divinorum* in each of the solvents are shown in Figure 3.2. Methanol, acetone, and chloroform are currently used for extraction of salvinorin A from *S. divinorum* in forensic laboratories [3], whereas acetonitrile has also been reported [10]. 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 minute. The results of the solvent study are summarized in Table 3.2.

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

Solvent	Solubility Parameter ¹¹ (MPa ^{1/2})	Average Total Mass Extracted (mg/g) ^a	Average Mass Salvinorin A Extracted (mg/g) ^a	Interday RSD ^b		
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 ^c	ND ^c		
^a mg of material per g of leaf extracted, mean \pm standard deviation (n = 9)						

Table 3.2. Mass of material extracted from *S. divinorum* with a one-minute extraction time and different solvents.

^b RSD = relative standard deviation, mean \pm one standard deviation (n = 6)

^c ND = not detected



Figure 3.2. Representative total ion chromatograms of *S. divinorum* extractions in (A) methanol, (B) acetonitrile, (C) acetone, (D) chloroform, (E) dichloromethane, (F) hexane. IS = internal standard, 1) salvinorin B, 2) salvinorin A, 3) salvinorin C, 4) salvinorin D.

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 [12]. 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 either of these conditions. Hexane was therefore determined to be an 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 hours [10]. 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 hours. The extracts were analyzed in triplicate on two consecutive days and the results of the interday stability study are summarized in Table 3.2. A decrease in salvinorin A abundance over the two-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 salvinorin A. Consequently, dichloromethane was chosen as the optimal extraction solvent, as it extracted the greatest mass of salvinorin A, with the least extraction of other plant compounds, and the highest interday precision.

Dichloromethane may be carcinogenic and is hazardous when inhaled [8]. Both chloroform and acetone could be considered as alternatives, as they extract the next greatest mass of salvinorin A (Table 3.2). However, chloroform is more toxic than dichloromethane and has some acidic character that can degrade a GC column. Nevertheless, chloroform is commonly

used in many forensic laboratories as a solvent for the extraction of controlled substances. Acetone is the least toxic of these three solvents; however moisture from the air can result in condensation reactions, creating impurities in the solvent. Additionally, acetone has a low flash point (-20 °C) and is, therefore, highly flammable [7]. Acetone also had the highest interday variability, meaning that the acetone extract must be analyzed immediately or the solvent evaporated prior to storage. Thus, for the purpose of this research, dichloromethane was used as the extraction solvent for salvinorin A.

3.3.3 *Optimization of Extraction Time*

The extraction time was then optimized using dichloromethane as the solvent. Extraction times of 1, 3, 6, 30, 100, and 300 minutes were investigated. Since salvinorin A is found primarily in the trichomes on the *S. divinorum* leaves [4], shorter extraction times may be favorable to reduce breakdown of cell walls or extraction of other plant materials, such as plant pigments. Therefore, a one-minute extraction was chosen to represent the shortest extraction time that can be reasonably reproduced. The 300-minute 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 Figure 3.3.

The mass of salvinorin A extracted increased with exposure time to dichloromethane, reaching a maximum between 3 and 6 minutes $(0.0796 \pm 0.0153 \text{ and } 0.0882 \pm 0.0202 \text{ mg/g})$, then remained relatively constant from 30 to 300 minutes (Figure 3.3). There was no statistically significant difference in the mass of salvinorin A extracted using 3 and 6 minute extraction times. Therefore, an extraction time of 5 minutes was chosen for convenience. These results

confirm that shorter extraction times result in higher abundance of salvinorin A [4]. However, the standard deviations for triplicate extractions are generally higher for the shorter extraction times. Variation inherent in a short manual extraction could have a greater impact on the total variance than the longer extraction times, resulting in higher standard deviations. However, shorter extraction times are more practical in forensic laboratories. The decrease in abundance of salvinorin A extracted from 6 minutes to 30 minutes may have been due to the instability of salvinorin A in solution, as demonstrated by Tsujikawa *et al.* [10]. 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-minute extraction time is shown in Figure 3.4. Although salvinorins B, C, and D were also observed, salvinorin A was the most abundant compound extracted, resulting in a simple chromatogram that allows easy identification of salvinorin A (retention time = 26.62 min, molecular ion = m/z 432).

3.3.4 Linear Range and Extraction Efficiency of Salvinorin A from S. divinorum

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* [5]. An extraction efficiency of 97.6% was determined with 1.95 ± 0.20 mg/mL recovered from the 2.00 mg/mL spike. The calibration curve is shown in Figure 3.5. It was noteworthy that the analytical method was linear (y = 4 x 10^7 x - 5 x 10^6 , R² = 0.9951) over the concentration range investigated of 0.10 - 12.00 mg/mL salvinorin A.



Figure 3.3. Semi-logarithmic graph of the mass of salvinorin A versus extraction time. Error bars represent \pm one standard deviation (n=9).



Time (min)

Figure 3.4. Representative total ion chromatogram 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



Salvinorin A (mg/mL)

Figure 3.5. GC-MS calibration curve of salvinorin A, showing linearity ($y = 4 \times 10^7 x - 5 \times 10^6$, $R^2 = 0.9951$) over the concentration range investigated of 0.10 - 12.00 mg/mL salvinorin A.

3.4 Conclusion

The optimal extraction procedure, resulting in the greatest mass of salvinorin A and the highest interday precision, involved extracting *S. divinorum* leaves with dichloromethane for five minutes. This rapid and simple procedure provided an extraction efficiency of 97.6% and an interday precision of 9.6%. The information obtained from this investigation will be useful to forensic analysts in states where *S. divinorum* or salvinorin A are currently regulated. In addition, in the event of further state or potentially federal regulation of *S. divinorum* or salvinorin A, results from this research could form the basis for the development of a standardized extraction procedure.

