COMPARISON OF GAS CHROMATOGRAPHY-MASS SPECTROMETRY AND LIQUID CHROMATOGRAPHY-MASS SPECTROMETRY FOR DISCRIMINATION OF SALVIA DIVINORUM FROM RELATED SALVIA SPECIES USING CHEMOMETRIC PROCEDURES

By

Monica L. Bugeja

A THESIS

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

MASTER OF SCIENCE

Forensic Science

2012

ABSTRACT

COMPARISON OF GAS CHROMATOGRAPHY-MASS SPECTROMETRY AND LIQUID CHROMATOGRAPHY-MASS SPECTROMETRY FOR DISCRIMINATION OF SALVIA DIVINORUM FROM RELATED SALVIA SPECIES USING CHEMOMETRIC PROCEDURES

By

Monica L. Bugeja

Salvia divinorum is a hallucinogenic plant that has recently gained legislative attention due to an increase in its recreational use. *S. divinorum* and/or its active constituent, salvinorin A, are currently regulated in 26 states, including Michigan. As there are over 900 known *Salvia* species, discrimination of *S. divinorum* from other *Salvia* species is important in a forensic laboratory. Salvinorin A is known to be found only in *S. divinorum* thus far. Therefore, the presence of salvinorin A, detected using gas chromatography-mass spectrometry (GC-MS), serves as the current method for identifying *S. divinorum.* However, salvinorin A can be extracted from *S. divinorum* to prepare liquid extracts for recreational use. This can result in low levels of salvinorin A in the residual plant material that may or may not be detectable, causing possible difficulty with identification. Hence, an alternative procedure for identifying *S. divinorum*, such as using a full chemical fingerprint, would be beneficial.

The purpose of this research was to generate chemical fingerprints of *S. divinorum* and four related *Salvia* species using GC-MS to analyze the volatile compounds and liquid chromatography-mass spectrometry (LC-MS) to analyze the non-volatile compounds. The differentiation of *S. divinorum* from the other species based on chemical fingerprints was then investigated using statistical procedures.

ACKNOWLEDGMENTS

The completion of this thesis could not be possible without the help and encouragement of several individuals. First, I would like to thank my family for listening to me complain and stress throughout my years of schooling, supporting me financially as I continued my education, and raising me the way I am. I would not be the person I am today without everything my family has taught me. In addition, I would like to thank a good friend and boyfriend, Daniel Kelly, who has given me advice, support, and encouragement throughout this project. I would also like to thank him for the time he has spent coming to visit me and taking me away from this thesis project every once in a while to be able to laugh and relax. Furthermore, I would like to thank the Forensic Chemistry Group, who have supported and assisted me with this research. I would especially like to recognize John McIlroy, who has taught me so much throughout the year. I am so grateful for all of his help and could not thank him enough. I would like to recognize Melissa Bodnar Willard, whose previous research on *S. divinorum* created a path to this research. I would also thank her for all of her guidance. I would like to thank Suzanne Towner, who has been a wonderful, supportive, encouraging, and helpful friend throughout this project, and Tiffany VanDeMark, who has also been a supportive and kind friend. I would like to acknowledge my friends and fellow scientists at the Michigan State Police, who have continuously supported and motivated me throughout my schooling.

In addition, I would like to thank the Midwest Forensics Resource Center for funding this project. Dr. Daniel Jones and staff at Michigan State University's Mass Spectrometry Facility for access to and help with the liquid chromatograph-mass spectrometer and interpretation of mass spectra. I would also like to thank Dr. Kevin Walker for access to the *Salvia* plants and Jim Klug and staff at Michigan State University's research greenhouse for maintaining the *Salvia* plants throughout this project. Finally, I would like to thank my committee members, Dr. Ruth Waddell Smith, Dr. Victoria McGuffin, and Dr. Mahesh Nalla for their help, guidance, and patience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Introduction

1.1 What is *Salvia divinorum*?

Salvia divinorum is a hallucinogenic plant originating from Oaxaca, Mexico that has recently gained legislative attention due to an increase in its recreational use [1-8]. There are over 900 *Salvia* species that are divided into three clades based on genetics and plant structure, as well as geographical origin [9]. Of these three clades, *S. divinorum* belongs to clade II and is a member of the Lamiaceae family [6].

S. divinorum consists of several chemical compounds; however, only one chemical compound is known to contribute to its hallucinogenic nature, a trans-neoclerodane diterpene named salvinorin A [1, 2, 4, 5]. This compound was first isolated from *S. divinorum* by Ortega *et al.* in 1982 [5, 10]. Ortega *et al.* determined the structure of salvinorin A (Figure 1.1) using infrared (IR) spectroscopy, carbon and hydrogen nuclear magnetic resonance (^{13}CNMR and ^{1}HNMR, respectively), single-crystal X-ray analysis, and mass spectrometry (MS) [10]. Using these instrumental techniques, the molecular formula of salvinorin A was determined to be $C_{23}H_{28}O_8$. Using MS with electron ionization, salvinorin A had a molecular ion of *m/z* 432, and a base peak of *m/z* 94 [3].

Other chemical compounds identified in *S. divinorum* include salvinorins A through I, divinatorins A through F, salvidivins A through D, and hardwickiic acid [1, 2, 6, 11, 12]. Salvinorin A is typically the most abundant compound, followed by salvinorins B, C, and D. While salvinorins E and F, divinatorins A through C, and hardwickiic acid may be present, these compounds are typically at very low abundance. Salvinorins A through D, which are believed to be produced for protection of the plant, are located in the peltate glandular trichomes that appear

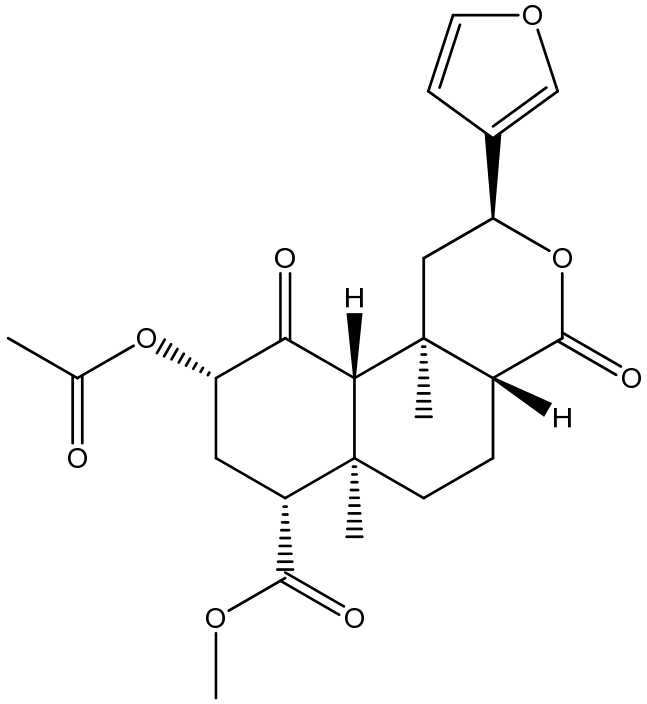

Figure 1.1: Chemical structure of salvinorin A.

as hair-like structures mostly on the leaves, but also on external stem tissues, and various other parts of the plant [8, 11]. Other compounds found in *S. divinorum* include salvincins A and B, loliolide, neophytadiene, stigmasterol, oleanolic acid, presqualene alcohol, peplusol, (E)-phytol, sugars, fatty acids, vitamins, plant sterols, alkanes, and tocopherol (vitamin E) [3, 12, 13]. Geographical location, environmental conditions, and age can affect the composition and concentration of these compounds [11, 14]. While numerous compounds have been identified, the entire chemical composition of *S. divinorum* is unknown [1, 2, 6, 11].

1.2 Use, Abuse, and Effects of *Salvia divinorum*

S. divinorum has previously been used by the Mazatec Indians for treatment of arthritis, headaches, and problems with excretion [11]. *S. divinorum* may be helpful for treatment of Alzheimer's, schizophrenia, and stimulant dependence; however, in the United States, there are currently no accepted medical uses [8, 12]. More recently, *S. divinorum* has been used recreationally.

The 2011 National Survey on Drug Use and Health Report approximated 175,000 people aged 12 or older used *S. divinorum* in their lifetime [15]. Methods of administration include chewing, smoking, and ingesting liquid extracts of *S. divinorum* [2]. Salvinorin A, the active compound, is absorbed through the oral mucosa, and eventually reaches the brain where it binds to the kappa opioid receptors causing several effects, which are similar to those caused by lysergic acid diethylamide (LSD) [8, 6, 11, 12, 16–18]. These effects include hallucinations, dizziness, slurred speech, and/or uncontrollable laughter and body movements. The onset of the effects can occur within seconds and, depending on the dose and route of administration, the

effects may last from minutes to an hour [4, 17]. However, there are currently no known toxic effects of salvinorin A [17].

1.3 Regulation

As a result of its recent recreational use as a hallucinogenic drug, several countries, which include Australia, Belgium, Denmark, Estonia, Finland, Italy, Japan, Spain, and Sweden, have regulated *S. divinorum* and/or salvinorin A [8, 19]. While the United States Drug Enforcement Administration (DEA) has included *Salvia divinorum* and salvinorin A on its Drugs and Chemical of Concern, they are not currently federally regulated [1, 4, 5, 7]. However, *Salvia divinorum* and/or salvinorin A are currently regulated in half of the United States [1, 5, 7, 19, 20]. As of October 1, 2010, *S. divinorum* and salvinorin A are regulated as schedule I substances in Michigan, which indicates that these substances have potential for abuse and no accepted medical use.

1.4 Extraction and Analysis of *Salvia divinorum*

1.4.1 Extraction Procedures

Due to its recent regulation in several states, the ability to identify *S. divinorum* and differentiate it from other plant materials is imperative in a forensic context [5, 9]. However, various extraction procedures are used to extract salvinorin A from *S. divinorum* and no standardized procedure currently exists [1].

Jermain and Evans compared different extraction procedures currently used by forensic laboratories to extract salvinorin A from *S. divinorum* [1]. The following extraction procedures were compared: (1) methanol added to *S. divinorum* leaves and vortexed at room temperature, as

used at the Tennessee Bureau of Investigations Central Laboratory in Nashville, (2) 0.1 N sodium hydroxide and methylene chloride added to leaves, as used at the Missouri State Highway Patrol Forensic Laboratory, (3) acetone added to *S. divinorum* leaves for 1 minute, as used at the Drug Enforcement Administration's Special Testing and Research Laboratory, (4) boiling chloroform added to *S. divinorum* leaves for 10 minutes, as used at the Oklahoma State Bureau of Investigation, (5) boiling chloroform added to leaves for 1 minute, which was modified from the above procedure, (6) and chloroform added to leaves at room temperature for 1 minute, which was also modified from the above procedure.

All extracts were analyzed using gas chromatography-mass spectrometry (GC-MS) [1]. Since *S. divinorum* is typically identified based on the presence of salvinorin A, the extraction procedures were compared based on the abundance of salvinorin A and the number of additional plant compounds extracted. The presence of salvinorin A was confirmed by comparing the mass spectrum from the extract to that of a salvinorin A standard. Of the six procedures, extracting with chloroform or acetone for one minute was determined to be optimal. In these extracts, maximum abundance of salvinorin A was extracted with minimal number of additional plant compounds. However, salvinorin A can be extracted from *S. divinorum* to prepare liquid extracts for recreational use [2, 6, 12]. This could result in low levels of salvinorin A in the residual plant material that may or may not be detectable, causing difficulties with identification. Hence, an alternative procedure for identifying *S. divinorum* based on extracting several plant compounds would be beneficial.

1.4.2 Analysis Procedures

In the previous study, all extracts were analyzed using GC-MS, which is one of the most common separation techniques used by forensic laboratories for controlled substance analysis. However, GC analysis is limited to the analysis of thermally stable, volatile compounds. Liquid chromatography can be used to analyze non-volatile, thermally unstable compounds that may not be suitable for GC analysis. For example, liquid chromatography (LC) with ultraviolet (UV) detection at 208 nm has also been used to analyze extracts of *S. divinorum.* Gruber *et al.* analyzed several *S. divinorum* extracts using this technique [21]. Based on the retention time of a reference standard, salvinorin A was detected in all of the extracts. Extracts of various other *Salvia* species (*i.e. S. concolor, S. blepharophylla, S. chiapensis, S. gregii, S. leucantha, S. membranacea,* and *S. recurva*) were also analyzed but salvinorin A was not detected in any of these species. However, LC-UV can only be used for screening purposes since UV detection does not provide definitive identification of compounds.

High performance liquid chromatography-electrospray ionization-tandem mass spectrometry (LC-ESI-MS^2) can provide detailed structural information of chemical compounds; therefore, compounds can be definitively identified. Tsujikawa *et al.* analyzed commercially available products of *S. divinorum* (*i.e.* leaf and concentrated extract products) by LC-ESI-MS^2, definitively identifying salvinorins A and B in each based on comparison with reference spectra and subsequent spectral interpretation [6]. Plant extracts of several *Salvia* species and varieties (*i.e. S. elegans* "Scarlet Pineapple", *S. forskaohlei, S.* "Indigo Spires", *S. glutinosa, S. guaranitica, S. blepharophylla, S. Mexicana* "Limelight", *S. buchananii*, *S. splendens*, and *S. farinacea*) were also analyzed using LC-ESI-MS^2, but neither salvinorin A nor salvinorin B was

identified. While definitive identification of salvinorin A in *Salvia* products was demonstrated, a procedure for identifying *S. divinorum* based on several plant compounds rather than salvinorin A alone would be helpful, as explained previously. To do this, additional *Salvia* species need to be analyzed and compared to *S. divinorum* using complete total ion chromatograms (TICs) and mass spectra.

As previously described, several researchers have demonstrated successful identification of salvinorin A in *S. divinorum* using GC-MS, LC-UV and LC-MS^2 [1, 6, 21]. Medana *et al.* compared the limits of detection of GC-MS, LC-UV, and liquid chromatography-multistage mass spectrometry (LC-MS^n), which included LC-MS^2 and LC-MS^3, for the detection of salvinorins A through D in reference standards [3]. Other compounds, which included salvinorins E and F, were analyzed using LC-UV to investigate the separation of salvinorins and using LC-MS^n to determine fragment ions to assist in identification of these compounds. Standards of divinatorins A through C were also analyzed using LC-MS^n. However, none of the afore-mentioned compounds were quantified. LC-UV not only had the poorest limit of detection (*i.e.* 367 ng/mL for salvinorin A) of the instrumental techniques investigated, but also the UV detector did not allow definitive identification of compounds. Both GC-MS and LC-MS^n provide structural information so that definitive identification is possible; however, LC-MS^2 and LC-MS^3 had lower limits of detection for salvinorins A through D than GC-MS (*e.g.*, 3 and 9 ng/mL for salvinorin A using LC-MS^2 and LC-MS^3, respectively compared to 40 ng/mL using GC-MS). Therefore, due to improved sensitivity, LC-MS would be more useful than GC-MS in cases where the salvinorins are present in low concentration. However, other objective comparisons

may be beneficial, such as the use of statistical procedures to evaluate and compare these techniques.

1.4.3 Statistical Procedures

Both GC and LC have been previously used to analyze extracts of *Salvia* species. However, the data produced from these analyses are very complex and comparison of chromatograms can be difficult and even subjective. In 2009, the National Academy of Sciences published a report titled "Strengthening Forensic Science in the United States: A Path Forward," which stated 13 recommendations to improve the field of forensic science [22]. Of these recommendations, one stated the need for more research to evaluate procedures, especially those that may be subjective. Therefore, more objective procedures, such as the use of statistical procedures, may be beneficial in forensic analyses [19, 22].

Statistical analyses have been previously applied to *Salvia* data. Zhou *et al.* extracted and analyzed 25 samples of 12 *Salvia* species (*S. divinorum* was not included in this data set) by ultra-high-performance liquid chromatography-tandem mass spectrometry [23]. The resulting chromatograms were subjected to various pretreatment procedures (*i.e.* peak finding, peak alignment and peak filtering) to minimize non-chemical sources of variance in the data, prior to statistical analysis. Principal component analysis (PCA) was then performed to determine which compounds could be used as chemical markers for the different species. In the resulting scores plot, the 12 species formed three clusters. The first cluster included samples extracted from the roots of plants, while the second cluster included samples extracted from the aerial parts of the plants. The third cluster included two samples: one extracted from the aerial part of the plant and one extracted from the whole plant. There was a number of different species within each cluster.

Although three clusters were apparent to the authors, it does not appear that any statistical assessment of the scores plot was performed to determine this clustering. According to the loadings plot, dihydrotanshinone I was the most variable compound and, thus, could be used as a chemical marker to differentiate samples extracted from roots of plants and those extracted from the aerial parts of the plants.

Carrer *et al.* also used PCA to assess *Salvia* data [14]. Leaves of *S. guaranitica* from seven different locations in Brazil were analyzed by GC-MS. Several compounds in *S. guaranitica* were identified, which included α-pinene, β-cubebene, β-caryophyllene, δ-cadinene, caryophyllene oxide, and phytol. The presence of the compounds varied based on the geographical origin of *S. guaranitica*. In the PCA scores plot, the *S. guaranitica* samples were grouped in three clusters based on variation in composition. However, no additional statistical assessment of the positioning of the clusters was performed.

Bodnar Willard *et al.* were the first to use PCA, along with various other statistical procedures, to assess the discrimination of *S. divinorum* from related *Salvia* species (*i.e. S. guaranitica, S. nemorosa, S. officinalis*, and *S. splendens*) based on TICs obtained using GC-MS [19]. To minimize non-chemical sources of variance, the data were background corrected, smoothed, retention time aligned, and normalized prior to statistical analysis. Using PCA, *S. divinorum* was easily discriminated from the other *Salvia* species through visual assessment of the scores plot.

Several additional statistical procedures were then conducted to provide a more objective interpretation of positioning of the *Salvia* species on the scores plot [19]. Euclidean distances were calculated between *S. divinorum* and each of the other species, as well as between replicates of the same species, based on the mean scores. Using this metric, *S. divinorum* was

distinct from the other species since the Euclidean distance was 4.7–17.5 times greater than the distance observed between replicates of the same species. Student's *t*-tests and Wilcoxon rank-sum tests were also performed on the mean PCA scores for each species. Using both tests, *S. divinorum* was statistically differentiated from the other *Salvia* species on both principal component 1 (PC1) and principal component 2 (PC2), at the 99.9 % confidence level. Hierarchical cluster analysis (HCA) was also performed on the mean PCA scores to assess similarity, or otherwise, of the species. This procedure showed that there was greater similarity among replicates of the same species (similarity index = 0.776–0.983) than between *S. divinorum* and the other *Salvia* species (similarity index = 0.000–0.528). Finally, to further assess similarity, PPMC coefficients were calculated between replicates of the same species, as well as between *S. divinorum* and the other *Salvia* species. Specifically, the coefficients were calculated for pair-wise comparisons of the loadings of each sample on PC1. Mean PPMC coefficients (r = 0.9477–0.9981) indicated that replicates were strongly correlated, while *S. divinorum* was weakly correlated to the other *Salvia* species (r = 0.2906–0.4428). Therefore, *S. divinorum* was differentiated from the other *Salvia* species using all these procedures, which offered an objective interpretation of the scores plot.

Although Bodnar Willard *et al*. demonstrated statistical discrimination of *S. divinorum* from other Salvia species based on the full TICs, the extraction procedure was optimized for the extraction of salvinorin A, rather than the entire chemical fingerprint [19]. In addition, the extracts were only analyzed by GC-MS, which is limited to the analysis of thermally stable, volatile compounds. As previously mentioned, LC-MS is generally more sensitive and can be used to analyze non-volatile compounds. Therefore, LC-MS may provide additional chemical information about the plant material.

1.5 Identification of other *Salvia* species

Over 900 *Salvia* species exist and, hence, the ability to identify *S. divinorum* and differentiate it from other plant material is imperative in a forensic context [5, 9]. Several *Salvia* species, which include *S. guaranitica, S. nemorosa, S. officinalis,* and *S. splendens*, have been analyzed previously using various analytical procedures such as gas chromatography-flame ionization detection (GC-FID), GC-MS, LC-MS^2, and ^{13}CNMR [24–27]. Many compounds from these species have been identified, some of which are common among species, including *S. divinorum.*

1.5.1 S. guaranitica

S. guaranitica, a plant used for decorative and medicinal purposes, belongs to the same clade as *S. divinorum*, clade II [7, 9, 14, 28]. Vallverdú *et al.* determined the chemical composition of *S. guaranitica* using GC-FID, GC-MS, and ^{13}CNMR to investigate possible medicinal uses [24]. A total of 49 compounds were identified, which included mostly sesquiterpenes, oxygenated sesquiterpenes, and monoterpenes. Phytol, an oxygenated diterpene, previously found in *S. divinorum* was also identified in *S. guaranitica* [3, 13, 14, 24].

1.5.2 S. nemorosa

S. nemorosa, a plant used for decorative purposes, belongs to clade I [7, 9, 28]. Malenčić *et al.* analyzed *S. nemorosa* by GC-FID and GC-MS to determine its chemical composition [25]. A total of 23 compounds were identified, which included monoterpenes, sesquiterpenes, and other hydrocarbons. Some of these compounds included α-pinene, α-cadinol, α-copaene, α-

humulene, *p*-cymene, δ-cadinene, and caryophyllene oxide, which have also been identified in *S. guaranitica*. Takeda *et al.* also identified salvinosides A through C in *S. nemorosa* by UV, ^{1}HNMR, ^{13}CNMR, negative ion fast atom bombardment-mass spectrometry, and circular dichroism spectroscopy [26]. Ulubelen *et al.* identified nemorosin and salvinemorol in *S. nemorosa* by UV spectroscopy, IR spectroscopy, ^{1}HNMR spectroscopy, and thin layer chromatography (TLC) [29]. Other compounds present in *S. nemorosa* included oleanolic acid, which was previously found in *S. divinorum*, sitosterol, and ursolic acid [3, 12, 13, 29].

1.5.3 S. officinalis

S. officinalis, an herb that has been researched greatly due to its medicinal properties, belongs to clade I [7, 9, 24, 30–32]. This plant material has antiviral, antitumor, and antibacterial properties [27, 30, 33]. Radulescu *et al.* analyzed *S. officinalis* by GC-MS to develop an extraction procedure and determine its chemical composition [30]. A total of 35 compounds were identified, some of which were common to other *Salvia* species. For example, β-pinene, β-caryophyllene, and β-elemene were also identified in *S. guaranitica*, while *p*-cymene and α-humulene were also found in *S. guaranitica* and *S. nemorosa*. In addition, oleanolic acid was identified in *S. nemorosa* and *S. divinorum*. Vitamin E was also found in *S. divinorum*. In a review of the polyphenolic compounds in *Salvia*, Lu *et al.* reported that cirsiliol was identified in both *S. officinalis* and *S. guaranitica* [32]. Djarmati *et al.* extracted *S. officinalis* and identified manool, rosmanol 7-ethyl ether, oleanolic acid, and carnosic acid 12-methyl ether-γ-lactone using combinations of IR, ^{1}HNMR, ^{13}CNMR, TLC, and X-ray diffraction [34]. Abreu *et al.* extracted and analyzed leaves of 60 different *Salvia* species, including *S. officinalis,* using LC-

MS^2 and identified carnosic acid, carnosol, methoxycarnosic acid, isorosmanol, and vitamin E [33]. Kivilompolo *et al.* analyzed *S. officinalis* by GC-MS, LC-UV, and LC-MS to compare qualitative and quantitative results of specific compounds [27]. Several compounds were detected; however, all compounds were not detected by every technique. For instance, rosmarinic acid was detected by LC-UV and LC-MS, but not by GC-MS. Other compounds that have also been identified in *S. officinalis* include α-thujone, β-thujone, camphor, and 1-borneol [31].

1.5.4 S. splendens

S. splendens, a plant used for decorative purposes, belongs to the same clade as *S. divinorum*, clade II [7, 9, 28]. Several compounds have been identified in *S. splendens*, some of which are structural congeners of salvinorin A [35]. These compounds included salvisplendins A through D, salviarin, splendidin, and splenolides A and B, which were extracted from *S. splendens* and analyzed using a combination of techniques such as IR, UV, ^{1}HNMR, ^{13}CNMR, and MS.