REFERENCE

References

- 1. US Department of Justice, D. E. A. *Drugs and Chemicals of Concern: Salvia divinorum and Salvinorin A.* December 2010.
- 2. House Bill 6038, Ninety-fifth Legislature State of Michigan, October 2010.
- 3. Jermain, J.; Evans, H., Analyzing *Salvia divinorum* and its Active Ingredient Salvinorin A Utilizing Thin Layer Chromatography and Gas Chromatography/ Mass Spectrometry. *J. Forensic Sci.* 2009, 54, 1-5.
- 4. Siebert, D., Localization of Salvinorin A and Related Compounds in Glandular Trichomes of the Psychoactive Sage, *Salvia divinorum*. *Ann. Bot. London*, 2004, 93, 763-771.
- Gruber, J. W.; Siebert, D. J.; Der Marderosian, A. H.; Hock, R. S., High Performance Liquid Chromatographic Quantification of Salvinorin A from Tissues of *Salvia divinorum* Epling & Jativa-M. *Phytochem. Analysis* 1999, 10, 22-25.
- 6. Methanol; MSDS No. M2015[Online]; Mallinckrodt Baker: Phippsburg, NJ, May 25, 2005, http://www.jtbaker.com/msds/englishhtml/m2015.htm (accessed March, 2010).
- 7. Acetone; MSDS No. A0446 [Online]; Mallinckrodt Baker: Phippsburg, NJ, May 25, 2005, http://www.jtbaker.com/msds/englishhtml/a0446.htm (accessed March, 2010).
- 8. Dichloromethane; MSDS No. M4420[Online]; Mallinckrodt Baker: Phippsburg, NJ, May 25, 2005, <u>http://www.jtbaker.com/msds/englishhtml/M4420.htm</u> (accessed March, 2010).
- 9. Hexane; MSDS No. H2381[Online]; Mallinckrodt Baker: Phippsburg, NJ, May 25, 2005, http://www.jtbaker.com/msds/englishhtml/h2381.htm (accessed March, 2010).
- Tsjikawa, K; Kuwayama, K.; Miyaguchi, H.; Kanamori, T.; Iwata, Y.; Yoshida, T.; Inoue, H., Determination of salvinorin A and salvinorin B in *Salvia divinorum*-related products circulated in Japan, *Forensic Sci. Int.* 2008, 180, 105-109.
- 11. Barton, A., Solubility parameters. Chem. Rev. 1975, 75, 731.

 Leger, C.; Lira, H.; Paterson, R. 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. Memb. Sci.* 1996, 120, 135-146.

CHAPTER 4

DIFFERENTION OF S. DIVINORUM FROM OTHER SALVIA SPECIES AND ASSOSIATION OF ADULTERATED PLANT MATERIALS TO S. DIVINORUM

4.1 Introduction

If *S. divinorum* or salvinorin A were to be further regulated, a means to differentiate *S. divinorum* from other *Salvia* species, as well as a method to associate adulterated plant materials to *S. divinorum*, would be highly beneficial for forensic laboratories. The *Salvia* genus is particularly difficult to characterize based on the physical morphology of the plant and it has been reported that differentiation between *S. divinorum* and other *Salvia* species is not possible through visual inspection of the plant alone [1]. In addition, a common strategy for distributing a controlled substance is to adulterate another regulated or non-regulated material with the substance. At least one case of this has been documented, where the toxic psychosis of a patient whose *Cannabis sativa* (marijuana) cigarette had been adulterated with the leaves and leaf extract of *S. divinorum* was reported [2].

In this chapter, the development of a method for the differentiation of *S. divinorum* from other *Salvia* species in order to prevent false positive and negative identifications is reported. Salvinorin A is currently only known to exist in *S. divinorum* [1]; therefore a visual comparison of chromatograms of questioned *Salvia* samples should provide a means of differentiation. A recent report by the U.S. 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 [3]. Hence, a more objective method for the differentiation

of *S. divinorum* from other *Salvia* species or adulterated plant materials is desirable. Multivariate statistical procedures can be used to assess association and discrimination among samples. 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 [4, 5].

Principal component 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 [6]. 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 [4] and for examining the resulting batch-to-batch variation in heroin cutting and distribution [5]. Despite this, PCA has not yet been investigated for discrimination of *S. divinorum* from other *Salvia* species or association of adulterated samples to *S. divinorum*.

As chromatograms of plant materials are very complex, containing hundreds of variables, PCA may not be able to identify the chemical sources of variance. Consequently, data pretreatment steps may be necessary prior to statistical analysis in order to minimize nonchemical sources of variance. Common pretreatment procedures for chromatographic data include background subtraction, smoothing, retention time alignment, and normalization [6, 7]. Background subtraction is often performed on chromatographic data to minimize low-frequency noise, primarily due to drift in the background signal. Smoothing is performed to minimize highfrequency 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.

Another potential problem with PCA is that interpretation of the 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, student t-tests, Wilcoxon rank-sum tests, hierarchical cluster analysis (HCA), and Pearson product moment correlation (PPMC) coefficients. The Euclidian distance represents the numerical distance between two points in multidimensional space [8]. 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 [8]. Additionally, HCA allows for a graphical representation of the association among samples in a data set, based on the similarity of the data. The student t-test and the Wilcoxon rank-sum test can be used to assess discrimination of samples at given confidence levels. The t-test assumes the data are normally distributed, whereas the Wilcoxon test makes no assumptions regarding the distribution of the data [8].

Pearson product moment correlation (PPMC) coefficients provide a means to compare

the similarity of the complete chromatogram. Coefficients are calculated for pairwise comparisons of chromatograms, for example, a questioned sample and a reference standard, allowing comparison of 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. A coefficient of ± 0.80 or greater indicates 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 [8]. As such, PPMC coefficients can be applied to assess similarity between samples, especially those positioned closely in the PCA scores plot.

The optimized extraction procedure developed in Chapter 3 will be used for the extraction of five different *Salvia* species (*S. divinorum, S. officinalis, S. guaranitica, S. splendens*, and *S. nemerosa*). Extracts will be analyzed by GC-MS and the resulting chromatograms subjected to various data pretreatment procedures prior to PCA, to investigate association and discrimination of the five *Salvia* species. Euclidean distances, student's t-tests, Wilcoxon rank-sum tests, HCA, and PPMC coefficients will be evaluated to provide an objective interpretation of the resulting PCA scores plot. Four different plant materials (*S. divinorum, S. officinalis, Cannabis sativa* and *Nicotiana tabacum*) will be spiked with *S. divinorum* extract to simulate an adulterated sample that might be encountered in a forensic laboratory. Similar statistical procedures will be used to assess association of the adulterated plant materials to *S. divinorum*. The multivariate statistical procedures will then be compared to determine the advantages and disadvantages of each for the association and discrimination of *S. divinorum*, particularly in forensic casework.

4.2 Materials and Methods

4.2.1 Differentiation of S. divinorum from other Salvia Species

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 hours. Dried leaves of *S. officinalis* were purchased from a commercial supplier (Penzeys Spices, Brookfield, WI). The five *Salvia* species were then extracted in triplicate using the optimized procedure described in Chapter 3, spiked with the internal standard, and analyzed by GC-MS.

4.2.2 Adulteration of Plant Materials

To simulate adulterated samples that might be encountered in a forensic laboratory, a method for extracting *S. divinorum* leaves was obtained from an online source [9]. For the extraction, 15.6 g dried *S. divinorum* leaves were placed in approximately 600 mL of acetone and soaked for 5 min. The solution was filtered and evaporated as described in Section 3.2.1. The extract was then reconstituted in 26.1 mL acetone and 6.0 mL of extract were added to 0.8 g of *S. divinorum* and 6.7 mL were added to 0.8 g of three other plant materials: *S. officinalis*, *N. tabacum* (Bugler Original Turkish and Blended Cigarette Tobacco, Lane Ltd., Tucker, GA), and *C. sativa* (Michigan State Police, Forensic Science Division, East Lansing, MI) to simulate a 5x fortified sample. A smaller volume of extract was used for adulteration of *S. divinorum*, as salvinorin A is inherent in the plant material. The adulterated plant materials were dried using nitrogen and subsequently placed in a desiccator, after which the dried weight was recorded.