1.6 Objective

With the limitations in previous studies, the objective of this research was to investigate the association and discrimination of *S. divinorum* from other *Salvia* species (*i.e. S. guaranitica, S. nemorosa, S. officinalis,* and *S. splendens*) based on both volatile and non-volatile compounds. In order to achieve this objective, several goals had to be met. First, all species were extracted and analyzed using GC-MS and LC-MS to generate chemical fingerprints based on the volatile and non-volatile compounds, respectively. Compounds present in the TICs were tentatively

identified using available MS library databases and literature. Data pretreatment procedures were performed to minimize non-chemical sources of variance. To assess the association and discrimination of *S. divinorum* from the other four *Salvia* species, PCA was performed separately on the GC and LC data. To objectively evaluate the PCA scores plot, Student's *t*-tests were performed on the mean scores of each species. In addition, PPMC coefficients were calculated between TICs to further assess this association and discrimination. The results from the statistical analyses of the GC and LC data were then compared to determine the analytical technique that provided more discriminating information for *S. divinorum*. Since there are many *Salvia* species, the ability to discriminate *S. divinorum* from related *Salvia* species is important in a forensic laboratory. Currently, *S. divinorum* is identified by the presence of salvinorin A; however, salvinorin A can be extracted from the leaf material. Therefore, an alternative procedure to identify *S. divinorum* based on several compounds may be useful.

References

References

[1] Jermain JD, Evans HK. 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.

[2] 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; 20:131–6.

[3] Giroud C, Felber F, Augsburger M, Horisberger B, Rivier L, Mangin P. *Salvia divinorum:* an hallucinogenic mint which might become a new recreational drug in Switzerland. Forensic Sci Int 2000; 112:143–50.

[4] Lange JE, Reed MB, Ketchie Croff JM, Clapp JD. College student use of *Salvia divinorum*. Drug and Alcohol Depen 2008; 94:263–6.

[5] McDonough PC, Holler, JM, Vorce SP, Bosy TZ, Magluilo J Jr, Past MR. The detection and quantitative analysis of the psychoactive component of *Salvia divinorum,* Salvinorin A, in human biological fluids using liquid chromatography-mass spectrometry. J Anal Toxicol 2008; 32:417–21.

[6] Tsujikawa K, Kuwayama K, Miyaguchi H, Kanamori T, Iwata YT, Yoshida T, Inoue H. Determination of salvinorin A and salvinorin B in *Salvia divinorum*-related products circulated in Japan. Forensic Sci Int 2008; 180:105–9.

[7] Bodnar Willard MA. Forensic analysis of *Salvia divinorum* and related *Salvia* species using chemometric procedures. MS Thesis (2010).

[8] Drug Enforcement Agency. *Salvia divinorum* and salvinorin A. 2010. www.deadiversion. usdoj.gov/drugs_concern/**salvia**_d.pdf

[9] Walker JB, Sytsma KJ, 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; 91:1115–25.

[10] Ortega A, Blount JF, Manchand PS. Salvinorin, a new *trans*-neoclerodane diterpene from *Salvia divinorum* (Labiatae). J Chem Soc Perkin Trans 1 1982; 2505–8.

[11] Siebert, DJ. Localization of salvinorin A and related compounds in glandular trichomes of the psychoactive sage, *Salvia divinorum*. Annals of Botany 2004; 93:763–71.

[12] Grundmann O, Phipps SM, Zadezensky I, Butterweck V. *Salvia divinorum* and salvinorin A: An update on pharmacology and analytical methodology. Planta Med 2007; 73:1039–46.

[13] Bigham AK, Munro TA, Rizzacasa MA, Robins-Browne RM. Divinatorins A–C, new neoclerodane diterpenoids from the controlled sage *Salvia divinorum*. Journal of Natural Products 2003; 66:1242–4.

[14] Carrer RP, Vanderlinde R, Dutra S, Marcon A, Echeverrigaray S. Esstential oil varation among Brazilian accessions of *Salvia guaranitica* L.. Flavour Fragr. J. 2007; 22:430–4.

[15] Results from the 2011 national survey on drug use and health: detailed tables. 2011. http://www.samhsa.gov/data/NSDUH/2011SummNatFindDetTables/NSDUH-DetTabsPDFWHTML2011/2k11DetailedTabs/Web/HTML/NSDUH-DetTabsSect7seTabs1to45-2011.htm#Tab7.28C

[16] Kennedy JH, Wiseman JM. Direct analysis of *Salvia divinorum* leaves for salvinorin A by thin layer chromatography and desorption electrospray ionization multi-stage tandem mass spectrometry. Rapid Commun Mass Sp 2010; 24:1305–11.

[17] Lange JE, Daniel J, Homer K, Reed MB, Clapp JD. *Salvia divinorum*: Effects and use among YouTube users. Drug and Alcohol Depen 2010; 108:138–40.

[18] Drug Enforcement Administration, Office of Diversion Control, Drug and Chemical Evaluation Section. *Salvia divinorum* and Salvinorin A. December 2010.

[19] Bodnar Willard MA, McGuffin VL, Waddell Smith R. Forensic analysis of *Salvia divinorum* using multivariate statistical procedures. Part I: discrimination from related *Salvia* species. Anal Bioanal Chem 2012; 402:833–42.

[20] Enrolled House Bill No. 6038. State of Michigan 95th Legisature, Regular Session. 111 (2010).

[21] Gruber JW, Siebert DJ, Der Marderosian AH, Hock RS. High performance liquid chromatographic quantification of salvinorin A from tissues of *S. divinorum* Epling & Játiva-M. Phytochem. Anal. 1999; 10:22–5.

[22] Committee on identifying the needs of the forensic sciences community, National Research Council. Strengthening Forensic Science in the United States: A Path Forward. 2009. https://www.ncjrs.gov/pdffiles1/nij/grants/228091.pdf

[23] Zhou Y, Xu G, Fung Kei Choi F, Ding L, Bin Han Q, Song J, Feng Qiao C, Zhao Q, Xu H. Qualitative and quantitative analysis of diterpenoids in *Salvia* species by liquid chromatography coupled with electrospray ionization quadrupole time-of-flight tandem mass spectrometry. Journal of Chromatography A 2009; 1216:4847–58.

[24] Vallverdú C, Vila R, Lorenzo D, Paz D, Dellacassa E, Davies P, Villamil J, Tomi F, Casanova J, Cañigueral S. Composition of the essential oil of cultivated *Salvia guaranitica* from Uruguay. Flavour Fragr. J. 2005; 20:421–4.

[25] Malenčić D, Couladis M, Mimica-Dukić N, Popović M, Boža P. Essential oils of three *Salvia* species from the Pannonian part of Serbia. Flavour Fragr. J. 2004; 19:225–8.

[26] Takeda Y, Zhang H, Matsumoto T, Otsuka H, Oosio Y, Honda G, Tabata M, Fujita T, Sun H, Sezik E, Yesilada E. Megastigmane glycosides from *Salvia Nemorosa.* Phytochemistry 1997; 44:117–20.

[27] Kivilompolo M, Obůrka V, Hyötyläinen T. Comparison of GC-MS and LC-MS methods for the analysis of antioxidant phenolic acids in herbs. Anal Bioanal Chem 2007; 388:881–7.

[28] Tychonievich, J. *Salvia* Clades. Personal Communication. 2009.

[29] Ulubelen A, Topcu G, Sönmez U, Eris C. Terpenoids from *Salvia nemorosa*. Phytochemistry 1994; 35:1065–7.

[30] Kan Y, Gökbulut A, Kartal M, Konuklugil B, Yilmaz G. Development and validation of a LC method for the anlaysis of phenolic acids in Turkish *Salvia* species. Chromatographia 2007; 66:S147–52.

[31] Radulescu V, Chiliment S, Oprea E. Capillary gas chromatography-mass spectrometry of volatile and semi-volatile compounds of *Salvia officinalis.* Journal of Chromatography A 2004; 1027:121–6.

[32] Lu Y, Yeap Foo L. Polyphenolics of *Salvia*–a review. Phytochemistry 2002; 59:117–40.

[33] Abreu ME, Müller M, Alegre L, Munné-Bosch S. Phenolic diterpene and α-tocopherol contents in leaf extracts of 60 *Salvia* species. J Sci Food Agric 2008; 88:2648–53.

[34] Djarmati Z, Jankov RM, Djordjević A, Ribar B, Lazar D, Engel P. Carnosic acid 12-methyl ether-γ-lactone, a ferruginol-type diterpene from *Salvia officinalis*. Phytochemistry 1992; 31:1307–9.

[35] Fontana G, Savona G, Rodríguez B, Dersch C, Rothman RB, Prisinzano TE. Synthetic studies of neoclerodane diterpenoids from *Salvia splendens* and evaluation of opioid receptor affinity. Tetrahedron 2008; 64:10041–8.

Chapter 2: Theory

2.1 Instrumental Techniques

Chromatography is a separation technique, which is widely used in forensic science for a variety of applications including, but not limited to, the analysis of controlled substances [1]. The goal of any chromatographic technique is to separate the sample mixture into individual compounds [2]. To do this, all chromatographic techniques contain a stationary and a mobile phase. The stationary phase is either a solid or a liquid film, which can have different chemical compositions. There are many polar and non-polar stationary phases that can be used and the selection of the stationary phase is dependent on the nature of compounds to be separated [3, 4]. While the stationary phase is static, the mobile phase, which is composed of either a gas or a liquid, is continuously flowing and carries the sample through the instrument [5]. Compounds in a sample mixture will interact with the stationary and mobile phases to achieve separation. In this research, both gas and liquid chromatography coupled with mass spectrometry was used.

2.1.1 Gas Chromatography-Mass Spectrometry

In modern gas chromatography (GC), which is used to analyze thermally stable and volatile compounds, the stationary phase is a liquid film, while the mobile phase is a pressure-regulated gas, more commonly referred to as the carrier gas [3–5]. In addition to interactions between phases, separation in GC is also based on the boiling points of the compounds in the sample mixture. While solid, liquid, and gaseous samples can be analyzed using GC, typically in controlled substance analysis, the solid sample is dissolved in a suitable solvent, which generates a liquid for analysis. The sample is injected *via* syringe through a rubber septum into the GC

(Figure 2.1), which consists of a heated inlet, an oven containing a column coated with the stationary phase, and a detector, which was a mass spectrometer in this research [1, 3].

Upon injection, the sample is volatilized in the heated inlet, which contains a split valve that can be opened or closed during the injection [5]. If the valve is closed during injection, this is called a splitless injection [4]. Under these circumstances, all of the injected sample enters the column. As a result, splitless injections are used to analyze samples at low concentrations [1]. However, when using a splitless injection, compounds may thermally degrade due to the time they spend in the inlet. Alternatively, if the valve is open during injection, this is called a split injection, in which case, only a portion of the sample enters the column. For example, a split ratio of 50:1 indicates that for every 1 part of sample that enters the column, 50 parts are vented to waste. In particular, the higher molecular weight compounds are lost since they do not volatilize as easily as the lower molecular weight compounds and, hence, this type of injection may not be fully representative of the sample [4]. Split injections are used to analyze highly concentrated or dirty samples.

The carrier gas flows throughout the system at a rate ranging from 0.5 to 2 mL/min. The role of the carrier gas is to carry the volatilized sample onto the column. Common carrier gases include nitrogen, helium, and hydrogen [2, 4, 5]. However, high purity helium is primarily used for gas chromatography-mass spectrometry (GC-MS) analysis because it is light and inert [2, 5]. Therefore, it is efficiently pumped away by the mass spectrometer vacuum system, which enables the direct connection of the gas chromatograph and mass spectrometer.

Ideally, the sample mixture is injected onto the column in as narrow of a band as possible [3, 4]. The column is capillary tubing with a thin coating of the liquid stationary phase on the

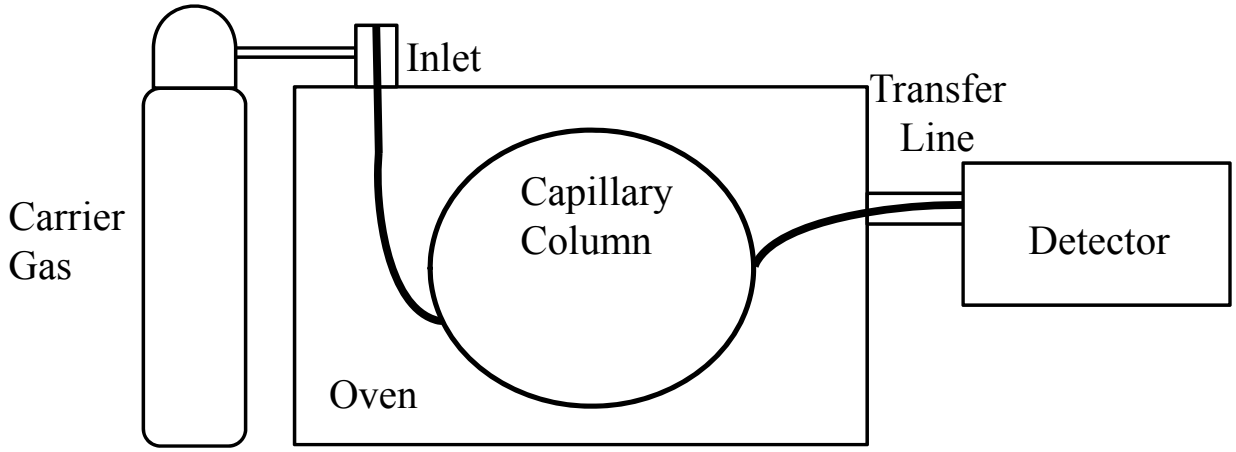

Figure 2.1: Schematic diagram of a gas chromatograph.

inner walls allowing the carrier gas to flow through the center. As previously stated, there are many stationary phases with different chemical compositions that are used for different applications. For example, a non-polar stationary phase composed of a polysiloxane backbone with 5% phenyl and 95% methyl substitutes is useful when separating non-polar compounds such as hydrocarbons [4]. Sample compounds partition between the carrier gas and the stationary phase and, since the carrier gas is inert, the interactions between the compounds in the sample and the stationary phase lead to separation [1, 4, 5]. Possible interactions that may occur include dispersion forces, dipole induction forces, dipole orientation forces, and hydrogen bonding [4]. Therefore, separation is affected by properties of the compounds, such as size, polarizability, and acidity or basicity.

The efficiency of the separation is affected by the specifications of the column, mainly length (typical range 10–30 m), internal diameter (typical range 0.1–0.32 mm), and stationary phase thickness (typical range 0.25 or 0.32 μm) [2, 4, 5]. In an efficient separation, compounds should travel through the column in as narrow a band as possible, resulting in narrow, Gaussian-shaped peaks in the final chromatogram [4].

Narrower internal diameters can also result in a more efficient separation [2]. A capillary column with a narrow internal diameter allows the carrier gas to flow through the column unrestricted without major pressure drops. As a result, longer capillary columns are used, which provide more time for interactions between the compounds and the stationary phase and thus more separation. However, longer time spent in the column leads to greater broadening from molecular diffusion. Molecular diffusion occurs when molecules of the same compound diffuse from regions of high concentration to regions of low concentration in the mobile and stationary phases [4, 5]. Due to molecular diffusion, band broadening increases with time.

A thinner stationary phase can also provide a more efficient separation due to rapid mass transfer [5]. Mass transfer occurs when molecules of a compound move between the mobile and stationary phases [4, 5]. To avoid band broadening, ideally the molecules in the mobile and stationary phases should be in equilibrium with one another. However, since the mobile phase is continuously moving, the system is dynamic, and thus equilibrium must be re-established constantly. The faster this is achieved, the less band broadening occurs; however, sufficient time for separation is still needed. Therefore, there must be a compromise between column length, internal diameter, and thickness of the stationary phase to obtain an efficient separation,

To further control the separation and minimize band broadening, the temperature of the oven can be adjusted. The oven, in which the column is located, can be set to maintain a constant temperature for an entire analysis, which is called an isothermal analysis, or to vary temperature throughout the analysis, which is called temperature programming [3–5]. This type of programming allows for the separation of complex mixtures with a wide range of boiling points. Therefore, temperature programming can decrease analysis time, which limits molecular diffusion, decreases mass transfer time, and minimizes band broadening. However, a decrease in analysis time can result in a loss in resolution and, thus, there must be a compromise between analysis time and resolution.

Many different detectors are available for GC, each of which generates a chromatogram of the sample [2]. This is a plot with retention time, which is the time it takes for a compound to travel through the column and reach the detector, on the x-axis and abundance on the y-axis [1, 3]. Each detected compound is represented by a peak in the chromatogram with a specific abundance at a given retention time.

Despite the number of detectors available, specifically in this research, a mass spectrometer (MS) detector was used [2]. The GC column passes through a transfer line into the mass spectrometer. The transfer line is typically heated to the equivalent of the highest GC oven temperature to prevent the loss of separated compounds due to condensation. The mass spectrometer is maintained under vacuum (approximately 10^{-4} to 10^{-7} Pa) to reduce ion collisions with gas molecules, to increase the mean free path of the ions, and to protect the mass spectrometer from moisture that would cause corrosion and degrade performance [1, 3]. Due to the low flow rates used with capillary GC columns and the low molecular weight of helium, the carrier gas is efficiently pumped away by the MS vacuum system. Therefore, it is possible to directly couple the GC to the MS through the transfer line, avoiding the need for a specialized interface [2].

The mass spectrometer consists of an ionization source, a mass analyzer, and a detector [3]. The source ionizes the separated compounds from the GC column and fragments those ions. The mass analyzer separates the ions according to mass-to-charge (*m/z*) ratio, and the detector converts the ions into an electric signal to be detected.

Many types of ionization methods are available but, in this research, electron ionization (EI) was used. Figure 2.2 shows a schematic diagram of an EI source. During EI, a filament produces an electron beam (typically 70 eV) that is attracted to a positive plate on the opposite side of the source [1, 3]. Due to the presence of a magnet in the source, the electron beam travels in a spiral path, which increases the chance of interaction between the beam and the separated compounds from the GC column. The separated compounds pass by the beam perpendicularly and the electron beam causes the ejection of an electron from the compounds, leaving them with a positive charge [1, 3]. This ion has the same molecular mass as the original compound and is

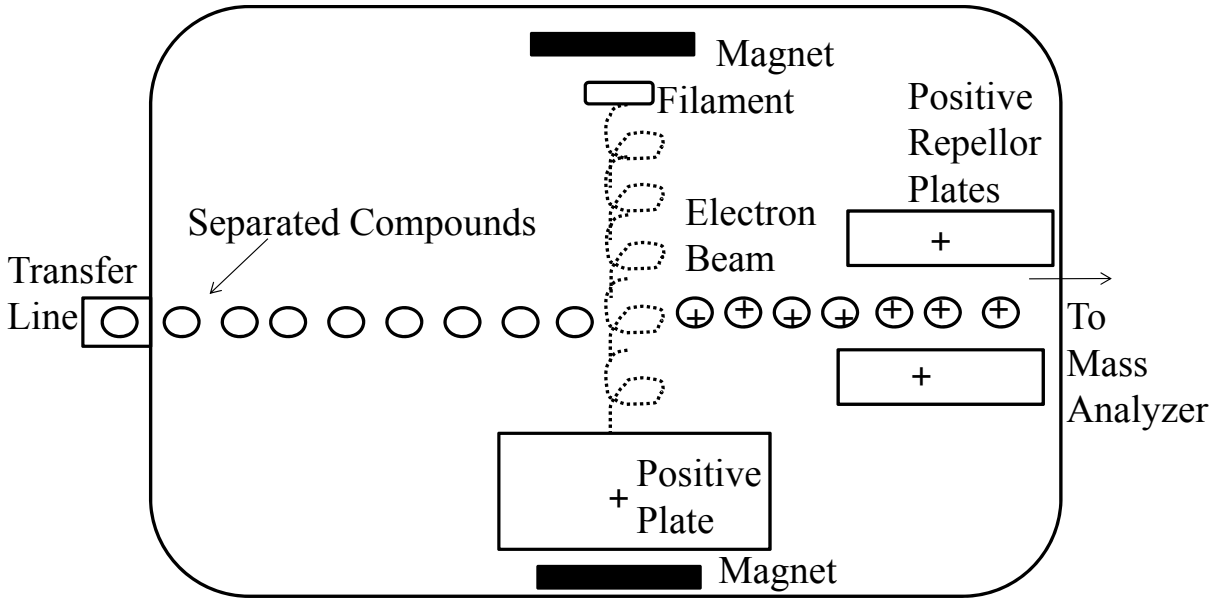

Figure 2.2: Schematic diagram of an electron ionization source adapted from [1].

known as the molecular ion [1]. The amount of energy needed to ionize a compound, known as the ionization energy, varies based on the structure of the compound; however, approximately 5 to 25 eV is a sufficient amount to ionize most compounds [6]. Therefore, the excess energy causes the molecular ion to fragment. The way in which the ion fragments is unique to that ion and will occur in a reproducible manner under the same ionization conditions. Thus, EI is considered a hard ionization method due to the extent of fragmentation that occurs.

The ions formed are then repelled by positive repellor plates into the mass analyzer, which separates the ions based on their *m/z* value [3]. Several different types of mass analyzers are available but, in this research, a quadrupole analyzer was used (Figure 2.3). The quadrupole consists of four rods with opposite rods connected as a pair. Alternating direct current (DC) and radio frequency (RF) voltages are applied to the two pairs of rods. A specific ratio of DC and RF voltages will allow only ions with a narrow range of *m/z* values to pass through the quadrupole unhindered and reach the detector [1, 3]. Ions with a *m/z* value outside this range will hit the rods of the quadrupole, be neutralized, and pumped away by the vacuum system [3]. The quadrupole can operate in full scan mode, during which both the DC and RF voltages are increased while still maintaining the same DC/RF ratio [1]. At each DC/RF voltage, ions of different *m/z* values will pass through the quadrupole to the detector. In this way, a wide range of *m/z* values will reach the detector (*e. g. m/z* 50–550) to obtain a full scan.

In bench-top GC-MS instruments, continuous dynode electron multipliers are commonly used as the detector (Figure 2.4). This detector is composed of a curved glass tube that is coated with a substance (*e.g.* lead oxide) that readily emits secondary electrons [1, 2, 5]. The positively charged ions from the mass analyzer hit the surface of the electron multiplier, causing ejection of electrons from the multiplier surface [1, 3]. A potential gradient exists along the length of the

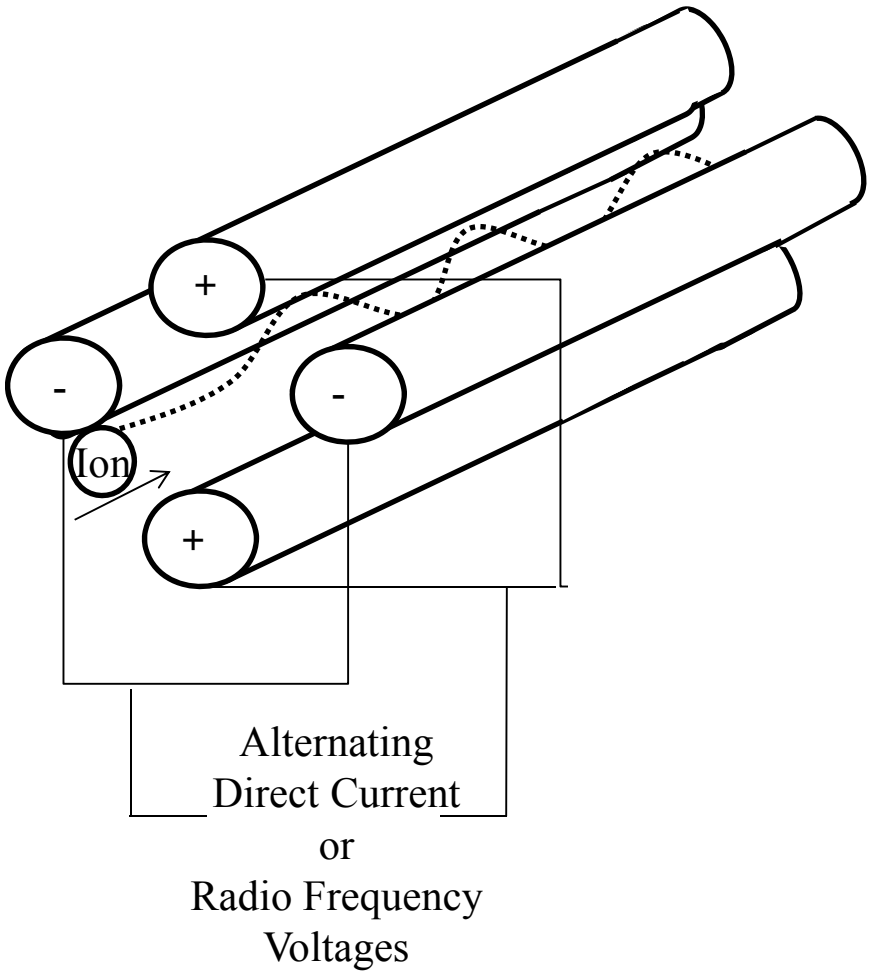

Figure 2.3: Schematic diagram of a quadrupole mass analyzer adapted from [1, 2].