Each of the adulterated plant materials was divided into 0.2 g samples and extracted using the optimized extraction method (Section 3.2.1). The triplicate extractions were then spiked with the internal standard and analyzed by GC-MS.

4.2.3 GC-MS Analysis

All extracts were analyzed by GC-MS in triplicate using the same instrument parameters as previously described in Section 3.2.3.

4.2.4 Data Analysis

Chromatograms of all extracts were subjected to pretreatment prior to data analysis. Firstly, background subtraction 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, 355) were generated in 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 2 data point window (OriginPro, version 7.5853, Origin Lab Corp., Northampton, MA).

After background correction and smoothing, the chromatograms were divided into two separate data sets. The first set comprised all *Salvia* species (*S. divinorum*, *S. officinalis*, *S.*

guaranitica, *S. splendens*, and *S. nemorosa*), while the second set consisted of the plant materials and adulterated plant materials (*S. divinorum*, *S. officinalis*, *C. sativa*, and *N. tabacum*). The two data sets were separately subjected to retention time alignment and normalization prior to data analysis. The chromatograms were retention time aligned in Matlab (version 7.4.0.287, MathWorks, Natick, MA) using a peak-matching algorithm [10] with a window size of 4 data points. The target chromatogram used for alignment was created by summing one replicate chromatogram of each plant and dividing by the total number of plants in the respective data set.

Chromatograms in each data set were then normalized. Normalization of the chromatograms was problematic due to differences in the number and abundance of compounds extracted from each plant material. Initially, normalization was performed using progesterone an internal standard (IS). However, in the second data set, PCA identified a single component, delta-9-tetrahydrocannabinol (THC) in C. sativa, as 99.0% of the total variance. These results were caused by the significant difference in the absolute abundance of THC extracted from C. sativa and the abundances of the components extracted from the other plant materials. Total area normalization was also investigated; however PCA identified salvinorin A as 98.5% of the total variance in both the data sets. As salvinorin A was, in effect, the only component extracted from S. divinorum, it accounts for the majority of the total area. Total area normalization, therefore, was ineffective in normalizing the full data set. To account for the significant difference in the absolute abundance of THC extracted from C. sativa and the components extracted from the other plant materials, maximum peak normalization was also investigated. 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 respective 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.

Each pretreated data set was then subjected to principal components analysis. For the first data set, the PCA scores for the *Salvia* species were calculated in Matlab and plotted in Microsoft Excel to generate the scores plot. Similarly, eigenvectors generated in Matlab were plotted against retention time in Microsoft Excel to generate the loadings plots. The 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).

For the second data set, PCA was performed on the chromatograms of the unadulterated plant materials to generate eigenvectors based only on the intrinsic differences in chemical composition of the plants. PCA scores and loadings plots were generated in Microsoft Excel, as described above. The TICs of the adulterated plant materials were then mean centered relative to the TICs of the unadulterated materials. To do so, the average abundance at each retention time in the unadulterated plant materials was subtracted from the corresponding data point in each replicate of the adulterated plant materials. The mean-centered data for each replicate were multiplied by the eigenvectors generated for the unadulterated materials. These values were then summed, resulting in a score for that replicate that was then projected onto the scores plot of the unadulterated plant materials. The scores were projected in this way so that differences between adulterated and unadulterated plant materials were primarily due to the adulterant, salvinorin A, rather than variations in the analytical methodology and analysis procedures.

Several procedures were evaluated to statistically assess the PCA scores plots for both

data sets. Mean scores on principal components 1 and 2 (PC1 and PC2, respectively) were calculated for all samples. Euclidian distances were calculated (Microsoft Excel) between the mean score of each plant material and *S. divinorum* in PC1 and PC2. Hierarchical cluster analysis was performed in Pirouette (version 2.02, Infometrix, Woodinville, WA) using Euclidian distance with single-cluster linkage to assess similarity of the PCA scores. Finally, the unequal-variance student t-test and Wilcoxon rank-sum test were performed (Microsoft Excel) on the mean scores for each plant material in PC1 and PC2.

PPMC coefficients derived from PCA loadings were calculated in Matlab for each data set from the full chromatogram [8]. The chromatograms that had been pretreated and mean centered to the respective data set were multiplied by the eigenvector, PC1 or PC2, that allowed the greatest differentiation of the extracts. This method enabled a pairwise comparison of the chromatograms based on the variance described by the associated PC.

4.3 Results and Discussion

4.3.1 Differentiation of S. divinorum from other Salvia Species

The optimized extraction procedure described in Chapter 3 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 Figure 4.1 with chemical identities of selected compounds defined in Table 4.1. Many of the chemical compounds in the *Salvia* chromatograms are terpenes or alkanes



Plant Material	Peak	Retention Time (min)	Chemical Identity
S. 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
S. divinorum	8	26.620	Salvinorin A
S. nemorosa	9	23.100	Pyridaben
	7	25.623	Octacosane
		26.858	Eicosane
S. splendens	9	23.100	Pyridaben
	10	23.707	Squalene
	7	24.302	Hexacosane
		25.623	Octacosane
		26.858	Eicosane
S. guaranitica	11	25.322	Falcarinol
	12	26.334	Columbin
	7	24.302	Hexacosane
		25.623	Octacosane
		26.858	Eicosane

Table 4.1. Tentative chemical identification of compounds of Salvia species in Figure 4.1.

commonly found in plant materials.

The chromatograms of *S. divinorum* (Figure 3.4) and the other *Salvia* species (Figure 4.1) 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.

Principal component analysis was then investigated as an objective method for differentiation and the resulting scores plot is shown in Figure 4.2. 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 the PCA 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 (Figure 4.3A and 4.3B, respectively). 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



Figure 4.2. Scores plot of principal components 1 and 2 for *Salvia* species. Representative example shown of Euclidian distance between *S. divinorum* and *S. guaranitica*. *S. divinorum*, *S. officinalis*, *S. guaranitica*, *S. nemorosa*, *S. splendens*


Figure 4.3. Loadings plots of (A) principal component 1 and (B) principal component 2 for scores plot of *Salvia* species.

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.

While a visual examination of the scores plot can give an indication of the association and discrimination of samples, such visual interpretation is still somewhat subjective. 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 4.2 with one example also demonstrated on the PCA scores plot (Figure 4.2). In each case, the distance between the mean scores of *S. divinorum* and the 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 [8]).

Hierarchical cluster analysis was then performed on the scores of each Salvia species.