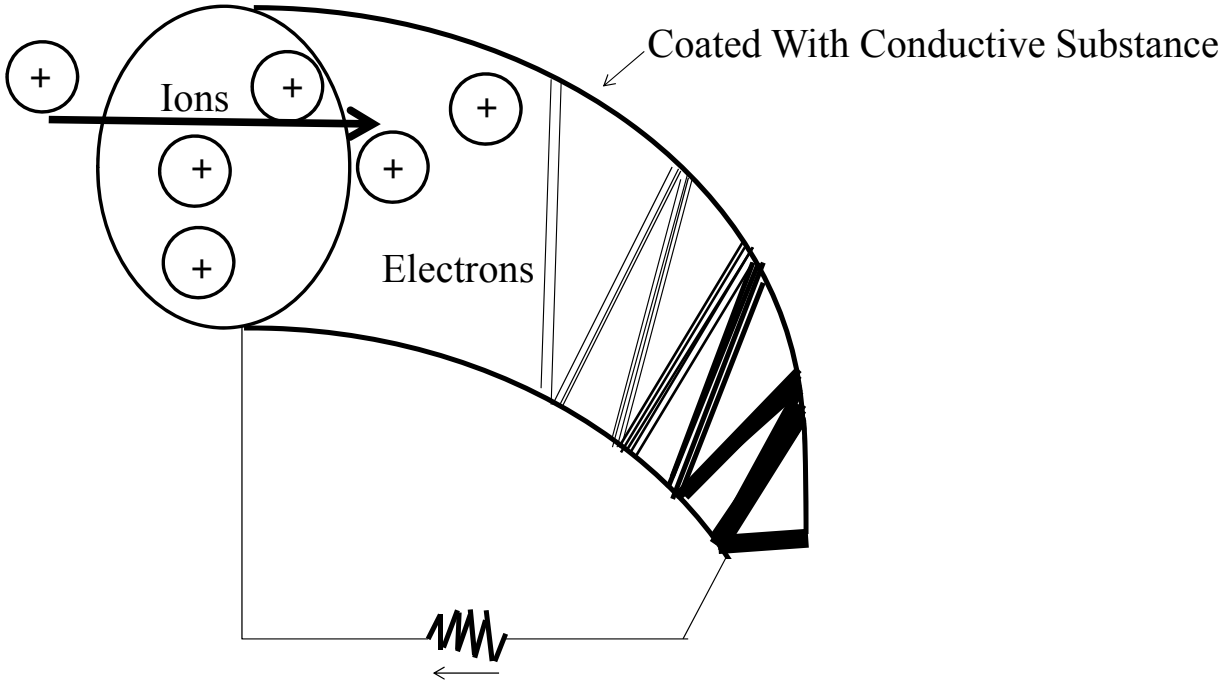

Figure 2.4: Schematic diagram of a continuous dynode electron multiplier adapted from [1].

multiplier, attracting electrons down the multiplier toward the ground potential [1]. As the electrons travel, further collisions with the surface occur, which causes ejection of more electrons [1, 3]. For every n collisions with the surface, the signal is amplified by 2^n. When the cascade of electrons reaches the end of the multiplier, the signal has been amplified by approximately 10^5 to 10^8 [1, 2, 5].

The current is measured and digitized, producing the mass spectrum of the compound. The spectrum is a plot of ion abundance versus *m/z* value. In many cases, the ion with the highest *m/z* value is the molecular ion, which is useful in identifying the compound's molecular mass. Every mass spectrum also has a base peak, which is the most abundant ion. In addition, the masses of the fragment ions are used to determine the structure of the compound. Figure 2.5 shows the mass spectrum of salvinorin A, which is the hallucinogenic compound in *Salvia divinorum*, generated using EI with the molecular ion, base peak, and major fragments labeled [7, 8].

As previously stated, GC can be coupled with several detectors; however, these detectors are not specific for identification purposes since only a chromatogram, and hence retention time information, is obtained for each compound in the sample mixture [2]. Therefore, only retention times between compounds can be compared. In GC-MS, the chromatogram is more correctly called the total ion chromatogram (TIC), since the chromatogram is plotted as the sum of all the ion signals in the sample [3]. Using GC-MS, every compound in the TIC has a given retention time and a mass spectrum; thus, definitive identification of unknown compounds is possible. Therefore, GC-MS is widely used in forensic science to identify unknown substances.

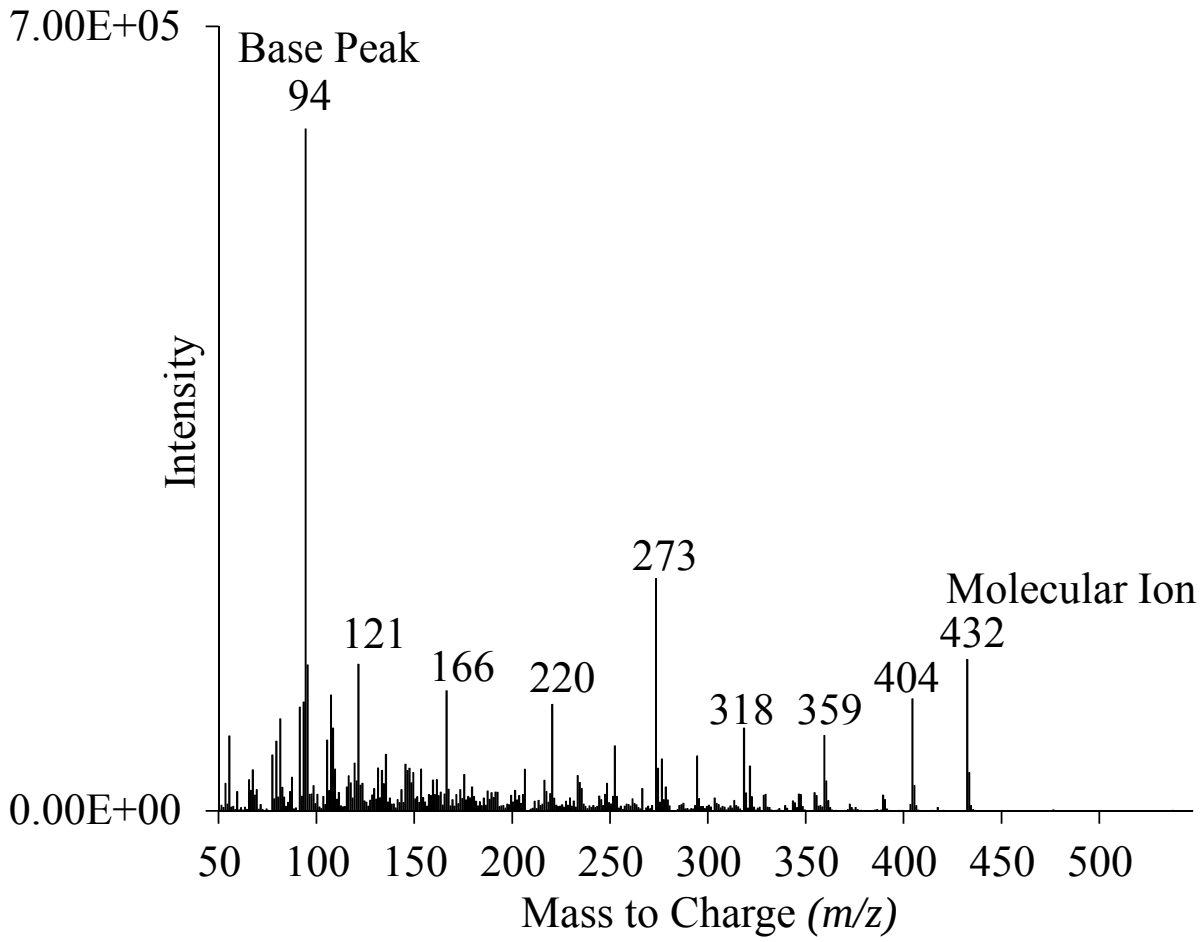

Figure 2.5: Mass spectrum of salvinorin A analyzed using electron ionization.

2.1.2 Liquid Chromatography-Mass Spectrometry and Tandem Mass Spectrometry

Currently, liquid chromatography-mass spectrometry (LC-MS) is not routinely used in forensic science; however, it has application for separation and identification of non-volatile and thermally degradable compounds that may not be detected by GC-MS [4]. A liquid chromatography system (Figure 2.6) consists of a fixed-volume injection loop, a liquid mobile phase, pumps for the mobile phase, a column containing a packed solid stationary phase, and a detector.

Since the sample is injected onto the column as a liquid, samples not in liquid form must first be prepared. Once prepared, the sample is drawn into a syringe and then injected into the fixed-volume injection loop [4, 5]. The injection loop typically contains 1 to 25 μL of sample in the 'load' position; however, it can contain up to 100 μL of sample [2]. The loop is then placed into the 'inject' position, during which the mobile phase enters the loop and carries the sample onto the column [2, 4, 5]. Fixed volume injection loops are advantageous since injection volume is more reproducible, which leads to improved analytical precision [2].

Columns in liquid chromatography are typically 10 to 30 cm in length, 4 to 5 mm in internal diameter, and contain a solid stationary phase with particle sizes ranging from 3 to 5 μm [4, 5]. Since packed columns are used in LC, peak broadening can result not only from molecular diffusion and mass transfer, but also from "multiple path" contribution. This occurs when molecules of the same compound take different paths through the packed particles of the stationary phase [5]. In order to minimize peak broadening due to "multiple path" contribution, a smaller particle size can be used, enabling tighter packing and reducing the number of pathways available. In some LC systems, the column may be used inside an oven [5]. For LC, temperature

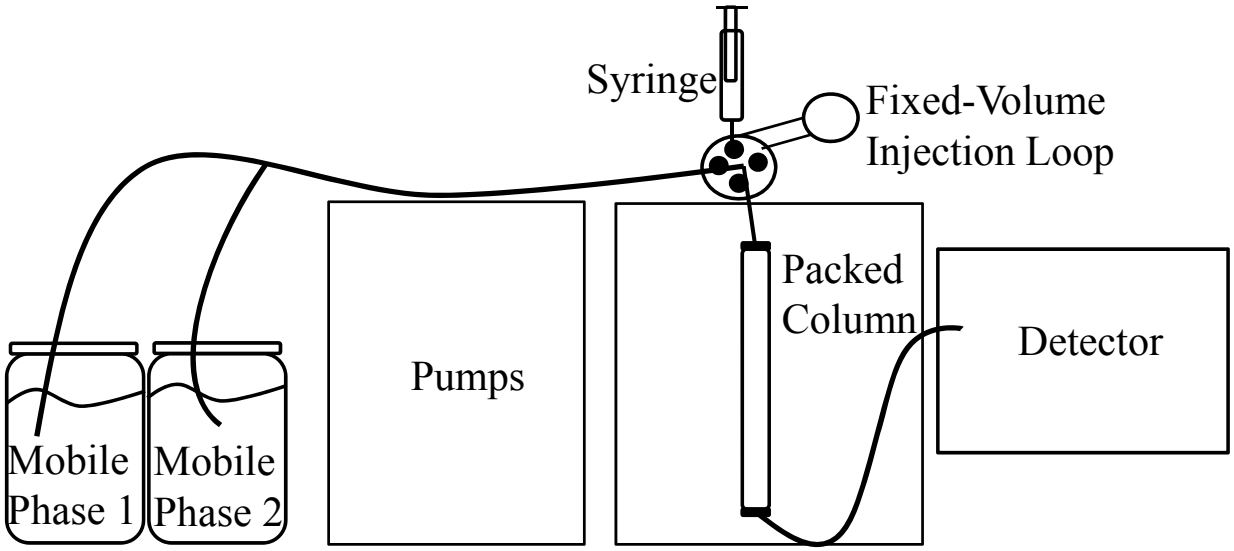

Figure 2.6: Schematic diagram of a liquid chromatograph adapted from [5].

programming is not commonly used; instead, a constant oven temperature is used to prevent temperature variations that could affect retention times [4, 5].

Efficient separation in LC results from interactions between the compounds in the sample, the stationary phase, and the mobile phase [4, 5]. Several types of interactions may occur and various properties of the compounds, stationary phase, and the mobile phase affect separation [4]. However, the separation in LC is ultimately based on polarity and solubility [4, 5]. For example, in normal-phase chromatography, a polar stationary phase and non-polar mobile phase are used to separate compounds, while in reversed phase chromatography, a non-polar stationary phase and a polar mobile phase are used [4, 9]. Reversed phase chromatography is more commonly used for controlled substance analysis and has been shown to separate a variety of compounds such as depressants and pain killers, which makes it useful for forensic science analyses [4, 5].

To further assist with separation, the composition of the mobile phase, which can consist of different solvents with varying polarities, can be programmed to change over time [5]. This is known as solvent programming, which is analogous to temperature programming used in gas chromatography. In solvent programming, the concentration of one solvent is programmed to increase during the analysis while another solvent's concentration decreases. This consequently changes the polarity of the mobile phase. It is common in reversed phase chromatography to decrease the polarity of the mobile phase by decreasing the concentration of the more polar solvent and increasing the concentration of the less polar solvent [4, 9]. In general, this decrease in polarity of the mobile phase will gradually cause the non-polar compounds to elute from the stationary phase after the polar compounds have been separated. Therefore, solvent

programming allows for the efficient separation of compounds with a wide range of polarities and also decreases analysis time [5].

The separated compounds then enter the detector. As for GC, a number of different types of detectors are available for LC; however, in this research, a mass spectrometer was used as the detector [2]. The column effluent first enters a heated capillary interface, which leads into the ionization source [3]. While many ionization methods are available, electrospray ionization (ESI) was used in this research. This is a common ionization method for analysis by LC-MS because ionization occurs in solution at atmospheric pressure, which makes ionization of non-volatile compounds possible [1]. Contrary to EI, ESI is a soft ionization method, during which little fragmentation of the pseudomolecular ion occurs [1, 10]. As a result, the pseudomolecular ion is typically apparent; however, it is difficult to identify compounds since less fragment ions are available for structural elucidation [10].

In ESI, ion formation is either the result of compound interaction with the solvent solution in which it is dissolved or with the mobile phase [1]. ESI can be performed in positive or negative mode, which results in both singly charge ions, such as the pseudomolecular ion, and multiply charge ions. In negative mode, ionization occurs by deprotonation or by forming a negatively charged adduct; therefore, this mode is useful for ionizing acidic compounds [1, 10]. In positive mode, ionization occurs by protonation or by forming a positively charged adduct; therefore, this mode is useful for ionizing basic compounds. In this research, positive ionization was used and will be discussed further.

To couple the LC system to the mass spectrometer, the column effluent first enters a heated capillary interface, which leads into the ionization source [3]. The ESI source (Figure 2.7) consists of a needle and a counter electrode [1, 2, 10]. A potential difference, known as the ion

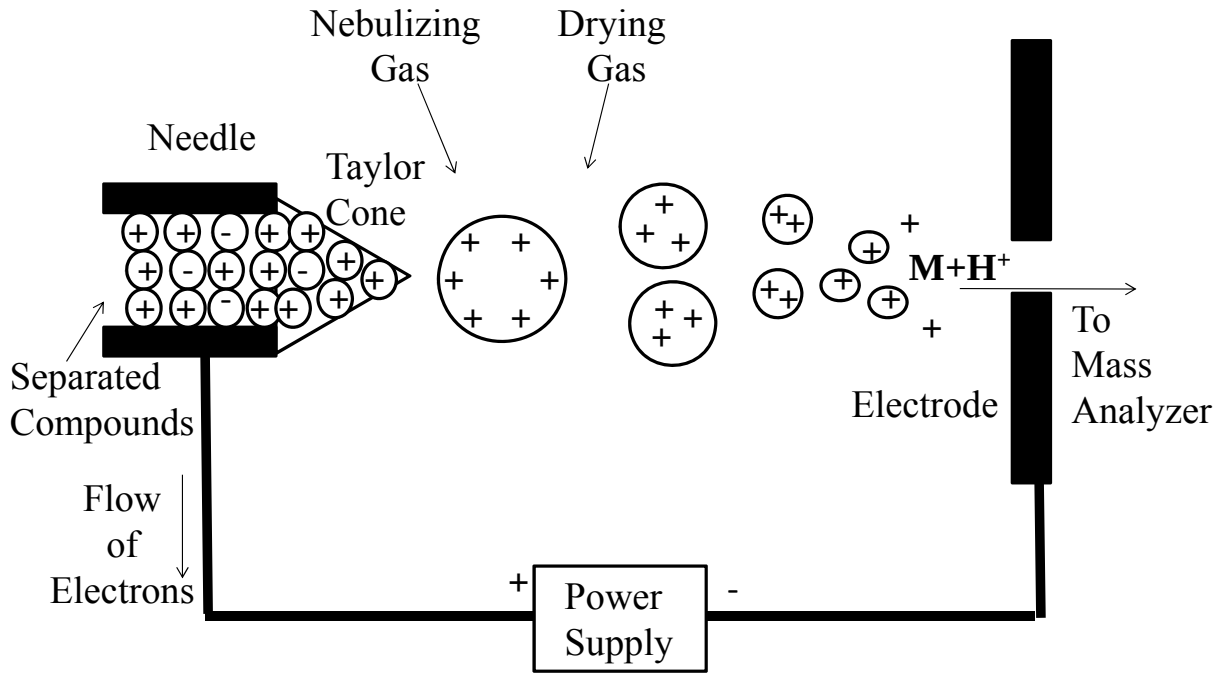

Figure 2.7: Schematic diagram of electrospray ionization source operating in positive ionization mode adapted from [1].

spray voltage, is applied between the two, which generates an electric field [1, 10]. Charge is accumulated as liquid, which consists of the separated compounds and mobile phase, emerges from the needle. Due to the potential difference between the needle and counter electrode, the liquid forms a"Taylor cone" [1, 6]. The source also contains a nebulizing gas, which assists with droplet formation, and a drying gas, which assists with solvent evaporation [1, 10]. As the droplets get smaller, the charge density increases and the ions begin to repel one another [1, 3, 10]. Ions in the gas phase are formed as the charged droplets from the Taylor cone undergo solvent evaporation and Coulomb fissions. The exact mechanism of ionization is unknown; however, it is commonly thought that these Coulomb fissions occur either by the charge residue mechanism or the ion evaporation mechanism [10, 11].

In the charge residue mechanism, repulsion forces overcome surface tension forces [6, 10, 11]. As a result, the droplets become smaller and smaller. This process continues until the droplet contains only one ion. Throughout the mechanism, the drying gas evaporates the solvent and eventually the ion enters the gaseous state. In the ion evaporation mechanism, the electric field on the surface of the droplet overcomes surface tension forces. As a result, the ion escapes the droplet and immediately enters the gaseous state. Then, the ions pass through a capillary into differentially pumped vacuum regions, and are focused into the mass analyzer [1, 2].

In this research, a hybrid triple quadrupole/linear ion trap (Figure 2.8) was used as the mass analyzer. This consists of three quadrupoles (Q_0, Q_1, and Q_2) and a hybrid quadrupole (Q_3), which is a linear ion trap (LIT) [10, 12]. The first quadruple (Q_0) is not used as a mass analyzer, but is used to focus and collect the ions, which ultimately increases the ion signal [1, 9, 10, 12]. The second quadrupole (Q_1) is capable of mass analysis, while the third quadrupole (Q_2)

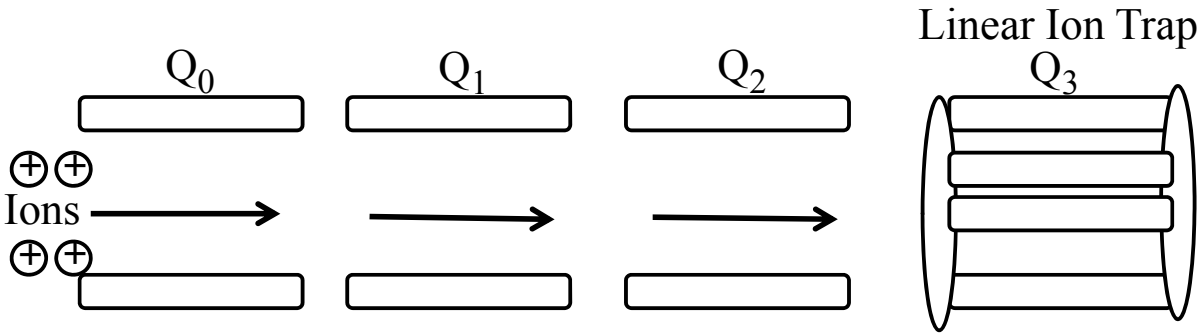

Figure 2.8: Schematic diagram of a hybrid triple quadrupole/linear ion trap adapted from [1].

is used as a collision cell. The LIT is also capable of mass analysis. It consists of four rods, arranged similarly to a quadrupole, with electrodes at each end [1]. Potentials are applied to these electrodes, which allow the LIT to confine ions. Linear ion traps have a greater ion trapping capacity than 3-dimensional ion traps and, therefore, have an improved trapping efficiency, which enhances ion signal. In this research, a dynamic fill time was used, which allows the trap to fill until a certain ion signal is reached as opposed to using a static fill time, which allows the trap to fill for a specified time, irrespective of ion signal achieved during that time [9, 12]. Space-charge effects, where ions are ejected from the trap at the wrong time due to overfilling, and repulsion between ions, are reduced using a dynamic fill time because the trap can be filled several times with smaller mass ranges [1, 12]. Thus, fewer ions are in the trap at one time, which reduces repulsion forces between ions and prevents overfilling. Therefore, using a dynamic fill time can enhance the accuracy of mass spectra and increase the signal [1, 9, 12].

In this research, an enhanced mass spectrum scan was performed, during which all ions are transmitted through the three quadrupoles (Q_0, Q_1, and Q_2) without separation [9, 12]. To do this, RF voltages are applied to these quadrupoles without DC voltages [1, 12]. Then, the ions are trapped and separated by the LIT [12]. To do this, RF voltages are scanned. For a given RF, only ions with a specific *m/z* value are ejected [1, 12]. Therefore, by scanning the voltage, all *m/z* values in the desired range can be ejected. The ejected ions then enter the detector, which was a continuous dynode electron multiplier in this research, similar to that previously described in Section 2.1, to generate the mass spectrum. Figure 2.9 shows the mass spectrum of salvinorin A generated using ESI with the pseudomolecular ion, base peak, and major fragments labeled.