Table 4.2. Euclidian distance between *Salvia* species and *S. divinorum* relative to Euclidian distance between replicates of each species.^a

	Euclidian Distance between Replicates ^b	Euclidian Distance to S. divinorum	Euclidian Distance Ratio
S. divinorum	$3.03 \pm 1.87 \times 10^4$		
S. officinalis	$2.81 \pm 1.65 \text{ x10}^5$	1.62×10^{6}	5.8
S. guaranitica	$1.70 \pm 1.14 \text{ x} 10^5$	7.92×10^5	4.7
S. nemorosa	$7.57 \pm 4.50 \text{ x}{10}^4$	1.33×10^{6}	17.5
S. splendens	$1.80 \pm 1.05 \text{ x10}^5$	1.48×10^{6}	8.2

^a Euclidian distance calculated based on scores from principal component analysis

^b mean \pm standard deviation (n = 36)

The branch lengths of the dendrogram, shown in Figure 4.4, represent the similarity index among the plant materials [8]. The similarity index between replicates from all extractions ranged from 0.983 for *S. divinorum* to 0.776 for *S. officinalis*, indicating that *S. divinorum* showed the least variability and *S. officinalis* the most. 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 [12].

The student t-test and the Wilcoxon rank-sum test were performed, comparing the mean score of *S. divinorum* with each of the four *Salvia* species, and the results are reported Tables 4.3 and 4.4, respectively). Using the student 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 PC to provide a quantitative assessment of the scores plot. The mean PPMC coefficients (R_{mean}) for PC1 and PC2 are summarized in Table 4.5 A and B. For PC1, strong correlations were observed among replicate extractions for each species ($R_{mean} = 0.9477 - 0.9981$). When compared with



Figure 4.4. Hierarchical cluster analysis dendrogram showing similarity among Salvia species.

	S. divin	norum	S. officinalis		S. guaranitica		S. nemorosa		% Confidence
	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2	Level
S. divinorum									99.9
S. officinalis	19.96 (8)	23.12 (8)							99.0
S. guaranitica	14.41 (8)	22.61 (9)	24.84 (14)	10.88 (12)					98.0
S. nemorosa	22.70 (12)	82.28 (9)	25.20 (9)	6.27 (10)	5.91 (10)	31.32 (15)			95.0
S. splendens	12.57 (9)	44.24 (8)	23.82 (13)	7.69 (15)	2.65 (16)	24.58 (14)	2.92 (10)	3.55 (11)	90.0
									50.0
									20.0

0.0

Table 4.3. T-statistic (with associated degrees of freedom) for comparison of *Salvia* species. Cell color represents confidence level on heat map.

	S. divinorum		S. offi	S. officinalis S. gua		ranitica	S. nemorosa	
	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2
S. divinorum								
S. officinalis	-3.58	-3.58						
S. guaranitica	3.58	-3.58	3.58	3.58				
S. nemorosa	3.58	-3.58	3.58	-3.58	3.58	-3.58		
S. splendens	3.58	-3.58	3.58	-3.58	3.58	-3.58	-2.96	2.96

Table 4.4. Wilcoxon rank-sum statistic for comparison of *Salvia* species. Cell color represents confidence level on heat map in Table 4.3.

	S. divinorum	S. officinalis	S. guaranitica	S. nemorosa	S. splendens
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. nemorosa	0.2906 ± 0.1233	-0.8673 ± 0.0192	0.4825 ± 0.0734	0.9675 ± 0.0195	
S. splendens	0.1567 ± 0.0922	-0.8432 ± 0.0361	0.5343 ± 0.0556	0.9235 ± 0.0416	0.9616 ± 0.0194

Table 4.5A. Pearson product moment correlation coefficients derived from principal component 1 for Salvia species.^a

Table 4.5B. Pearson product moment correlation coefficients derived from principal component 2 for Salvia species.^a

	S. divinorum	S. officinalis	S. guaranitica	S. nemorosa	S. splendens
S. divinorum	0.9919 ± 0.0067	-0.1689 ± 0.2525	0.1130 ± 0.0284	-0.7914 ± 0.0110	-0.7652 ± 0.0342
S. officinalis		0.8626 ± 0.0745	0.0190 ± 0.2968	-0.0942 ± 0.3703	-0.1262 ± 0.3437
S. guaranitica			0.9787 ± 0.0114	-0.6021 ± 0.0251	-0.5920 ± 0.0393
S. nemorosa				0.9861 ± 0.0117	0.9154 ± 0.0304
S. splendens					0.9549 ± 0.0517

^a mean \pm standard deviation (n = 81)

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 Figure 4.2. In addition, there was strong correlation between *S. splendens* and *S. nemorosa*, 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.

For PC2, strong correlations were observed among replicate extractions for each species $(R_{mean} = 0.8626 - 0.9919)$. Weak to moderate correlations were observed between *S. divinorum* and the other *Salvia* species ($R_{mean} = -0.7652 - 0.1130$) in agreement with their position on the PCA scores plot in Figure 4.2. The strong correlation between *S. splendens* and *S. nemorosa* was still observed. PPMC coefficients calculated from both PC1 and PC2 had low to moderate correlation of *S. divinorum* to the other *Salvia* species. As both PC described the variance between the species in a relatively similar manner, either could be used for placing a statistical confidence behind the PCA results.

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 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 [8]. The student 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 [8]. In contrast, the Wilcoxon rank-sum test is nonparametric and does not assume a normal distribution. This test is considered to be as effective as the student t-test for normally distributed data and, in cases where the data are not normally distributed, more effective than the t-test [8]. Similar conclusions were obtained for the student ttest 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 the 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.

4.3.2 Adulteration of Plant Materials

Representative chromatograms of each of the unadulterated and adulterated plant materials are shown in Figures 4.5 and 4.6, respectively, with chemical identities of selected compounds defined in Table 4.6. The resulting PCA scores and loadings plots are shown in



Figure 4.5 Representative total ion chromatograms of unadulterated (A) C. sativa, (B) N.

tabacum.



Figure 4.6 Representative total ion chromatograms of adulterated plant materials (A) *S. officinalis*, (B) *S. divinorum*, (C) *C. sativa*, (D) *N. tabacum*.

Plant Material	Peak	Retention Time (min)	Chemical Identity
C. sativa	13	20.788	Cannabidiol
	14	21.612	Tetrahydrocannabinol
	15	22.093	Cannabinol
N. tabacum	16	3.228	Xylene
	17	9.596	Nicotine
	18	15.358	Eicosyne
	19		Alkanes

Table 4.6. Tentative chemical identification of compounds in plant materials in Figure 4.5.

Figures 4.7 and 4.8, respectively.

The first two PCs account for 76% of the total variance among the unadulterated plant materials. Replicates of each individual extract of the unadulterated plant materials are clustered closely and are distinct from *S. divinorum* based on visual inspection. Extracts of *S. officinalis* are positioned positively on PC1 due to the chemical compounds previously mentioned that load positively on this PC (Figure 4.8A). In contrast, the other plant materials are positioned negatively on PC1, as their active compounds (nicotine, tetrahydrocannabinol (THC), and salvinorin A) load negatively. Extracts of *N. tabacum* are positioned positively on PC2 due to the presence of xylene, nicotine, eicosyne, and various alkanes that load positively on PC2 due to the presence of cannabidiol, THC, and cannabinol. Compounds from *S. divinorum* and *S. officinalis* do not contribute to the variance described by PC2 and, as a result, both species are positioned near zero on the scores plot.