Since ESI is a soft ionization method and doesn't result in substantial fragmentation of the molecular ion, identification of compounds based on structure is more difficult [1]. To

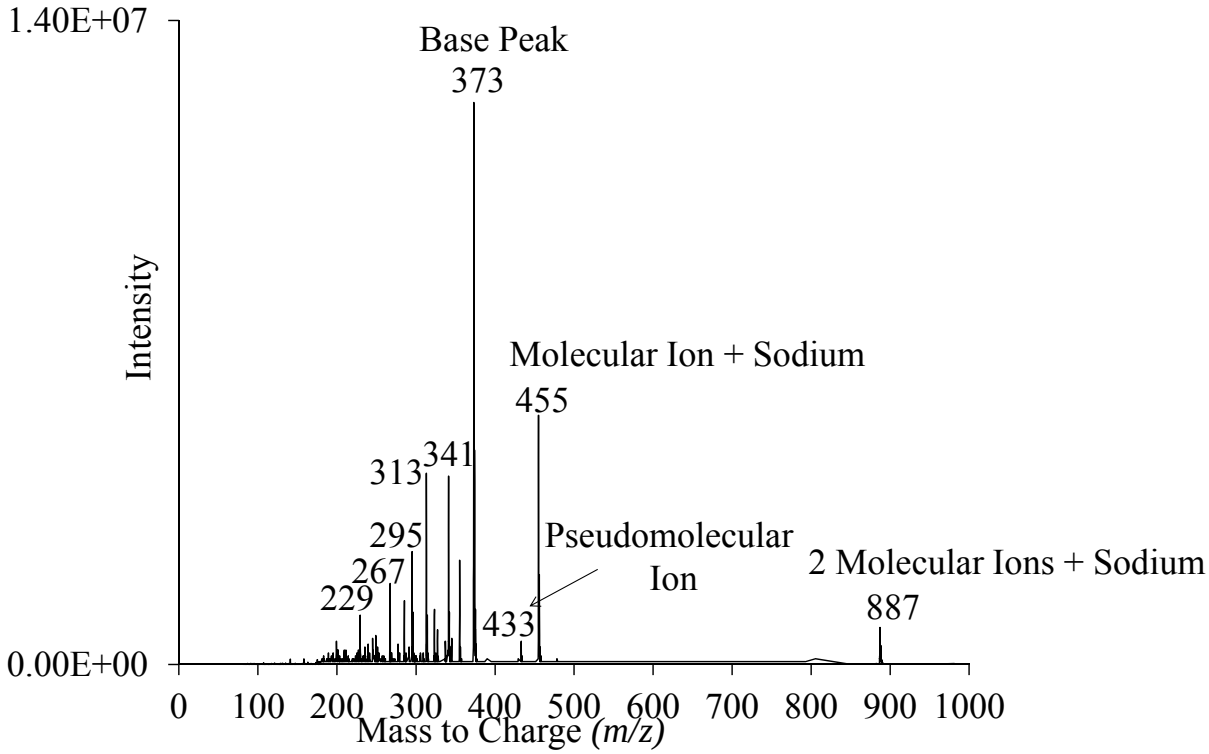

Figure 2.9: Mass spectrum of salvinorin A analyzed using electrospray ionization.

overcome this, tandem mass spectrometry (MS^2) was also performed in this research. In MS^2, either a selected ion or all the ions in the original mass spectrum are fragmented further, therefore, obtaining another mass spectrum. This provides more information about the fragmentation patterns of a compound and assists in the determination of an unknown compound's structure [1, 10]. This is especially useful with LC analyses because generally libraries used for identification purposes are not widely available.

To perform MS^2 in this research, an enhanced product ion scan was conducted after the enhanced mass spectrum scan. For the enhanced product ion scan, Q_0 collects the ions [10]. The second quadrupole (Q_1) is set to a specific DC and RF voltage to allow only ions with a specific *m/z* value to pass [1, 10]. In this research, the LC-MS/MS was programmed to individually select the top two abundant ions from the enhanced mass spectrum scan for each compound [9]. After the ions of a specific compound are selected in Q_1, the ions enter the third quadrupole (Q_2), which acts as a collision cell filled with nitrogen gas [1, 10]. The energy of the collision is transferred into internal energy, which causes the ions to fragment. This collision energy is a controlled parameter and can vary slightly by using a collision energy spread to allow the fragmentation of various ions [10]. The ions and their fragments exit the collision cell and are collected and separated by the LIT, prior to detection [1, 10].

2.2 Data Pretreatment

Chromatographic data collected over a period of time will contain non-chemical sources of variance, which can include rising chromatographic baselines, electronic noise, retention time drift, and differences in abundance due to differences in injection volume and detector response

[13]. Prior to any statistical procedures, data pretreatment should be performed to reduce these non-chemical sources of variance [13-16]. Once data pretreatment procedures are performed, patterns in the data due to actual chemical differences become more recognizable. In this research, four pretreatment procedures were applied to the data: background correction, smoothing, retention time alignment, and normalization. Each procedure is discussed in more detail below.

2.2.1 Background Correction

When analyzing samples by GC-MS with high oven temperatures, the stationary phase can degrade or "bleed" [1, 4]. There are several common bleed ions; however, the identity of these ions will depend on the type of column used. For example, some common column bleed ions for a non-polar stationary phase composed of a polysiloxane backbone include *m/z* 207, 281, and 355 [1]. Column bleed causes chromatograms to have a rising baseline that is not reproducible, even for replicate injections of a sample. To minimize this rising baseline, background correction can be performed [1, 16]. Although there are a number of ways to perform background correction, in this research, the correction was performed using instrument software [16]. Using this method, the mass spectrum of the background ions must first be obtained [16, 17]. This was done by selecting a section of the TIC on the plateau of the rising baseline that did not contain any sample peaks. The ions in the corresponding mass spectrum are then subtracted from the full TIC [17]. This will reduce the rising baseline; however, caution is needed because smaller peaks in the TIC may also be subtracted.

This background correction method is fast, simple, and reproducible. In addition, the

instrument software is available in most forensic laboratories and, thus, this background correction method could be easily incorporated into existing laboratory procedures.

2.2.2 Smoothing

Chromatograms contain electronic noise, which is random short-term fluctuations in detector response [13, 18]. Smoothing algorithms can be applied to reduce this non-chemical source of variance [13]. While several types of smoothing algorithms are available, a Savitzky-Golay smooth was used in this research [13, 18].

The Savitzky-Golay algorithm fits a polynomial to each section of the chromatogram using the least squares method [13, 18, 19]. The number of data points in a section of the chromatogram and order of the polynomial are defined by the user [13]. One factor to consider in determining the number of data points per section and the order of the polynomial is the number of data points within a peak. For each section, the polynomial is fit to the data points [13, 18]. To do this, the algorithm determines the coefficients of the polynomial so that the squared difference between the y values of the polynomial and the data points is minimal [13, 18, 20]. Using this polynomial, the best fit data point for the center of the section is then calculated [13, 18– 20]. This is done by substituting the x value of the original center data point into the polynomial to determine the y value for the new center data point [13, 18]. This new center data point will replace the original. Then, the section is moved one data point further and the process is repeated [18]. Although smoothing can reduce noise, chromatograms can be over smoothed, which can reduce the abundance of compounds [13]. Therefore, a balance must be reached between reducing noise and minimizing peak reduction.

2.2.3 Retention Time Alignment

Retention time drift is a common observation in chromatographic data. Variations in retention time can be due to slight differences in the mobile phase flow rate, mobile phase composition, oven temperature, and changes in the stationary phase composition due to degradation [15, 21]. Different algorithms are available to reduce retention time drift. In this research, a correlation optimized warping (COW) algorithm and a peak matching algorithm were used.

For both algorithms, the sample TICs are aligned to a target TIC, which should include as many peaks as possible from the TICs to be aligned [22]. The target TIC can be selected at random from the data set, generated as the average of all TICs in the data set, or prepared from a sample containing all analyte compounds that is then analyzed to generate the target [14, 21, 22].

For the COW algorithm, the target and the sample TICs are each divided into segments [14, 15, 21]. Two parameters can be adjusted: segment size and warp size [14, 21, 22]. Segment size is the number of data points per segment, while warp size is the maximum number of data points that can be added or subtracted from a segment in order to adjust the alignment of a peak [14, 15, 21, 22]. Each segment in the sample TIC is separately aligned to the target starting at the end of the TIC and the number of data points defined by the warp is added or subtracted [14]. For example, if the warp size is set to 3, then 0, 1, 2, or 3 data points can be added or subtracted from a segment in the sample TIC. Local correlation coefficients are calculated between the segment in the sample TICs and the target TIC for each warp [14, 21]. The warp giving the highest local correlation coefficient is used, although all coefficients for all warp sizes in that segment are also stored. The algorithm moves to the next segment to repeat the procedure. After aligning each segment, a global correlation coefficient is calculated, which considers all

combinations of the local correlation coefficients for each segment [14]. The combination of warps that give the highest global correlation coefficient is considered the optimal alignment [21].

For the peak matching algorithm, noise and peak identification thresholds are first defined [23]. The noise is defined as the standard deviation of a range of data points toward the beginning or end of the TICs. In this research, the noise was defined as the standard deviation of 88 data points towards the end of the TICs. The peak identification threshold, which is a threshold used to define a peak, is calculated by multiplying the noise by a constant. In this research, two times the noise was used. Any signal above this value was identified as a peak.

The algorithm then finds the peaks in the target TIC [23]. This is done by calculating the first derivative at each data point in the TIC. The leading edge of a peak is identified when the difference in abundance between consecutive data points is greater than the peak identification threshold. The algorithm then searches for the zero crossing in the first derivative, which represents the peak maximum, while the next zero crossing represents the tailing edge of that peak. The retention time corresponding to the zero crossing is interpolated. This retention time, representing the apex of the peak, is added to a list containing retention times of all peaks identified in the target. The algorithm is then applied to each sample TIC, which generates tables of peaks identified in each.

Once peaks are identified, the retention times of identified peaks in the sample TICs are compared to those identified in the target TIC [23]. If a peak is present in the target and sample TIC with retention times within a defined window size, which is a specific number of data points, then the peaks are considered a “match.” The algorithm then interpolates the retention

time axis to align the peak maxima. This procedure is repeated for each sample TIC in order to align the data.

2.2.4 Normalization

Slight variation in injection volume and detector response can lead to differences in abundance of the same compounds in replicate injections [13]. To account for these differences, normalization can be performed prior to data analysis. While several types of normalization are available, total area and internal standard normalization were used in this research. For total area normalization, the total area of each TIC is first calculated [13, 24]. Then, the average total area of all TICs in the data set is calculated. For each TIC, the abundance at each retention time is divided by the total area of that TIC, and then multiplied by the average area of all TICs in the data set.

For internal standard normalization, the peak abundance of the internal standard in each TIC is obtained [13]. Then, the average peak abundance of the internal standard in all TICs in the data set is calculated. For each TIC, the abundance at each retention time is divided by the internal standard abundance for that TIC, and then multiplied by the average internal standard abundance of all TICs in the data set.

2.3 Data Analysis

2.3.1 Pearson Product Moment Correlation Coefficients

Pearson product moment correlation (PPMC) coefficients are used to make pair-wise comparisons between two samples [25]. The coefficients are calculated using equation 2.1,

$$r = \frac{\sum\{(x_i - \bar{x})(y_i - \bar{y})\}}{\{[\sum(x_i - \bar{x})^2][\sum(y_i - \bar{y})^2]\}^{1/2}} \quad (2.1)$$

where r is the correlation coefficient, x_i and y_i represent the individual data points for two different samples being compared, and $\bar{x}$ and $\bar{y}$ are the average data from the samples. In this research, PPMC coefficients were calculated between chromatograms of the *Salvia* species. Therefore, x_i and y_i represent the abundance of data point *i* in each chromatogram (*x* and *y*), while $\bar{x}$ and $\bar{y}$ are the average abundance of all data points in each chromatogram. The coefficient ranges from 0 to ±1, with the sign indicating a positive or negative correlation. A positive correlation means the data points in both chromatograms increase or decrease at the same time, while a negative correlation means the data points of one chromatogram increase while the data points of the other decrease [26]. A coefficient of 0 indicates no correlation between the two samples being compared, while a coefficient of 1 indicates a perfect correlation. Coefficients between ±0.8 to ±1 indicate strong correlation, ±0.5 to ±0.79 indicate moderate correlation, and less than ±0.5 indicate weak correlation [27].

2.3.2 Principal Components Analysis

Principal components analysis (PCA) is an unsupervised, multivariate statistical procedure used to reduce the dimensionality of the data set [13, 16, 26, 28, 29]. Therefore, PCA simplifies the data and reveals underlying patterns in the data.