Scores for the adulterated plant materials were calculated using the eigenvectors for the unadulterated plant materials and then projected onto the scores plot (Figure 4.6). Each of the adulterated plant materials is clustered closer to unadulterated *S. divinorum* than to the corresponding unadulterated plant material based on visual assessment of the scores plot.

Euclidian distances were calculated between the mean scores of the unadulterated plant materials, adulterated plant materials, and the corresponding replicates. The distances are shown in Table 4.7 with one example also demonstrated on the PCA scores plot (Figure 4.7). In each case, the distance between the mean scores of *S. divinorum* and the unadulterated plant materials was 2.4 - 6.1 times greater than that between replicates of those species. Therefore, *S. divinorum*



Figure 4.7. Scores plot of principal components 1 and 2 for unadulterated plant materials (closed interior). Scores for adulterated plant materials (open interior) are projected onto the plot (see text). Representative example shown of Euclidian distance between adulterated *C. sativa* and both unadulterated *S. divinorum* and unadulterated *C. sativa*.

📕 S. divinorum, 🔺 S. officinalis, 🛛 C. sativa, 🔻 N. tabacum



Figure 4.8. Loadings plots of (A) principal component 1 and (B) principal component 2 for scores plot of unadulterated and adulterated plant materials.

Table 4.7. Euclidian distance of (A) unadulterated samples to *S. divinorum* relative to Euclidian distance between replicates, (B) adulterated samples to *S. divinorum* relative to Euclidian distance between adulterated and unadulterated samples.^a

(A) Unadulterated	Euclidian Distance	Euclidian Distance to	Euclidian Distance
	between Replicates	S. divinorum	Ratio
			-
S. divinorum	$3.08 \pm 1.83 \text{ x10}^4$		
S. officinalis	$7.30 \pm 4.93 ext{ x10}^5$	4.43×10^{6}	6.1
C. sativa	$3.00 \pm 1.72 \text{ x10}^4$	1.30×10^{6}	4.3
N. tabacum	$8.07 \pm 5.37 \text{ x}{10}^4$	1.90×10^{6}	2.4
	Euclidian Distance to		
(B) Adulterated	Unadulterated Matrix		
	_	-	
S .divinorum	6.01×10^{2}	6.01×10^{2}	
S. officinalis	3.36×10^{6}	1.07×10^{6}	0.32
C. sativa	1.55×10^{6}	8.10×10^{5}	0.52
N. tabacum	1.68 x10 ⁶	5.73×10^{5}	0.34

^a Euclidian distance calculated based on scores from principal component analysis

^b mean \pm standard deviation (n = 36)

is distinct from the other plant materials as the distance is clearly more than three times the standard deviation among replicates. The distance between *S. divinorum* and the adulterated plant materials was 0.32 - 0.52 times less than the distance between the corresponding unadulterated plant materials and the adulterated plant materials. Thus, plant materials adulterated with extracts of *S. divinorum* are more closely associated to *S. divinorum* than to the respective unadulterated plant material.

Hierarchical cluster analysis was performed on the scores of each of the unadulterated and adulterated plant materials and the dendrogram is shown in Figure 4.9. The similarity index between replicates ranged from 0.684 for unadulterated *S. officinalis* to 0.989 for unadulterated *C. sativa*, indicating that the variability in replicate extractions was greatest for *S. officinalis* and least for *C. sativa*. The similarity index between *S. divinorum* and the other plant materials ranged from 0.000 for *S. officinalis* to 0.374 for *C. sativa*, indicating that *S. divinorum* was most similar to *C. sativa* and least similar to *S. officinalis*. Although it may be surprising that two plants from the same species are the least similar, *S.divinorum* does not contain the many volatile compounds that are present in *S. officinalis* (Figures 3.4 and 4.1). It is noteworthy that the adulterated plant materials were most closely associated to *S. divinorum*, with a similarity index of 0.818, rather than to the corresponding unadulterated plant materials, which is consistent with the results of the Euclidian distances above.

The student t-test and the Wilcoxon rank-sum test were performed, comparing the mean score of each unadulterated and adulterated plant material with *S. divinorum* and the results are reported in Tables 4.8 and 4.9, respectively. The unadulterated plant materials were statistically distinguishable from each other and from *S. divinorum* at the 99.9% confidence level on both



Figure 4.9. Hierarchical cluster analysis dendrogram showing similarity among unadulterated and adulterated plant materials.

Table 4.8. Unequal variance t-statistic (with associated degrees of freedom) for comparison of (A) unadulterated plant materials, (B) unadulterated plant materials with adulterated plant materials, and (C) adulterated plant materials. Cell color represents confidence level on heat map in Table 4.3.

(A)	S. divinorum		<i>C. s</i>	C. sativa		S. officinalis		
	PC1	PC2	PC1	PC2	PC1	PC2		
S. divinorum								
C. sativa	9.36 (9)	140.35 (13)						
S. officinalis	21.55 (8)	0.17 (9)	21.24 (8)	49.13 (10)				
N. tabacum	31.95 (14)	89.48 (9)	31.58 (8)	145.40 (10)	19.58 (8)	57.17 (15)		
(B)	S. divi	norum	<i>C. se</i>	ativa	S. offi	cinalis		
	PC1	PC2	PC1	PC2	PC1	PC2		
S. divinorum								
C. sativa	6.78 (13)	0.75 (10)						
S. officinalis	6.00 (8)	0.67 (11)	4.35 (8)	0.14 (16)				
N. tabacum	0.96 (15)	6.16 (14)	5.56 (8)	6.68 (9)	6.87 (15)	6.65 (9)		
(C)	S. divi	norum	<i>C. s</i>	ativa	S. offi	cinalis	N. tal	расит
	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2
S. divinorum	16.92 (9)	4.57 (11)	15.31 (8)	96.74 (14)	18.39 (8)	1.90 (11)	5.42 (9)	77.33 (13)
C. sativa	17.63 (8)	10.14 (16)	16.63 (8)	149.28 (13)	15.69 (9)	2.41 (9)	10.91 (8)	86.49 (9)
S. officinalis	107.46 (11)	9.79 (16)	226.85 (13)	147.54 (13)	17.63 (8)	2.37 (9)	37.26 (9)	86.35 (9)
N. tabacum	11.27 (8)	10.34 (9)	9.97 (8)	78.24 (11)	18.43 (9)	5.89 (14)	2.88 (8)	62.96 (16)

Table 4.9. Wilcoxon rank-sum statistic for comparison of (A) unadulterated plant materials, (B) unadulterated plant materials with adulterated plant materials, and (C) adulterated plant materials. Cell color represents confidence level on heat map in Table 4.3.

(A)	S. div	vinorum	С	sativa	S. off	ficinalis		
	PC1	PC2	PC1	PC2	PC1	PC2		
S. divinorum								
C. sativa	3.58	3.58						
S. officinalis	3.58	0.84	3.58	-3.58				
N. tabacum	3.58	-3.58	3.58	-3.58	-3.58	-3.58		
(B)	S. div	vinorum	С	sativa	S. off	ficinalis		
	PC1	PC2	PC1	PC2	PC1	PC2		
S. divinorum								
C. sativa	3.58	0.13						
S. officinalis	3.58	-0.04	3.58	-0.31				
N. tabacum	-0.4	-3.58	-3.58	-3.58	-3.58	-3.58		
(C)	S. div	vinorum	<i>C</i> . <i>s</i>	sativa	S. off	ficinalis	N. tal	bacum
	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2
S. divinorum	3.58	-3.31	3.58	-3.31	-3.58	-1.46	3.58	3.58
C. sativa	3.58	-3.58	3.58	-3.58	-3.58	-1.19	3.58	3.58
S. officinalis	3.58	-3.58	3.58	-3.31	-3.58	-1.19	3.58	3.58
N. tabacum	3.58	-3.58	3.58	-3.31	-3.58	-3.58	2.25	3.58

PC1 and PC2. The sole exception is *S. officinalis* and *S. divinorum*, which were distinguishable at the 99.9% confidence level on PC1 but were not distinguishable on PC2. As previously mentioned, compounds from these species do not contribute to the loadings for PC2 (Figure 4.8B) which prevents their differentiation.

Using the student t-test, the adulterated plant materials were statistically distinguishable from *S. divinorum* at the 99.9% confidence level on both PC1 and PC2. Unfortunately, this indicates that the adulterated samples cannot be directly associated to *S. divinorum*, which would be desirable for forensic identification. However, adulterated *S. divinorum* (which has a higher concentration of salvinorin A) could not be distinguished from the adulterated plant materials at the 99.9% confidence level on either PC1 or PC2. In fact, there was such high association that adulterated *S. divinorum* was only statistically distinguished from adulterated *N. tabacum* and *C. sativa* at the 50% and 20% confidence levels, respectively, on PC2. Adulterated *S. divinorum* could not be distinguished from *S. officinalis* at any confidence level on PC1. Hence, an enriched enriched sample of *S. divinorum* could be used as the standard to identify such adulterated plant materials. Analogous statistical association and differentiation was provided by the Wilcoxon test at similar confidence levels.

The PPMC coefficients were calculated for all pairwise combinations of unadulterated plant materials using the product of the mean-centered chromatogram and the eigenvector of each PC and resulting coefficients are shown in Table 4.10 and 4.11. Although PC1 accounted for 56% of the variance in the data set, the loadings plot (Figure 4.8) was dominated by the compounds in *S. officinalis*. In contrast, PC2 accounted for only 20% of the variance, but contained significant contributions from the other plant materials. For PC1, strong correlations

Table 4.10. Pearson product moment correlation coefficients derived from principal component 1 for (A) unadulterated plant materials, (B) adulterated plant materials with their corresponding unadulterated plant materials and *S. divinorum*.^a

(A) Unadulterated				
Unadulterated	S. divinorum	S. officinalis	C. sativa	N. tabacum
S. divinorum	0.9886 ± 0.0095		-	-
S. officinalis	-0.6647 ± 0.0060	0.9848 ± 0.0057		
C. sativa	-0.0399 ± 0.0105	$-0.4950 \pm \ 0.0073$	0.9938 ± 0.0048	
N. tabacum	0.0713 ± 0.0184	-0.7434 ± 0.0494	0.4205 ± 0.0354	0.9946 ± 0.0040
(B) Adulterated	S. divinorum	S. officinalis	C. sativa	N. tabacum
Unadulterated		-0.4486 ± 0.1072	0.4598 ± 0.0063	0.4718 ± 0.3173
S. divinorum	0.3722 ± 0.2513	0.8222 ± 0.1381	0.0968 ± 0.0183	0.5812 ± 0.3340

Table 4.11. Pearson product moment correlation coefficients derived from principal component 2 for (A) unadulterated plant materials, (B) adulterated plant materials with their corresponding unadulterated plant materials and *S. divinorum*.^a

(A) Unadulterated				
Unadulterated	S. divinorum	S. officinalis	C. sativa	N. tabacum
S. divinorum	0.9972 ± 0.0016			
S. officinalis	-0.3900 ± 0.0046	0.9664 ± 0.0145		
C. sativa	0.2932 ± 0.3017	-0.5771 ± 0.0752	0.9898 ± 0.0078	
N. tabacum	-0.6995 ± 0.2932	0.3291 ± 0.2373	-0.3125 ± 0.0280	0.9961 ± 0.0030
(B) Adulterated	S. divinorum	S. officinalis	C. sativa	N. tabacum
Unadulterated		0.3058 ± 0.2871	-0.3425 ± 0.0040	0.5028 ± 0.1531
S. divinorum	0.9009 ± 0.0309	0.9762 ± 0.0150	0.8667 ± 0.0103	0.7959 ± 0.1028

^a mean \pm standard deviation (n = 81)

0.9664 - 0.9972). Weak correlations were observed between unadulterated *S. divinorum* and each of the unadulterated plant materials ($R_{mean} = 0.2932 - 0.3291$), in agreement with their position on the PCA scores plot in Figure 4.7. Similarly for PC2, strong correlations were observed among replicate extractions for each unadulterated plant material ($R_{mean} = 0.9664 - 0.9972$). Weak to inversely moderate correlation was observed between unadulterated *S. divinorum* and each of the unadulterated plant materials ($R_{mean} = -0.6995 - 0.2932$), in agreement with their position on the PCA scores plot in Figure 4.7. For both PC1 and PC2, the adulterated plant materials had strong correlations to *S. divinorum* with weak to moderate correlations to the corresponding unadulterated plant material (Table 4.10 and 4.11).

The PPMC coefficients using PC1 showed strong correlation of adulterated *S. officinalis* to *S. divinorum* and weak to moderate correlation ($R_{mean} = 0.0968 - 0.5812$) of the other adulterated plant materials to *S. divinorum* (Table 4.8A). These results differ from the PCA observations. However, the PPMC coefficients using PC2 further confirm the PCA observations as each of the adulterated plant materials have a strong correlation (≥ 0.7959) to *S. divinorum* with weak correlation ($R_{mean} = -0.3425$ to 0.3058) to the corresponding neat plant material (Table 4.11). In this case, the two PCs do not describe the variance between the plant materials in a similar manner, and cannot be used equally for placing a statistical confidence behind the PCA results. The first PC identifies *S. officinalis* as the greatest variance in the data set, but does not differentiate the other plant materials (Figure 4.7). In contrast, the second PC differentiates *C. sativa, S. divinorum*, and *N. tabacum*, and is therefore superior to PC1 in describing the variance of the whole data set. The PPMC coefficients should be calculated using the PC that is

responsible for the greatest differentiation of the plant materials, in this case PC2.

In view of these results, plant materials adulterated with an extract of *S. divinorum* can be statistically associated to *S. divinorum* using PCA in combination with Euclidian distance, HCA, the student t-test, the Wilcoxon rank-sum test, or PPMC coefficients. Many of the inherent advantages and limitations described in the previous study were also apparent here. Accordingly, Euclidian distances and HCA would be excellent visual tools for courtroom presentation, whereas the student t-test would be useful when a more rigorous statistical evaluation of evidence was necessary.

4.4 Conclusions

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.

Four plant materials were then spiked with an extract of *S. divinorum* to simulate an adulterated sample that might be submitted to a forensic laboratory. The unadulterated and adulterated materials were extracted and analyzed by GC-MS. Again, by visual inspection of the chromatograms, the adulterated materials were associated to *S. divinorum* based on the presence of salvinorin A. Objective association was also demonstrated using PCA, where the adulterated plant materials were closely positioned to *S. divinorum* on the scores plot, but distinct from the native plant materials.

Several statistical methods were investigated to provide a numerical evaluation of the positioning in the PCA scores plot. These methods were used for discrimination, in the case of the *Salvia* species, as well as for association, in the case of the adulterated plant materials. Among these methods, hierarchical cluster analysis and Pearson product moment correlation 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 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 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 U.S. National Academies of Science, National Research Council [3].

REFERENCES

References

- 1. Munro, T. The Chemistry of *Salvia divinorum*. Ph.D. Dissertation, 2006, University of Melbourne, Melbourne, Australia, p 283.
- 2. Paulzen, M.; Grunder, G., Toxic Psychosis after Intake of the Hallucinogen Salvinorin A. *J. Clin. Psychiatry*. 2008, 69, 1501-1502.
- Strengthening Forensic Science in the United States: A Path Forward. National Research Council of the National Academies. The National Academies Press, Washington D.C. 2009.
- Esseiva, P.; Anglada, F.; Dujourdy, L.; Taroni, F.; Margot, P.; Pasquier, E.; Dawson, M.; Roux, C.; Doble, P., Chemical Profiling and Classification of Illicit Heroin by Principal Component Analysis, Calculation of Inter-Sample Correlation and Artificial Neural Networks. *Talanta* 2005, 67, 360-367.
- 5. Klemenc, S., In Common Batch Searching of Illicit Heroin Samples-Evaluation of Data by Chemometric Methods. *Forensic Sci. Int.* 2001, 115, 43-52.
- 6. Brereton, R. Applied chemometrics for students. West Sussex (England): John Wiley and Sons, 2007.
- Hupp, A. M.; Marshall, L. J.; Campbell, D. I.; Smith, R. W.; McGuffin, V. L., Chemometric Analysis of Diesel Fuel for Forensic and Environmental Applications. *Anal. Chim. Acta* 2008, 606, 159-171.
- 8. Devore, J. Probability and statistics for engineering and the sciences. Belmont, CA (United States): Duxbury Press, 1990.
- 9. Anonymous (Lint). Extraction success notes, *Salvia divinorum*. Erowid Experience Vaults Blog. http://www.erowid.org/experiences/exp.php?ID=7439, 2003.
- 10. Johnson K, Wright B, Jarman K, Synovec R. High-speed peak matching algorithm for

retention time alignment of gas chromatographic data for chemometric analysis. J Chromatogr A 2003; 996: 141-55.

- 11. Brereton, R., Applied chemometrics for students. England: John Wiley and Sons, 2007.
- 12. Walker, J.; Sytsma, K.; Treutlein, J.; Wink, M., *Salvia* (Lamiaceae) is not Monophyletic: Implications for the Systematics, Radiation, and Ecological Specializations of *Salvia* and Tribe Mentheae. *Am. J. Bot.* 2004, 91, 1115-1125.

CHAPTER 5

CONCLUSIONS AND FUTURE WORK

This research involved the investigation of three aspects of identification of *S. divinorum* for forensic purposes. First, extraction solvents and times were systematically compared to develop an optimized procedure for the extraction of salvinorin A from *S. divinorum*. This optimized procedure was then used to investigate both the subjective and objective differentiation of *S. divinorum* from other *Salvia* species. Lastly, the ability of chemometric procedures to associate adulterated plant matrices to *S. divinorum* plant material was investigated. Each of these three goals was successfully accomplished and the results are summarized in the following sections.

5.1 Development of an Optimized Procedure for the Extraction of Salvinorin A from *S*. *divinorum*

An extraction procedure was optimized through the comparison of six extraction solvents of varying polarities (methanol, acetonitrile, acetone, chloroform, dichloromethane, and hexane) and six extraction times (1, 3, 6, 30, 100, and 300 minutes). Dichloromethane extracted the greatest mass of salvinorin A, with the least extraction of other plant components and the highest interday precision. Dichloromethane was therefore chosen as the optimal extraction solvent and was used to investigate the different extraction times. The greatest mass of salvinorin A was extracted in the time range 3 to 6 minutes (0.0796 - 0.0882 mg/g). As there was no statistical difference between the mass of salvinorin A extracted at these two time points, an extraction time of 5 minutes was chosen for convenience. The extraction efficiency of the optimized procedure was determined using *S. officinalis* as the matrix, due to the inherent variability of

salvinorin A observed between leaf samples of *S. divinorum* [11]. Salvinorin A was found to be linear over the concentration range tested of 0.10 mg/mL to 12.00 mg/mL. An extraction efficiency of 97.6% was determined with 1.95 ± 0.20 mg/mL recovered from a 2.00 mg/mL spike.

Using this procedure, the greatest mass of salvinorin A, with the highest interday precision, was extracted with an average extraction efficiency of 97.6%. This procedure allows for a fast extraction that is simple to perform and has high reproducibility.

5.2 Differentiation of *S. divinorum* from other *Salvia* Species

Four other *Salvia* species (*S. officinalis, S. guaranitica, S. splendens* and *S. nemorosa*) in addition to *S. divinorum* were extracted using the optimized procedure and analyzed by gas chromatography-mass spectrometry (GC-MS). Chromatograms of the *Salvia* species were visually compared to *S. divinorum*. None of the other *Salvia* species had a peak at the same retention time as salvinorin A, nor were any of the mass spectra of the components of the other *Salvia* species consistent with that of salvinorin A. Therefore, it was determined that a subjective differentiation of the chromatograms of *Salvia* species is possible based on the presence or absence of the salvinorin A peak in the resulting chromatograms.

The combination of principal component analysis (PCA), Euclidian distance measurement, student t-test, Wilcoxon rank-sum, hierarchal cluster analysis (HCA), and Pearson product moment correlation (PPMC) coefficients were investigated as objective means of differentiating *Salvia* species. The *Salvia* species were separated from *S. divinorum* on the PCA scores plot and the statistical significance of the separation confirmed using Euclidian distances, student t-test, Wilcoxon rank-sum, HCA, and PPMC coefficients. The Euclidian distance

between *S. divinorum* and each of the other *Salvia* species in the scores plot was between 4.7-17.5 times the distance between the respective species replicates, indicating that the other *Salvia* species are distinct from *S. divinorum* in the scores plot.

Using the student t-test, each of the *Salvia* samples were statistically distinguishable from *S. divinorum* at a 99.9% confidence level on both PC1 and PC2. Analogous statistical differentiation was observed using the Wilcoxon test as with the student t-test, at similar confidence levels.

Using HCA, the similarity of *S. divinorum* with the other *Salvia* species ranged from 0.528 to 0.000, which was consistent with the clustering observed on the PCA scores plot and the association demonstrated with Euclidian distances. Weak to no correlation was observed for the PPMC coefficients of *S. divinorum* with each of the other *Salvia* species ($R_{mean} = 0.2906 - 0.4428$), which further confirmed the differentiation of the species observed in the scores plot. Therefore, differentiation of *S. divinorum* from other *Salvia* species was possible through both a visual comparison of chromatograms and the use of chemometric procedures.

5.3 Association of Adulterated Plant Materials to S. divinorum

Four plant matrices (*S. divinorum, S. officinalis, C. sativa* and *N. tabacum*) were adulterated with an extract of *S. divinorum* to simulate adulterated samples a forensic laboratory might encounter. The adulterated plant matrices were extracted using the optimized extraction method and analyzed by GC-MS. PCA, Euclidian distances, student t-test, Wilcoxon rank-sum, HCA, and PPMC coefficients were then performed on the resulting data. Each of the neat plant species are individually clustered from *S. divinorum* on the scores plot. Each of the adulterated plant matrices are clustered closer to *S. divinorum* than to their corresponding neat plant material. Euclidian distances were used as a statistical measure of the association of the adulterated plant material to *S. divinorum*. In each case, the distance between *S. divinorum* and the unadulterated plant materials was between 2.4- 6.1 times the distance between the respective unadulterated replicates, indicating that the unadulterated plant materials are distinct from *S. divinorum* in the scores plot. In addition, the distances between the adulterated plant materials and *S. divinorum* were between 0.52 and 0.32 times the distance between the adulterated plant materials and the corresponding unadulterated plant material. The Euclidian distances then indicate that the adulterated plant materials can be more closely associated to *S. divinorum* than to the respective unadulterated plant material.

Using the student t-test, the adulterated samples were statistically distinguishable from *S*. *divinorum* at the 99.9% confidence level on both PC1 and PC2, indicating that the samples cannot be associated with *S*. *divinorum*. The adulterated samples could not be distinguished from adulterated *S*. *divinorum* at the 99.9% confidence level in either PC1 or PC2, confirming the visual association observed in the PCA scores plot. Analogous statistical association and differentiation was observed using the Wilcoxon test as with the student t-test, at similar confidence levels.

Using HCA, the adulterated plant materials were most closely associated with *S*. *divinorum* at a similarity level of 0.818, which is consistent with the clustering observed on the PCA scores plot and the association demonstrated with Euclidian distances, the student t-test, and the Wilcoxon rank-sum test. The similarity with the other unadulterated plant materials ranged from 0.374 to 0.000.

PPMC coefficients further confirmed this association as each of the adulterated matrices had a strong correlation (≥ 0.7959) to *S. divinorum* with weak to no correlation ($R_{mean} = -0.3425$ to 0.3058) to their native populations. Therefore, the objective association of extracts of the adulterated plant material to *S. divinorum* was possible using chemometric procedures.

The multivariate statistical procedures were compared to determine the advantages and disadvantages of each for use in a forensic setting. 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 [1]. The student 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 [1]. 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 t-test for normally distributed data and, in cases where the data are not normally distributed, more effective than the t-test [1]. Similar conclusions were obtained for the student ttest 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 the 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.

These results preliminarily indicate that submitted samples adulterated with *S. divinorum* can be objectively associated with the adulterant using PCA in combination with the student t-test, the Wilcoxon test, HCA, or PPMC coefficients.

5.4 Future Work

While these studies are useful and applicable to the forensic field, there are many other studies that should be conducted to further these investigations. For instance, while extraction times and solvents were systematically investigated, only one extraction procedure was used in this study. It may be beneficial to investigate different types of extraction procedures, such as rotary agitation or sonication, in order to find a more efficient extraction procedure. In addition, only pure solvents were investigated in this research. An investigation examining the use of mixed solvents, for example dichloromethane and chloroform, in various ratios may also result in increased extraction efficiency.

This research examined the differentiation of *S. divinorum* from four other *Salvia* species, however close to one thousand *Salvia* species exist. Using the procedure developed in this research, it may be useful to examine the differentiation of a larger sample set of *Salvia* species. This would increase confidence in differentiation which would be beneficial for defending these results in a legal setting.

The ability to associate extracts of plant matrices adulterated with *S. divinorum* to *S. divinorum* was also demonstrated in this research. The method of extraction and spiking of *S. divinorum* was adapted from an online source to demonstrate a sample that might be encountered
by a forensic laboratory. However, since this was a proof-of-principle study, various adaptations were made. For example, acid-washed glassware and laboratory grade solvents were used to decrease the impurities. It would be necessary to investigate how, or if, these impurities affected the association of adulterated extracts to *S. divinorum*. As only one concentration of adulteration was investigated in this work, an additional investigation of the concentration range at which the plant materials adulterated with *S. divinorum* could still be associated to *S. divinorum* is also necessary.

Online suppliers sell *S. divinorum* fortified leaves in which concentrated extracts (50x, 20x, 10x, 5x and 1x) of *S. divinorum* are coated onto the dried leaves. Including a sampling of these fortified leaves in the data set to determine association to *S. divinorum* would be a useful addition to this study as well. This would demonstrate that a range of types and concentrations of submitted samples could still be associated or discriminated from *S. divinorum*.

While GC-MS is widely available in forensic laboratories for the identification of controlled substances, it is limited in its ability to only detect small, volatile components. Liquid chromatography- mass spectrometry (LC-MS) instrumentation is becoming more widely available and is used for the analysis of larger, more polar compounds. The investigation of *S. divinorum* using LC-MS would be an interesting addition to this research. LC-MS would allow identification of other component(s) that may prove to be more discriminating for *S. divinorum* than salvinorin A alone. In addition, the combination of GC-MS and LC-MS might allow for a chemical fingerprint of *S. divinorum* to be compiled. Chemometric procedures could then be used to increase the confidence in forensic classification of this plant material.

134

REFERENCES

References

1. Devore, J. Probability and statistics for engineering and the sciences. Belmont, CA (United States): Duxbury Press, 1990.