To perform PCA, the data are first mean-centered to set the average of the data set to zero [16, 26, 29]. To do this, the average of the data points across all samples is calculated and then the average is subtracted from each data point in each sample [26]. Next, the covariance between each pair of samples is calculated using equation 2.2,

$$Cov(X,Y) = \frac{\sum_{i=1}^{n}\{(x_i-\bar{x})(y_i-\bar{y})\}}{(n-1)} \quad (2.2)$$

where X and Y are different samples, $\bar{x}$ and $\bar{y}$ are the averaged data from the samples, and n is the number of samples [26]. The covariance for all pair-wise comparisons is displayed as a covariance matrix that is symmetrical around the diagonal [26, 30].

The eigenvectors and eigenvalues of the covariance matrix are then calculated [26, 29]. An eigenvector is a unit vector that can be multiplied by a matrix and results in a product that is a multiple of the unit vector [26]. The multiple of the eigenvector is called the eigenvalue. For PCA, each eigenvector represents a principal component, which describes the variance among the samples [16, 26]. Each successive eigenvector is perpendicular to the next [26]. The first principal component (PC1) describes the most variance in the data set, while the second principal component (PC2) describes the next greatest amount of variance. The amount of variance described is given by the eigenvalue for that eigenvector, in proportion to all other eigenvalues [13, 16, 26].

The outputs from PCA are scores and loadings plots [16]. The score for a sample on PC1 is calculated by multiplying the PC1 eigenvector by the mean-centered data for that sample, then summing the products [26]. This procedure can be repeated using the PC2 eigenvector to calculate the score for the sample on PC2. The scores of each sample, calculated in a similar manner, are then plotted to generate the scores plot, which typically shows PC1 on the x-axis and PC2 on the y-axis [16, 26, 28, 30]. The positioning of the samples on the scores plot can be used to associate and discriminate samples based on their relative location [16, 30]. Samples that are closely clustered in the scores plot are chemically similar, while those samples separated in the scores plot are chemically different [13, 16, 28].

The loading plots can be used to identify the most variable chemical compounds among the samples for each principal component [16]. The loadings plots for chromatographic data can

be generated by graphing the eigenvector for each principal component versus retention time. Compounds can either load positively or negatively in these plots and the height of each compound in the plots corresponds to its contribution to the positioning of the samples in the scores plot [13, 16, 28]. Since the loadings plots consist of the eigenvector for each PC plotted against retention time, it is possible to identify the compounds in the samples that are contributing to the variance described by each PC [16].

2.3.3 Student's t-tests

Student's *t*-tests are hypothesis tests used to determine if there is a statistically significant difference between two sample populations [25]. The null hypothesis (H_0) states that the two sample populations are equal, while the alternative hypothesis (H_1) states they are significantly different. In this research, Student's *t*-tests were calculated on the mean scores of each species to determine if *Salvia divinorum* could be statistically differentiated from four other *Salvia* species. All variances were assumed to unequal. Therefore, the Student's *t*-test for unequal variances was used (equation 2.3),

$$t_{calculated} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}} \tag{2.3}$$

where $\bar{x}_1$ and $\bar{x}_2$ are the average of sample populations 1 and 2, respectively, s_1 and s_2 are the standard deviation of sample populations 1 and 2, respectively, and n_1 and n_2 are the number of measurements in sample populations 1 and 2, respectively [25]. For unequal variance, the degrees of freedom (υ) is calculated using equation 2.4,

$$\upsilon = \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\left(\frac{s_1^4}{n_1^2(n_1-1)}+\frac{s_2^4}{n_2^2(n_2-1)}\right)} \quad (2.4)$$

The degrees of freedom is used to determine the $t_{critical}$ value from statistical tables at the desired confidence level [31]. When determining $t_{critical}$, it must be decided if a one-tailed or two-tailed *t*-test is necessary. A one-tailed *t*-test is used to determine whether one sample population is significantly greater or significantly less than the other [25]. A two-tailed *t*-test is used to determine whether the sample populations are significantly different, irrespective of one sample population being greater or less than the other. $T_{critical}$ from the table for the desired tailed *t*-test, confidence level, and degrees of freedom is then compared to $t_{calculated}$ from equation 2.3 [25, 31]. If $t_{calculated}$ is less than $t_{critical}$, then H_0 is accepted and the two sample populations are equal [25]. If $t_{calculated}$ is greater than $t_{critical}$, then H_0 is rejected, H_1 is accepted, and the two sample populations are significantly different from one another.

References

References

[1] Watson JT, Sparkman OD. Introduction to mass spectrometry: instrumentation, applications and strategies for data interpretation. 4th ed. West Sussex, England: John Wiley & Sons Ltd, 2007.

[2] Skoog DA, Holler FJ, Crouch SR. Principles of instrumental analysis. 6th ed. Belmont, California: Thomson Higher Education, 2007.

[3] McMaster M, McMaster C. GC/MS: A practical user's guide. New York: Wiley-VCH, 1998.

[4] Heftmann E, editor. Chromatography part a: fundamentals and techniques. 6th ed. Amsterdam, The Netherlands: Elsevier B. V., 2004.

[5] Willard HH, Merritt Jr. LL, Dean JA, Settle Jr. FA. Instrumental methods of analysis. 7th ed. Belmont, California: Wadsworth Publishing Company, 1988.

[6] Reid GE. Ionization methods: How are ions formed? Class notes from CEM 832. Spring 2010.

[7] Tsujikawa K, Kuwayama K, Miyaguchi H, Kanamori T, Iwata YT, Yoshida T, Inoue H. Determination of salvinorin A and salvinorin B in *Salvia divinorum*-related products circulated in Japan. Forensic Sci Int 2008; 180:105–9.

[8] Grilli M, Neri E, Zappettini S, Massa F, Bisio A, Romussi G, et al. Salvinorin A exerts opposite presynaptic controls on neurotransmitter exocytosis from mouse brain nerve terminals. Neuropharmacology 2009; 57:523–30.

[9] Jones AD. LC-MS training at Michigan State University Mass Spectrometry Facility. 2010.

[10] Jones AD. Background and theory of LC/MS/MS on QTRAP mass spectrometers. Powerpoint from Michigan State University Mass Spectrometry Training Facility.

[11] Gabelica V, De Pauw E. Internal energy and fragmentation of ions produced in electrospray sources. Mass Spectrometry Reviews 2005; 24:566–87.

[12] Jones AD. Personal Communication. December 2011.

[13] Blackledge RD, editor. Forensic analysis on the cutting edge: new methods for trace evidence analysis. New Jersey: John Wiley & Sons, 2007.

[14] Tomasi G, Berg FV, Andersson C. Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data. Journal of Chemometrics 2004; 18:231–41.

[15] Malmquist G, Danielsson R. Alignment of chromatographic profiles for principal component analysis: a prerequisite for fingerprinting methods. Journal of Chromatography A 1994; 687:71–88.

[16] Brereton RG. Applied chemometrics for scientists. West Sussex, England: John Wiley & Sons, 2007.

[17] MSD ChemStation Software version E.01.01.335 Agilent Technologies, Inc, 1987-2007.

[18] Savitzky A, Golay MJE. Smoothing and differentiation of data by simplified least squares procedures. Analytical Chemistry 1964; 36:1627–39.

[19] Gorry PA. General least-squares smoothing and differentiation by the convolution (Savitzky-Golay) method. Analytical Chemistry 1990; 62:570–3.

[20] Madden HH. Comments on the Savitzky-Golay convolution method for least-squares fit smoothing and differentiation of digital data. Analytical Chemistry 1978; 50:1383–6.

[21] Nielsen NPV, Carstensen JM, Smedsgaard J. Aligning of single and multiple wavelength chromatographic profiles for chemometric data analysis using correlation optimized warping. Journal of Chromatography A 1998; 805:17–35.

[22] Skov T, van den Berg F, Tomasi G, Bro R. Automated alignment of chromatographic data. Journal of Chemometrics 2006; 20:484–97.

[23] Johnson KJ, Wright BW, Jarman KH, Synovec RE. High-speed peak matching algorithm for retention time alignment of gas chromatographic data for chemometric analysis. Journal of Chromatography A 2003; 996:141–55.

[24] Rietjens M. Reduction of error propagation due to normalization: effect of error propagation and closure on spurious correlations. Analytica Chimica Acta 1995; 316:205–15.

[25] Miller JN, Miller JC. Statistics and chemometrics for analytical chemistry. 4th ed. Harlow, England: Prentice-Hall, 2000.

[26] Smith LI. A tutorial on principal components analysis. 2002. http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf

[27] Devore JL. Probability and statistics for engineering and the science. 4th ed. Duxbury Press, Belmont, CA, 1995.

[28] Bodnar Willard MA. Forensic analysis of *Salvia divinorum* and related *Salvia* species using chemometric procedures. MS Thesis (2010).

[29] Shlens J. A tutorial on principal component analysis: version 2. Systems Neurobiology Laboratory, Salk Insitute for Biological Studies. La Jolla, California 2005. http://www.snl.salk.edu/~shlens/pca.pdf

[30] Lee JLS, Gilmore IS, Seah MP. Proposed terminology for multivariate analysis in surface chemical analysis – vocabulary – part 1: general terms and terms for the spectroscopies. National Physical Laboratory. Teddington, Middlesex, UK. 2008. http://www.npl.co.uk/upload/pdf/multivariate_analysis.pdf

[31] Adam C. Essential mathematics and statistics for forensic science. West Sussex, England: John Wiley & Sons Ltd, 2010.

Chapter 3 Association and Discrimination of *Salvia divinorum* from Related *Salvia* Species Based on Chemical Fingerprints Generated Using Gas Chromatography-Mass Spectrometry

3.1 Introduction

This chapter describes the association and discrimination of *S. divinorum* from related *Salvia* species based on their volatile chemical compounds. *S. divinorum* and four other *Salvia* species were extracted in dichloromethane and extracts were analyzed by gas chromatography-mass spectrometry (GC-MS). The resulting total ion chromatograms (TICs) were background corrected, truncated, smoothed, retention time aligned, and normalized. Mean Pearson product moment correlation (PPMC) coefficients were then used to evaluate the precision of the extraction and analysis procedures, and to evaluate the association of *S. divinorum* extracts and the discrimination of *S. divinorum* from the other four *Salvia* species. Principal components analysis (PCA) was used to assess the discrimination of *S. divinorum* from the other four *Salvia* species in the scores plot. The PCA loadings plots were used to identify the compounds contributing most to the variance among the species. Finally, Student's *t*-tests were used to statistically evaluate the association and discrimination in the PCA scores plot.

3.2 Material and Methods

3.2.1 Extraction of Salvia Species

Leaves of *S. divinorum* were obtained from the research greenhouse at Michigan State University (MSU). *S. guaranitica, S. nemorosa,* and *S. splendens* were purchased from Christian's greenhouse in Haslett, Michigan. *S. officinalis* was purchased from Van Atta's Greenhouse and Flower Shop in Haslett, Michigan. Plants were transported to the MSU research

greenhouse, where they were transplanted as needed. Leaves were collected from each species, briefly rinsed with water to remove any dirt, and placed in a salad spinner (OXO, Chambersburg, PA) to remove the excess water. All leaves were dried in the same manner, one species at a time, with a food dehydrator (NESCO American Harvest model ED-75PR, Two Rivers, WI) at 35 °C. All food dehydrator shelves and liners were cleaned between drying each species. *S. guaranitica, S. nemorosa,* and *S. splendens* were dried for 12 hours, while *S. divinorum* and *S. officinalis* were dried for 16 hours and 36 hours, respectively. Leaves were then stored in sealed bags at 1 °C until analysis.

All glassware used for the extraction was acid-washed following standard procedures. Approximately 0.04 g dried leaves were placed in a 100 mL round bottom flask and 13 mL dichloromethane (J. T. Baker, Phillipsburg, NJ) were added. The flask was placed on a rotary agitator (Büchi Rotavapor-R, Flawil, Switzerland) for 16 hours at medium speed. The extract was then filtered using the experimental set-up in Figure 3.1. The watch glass was placed on top of the Büchner funnel containing the nylon mesh filter (Small Parts Inc., Miami Lakes, FL) to generate a stronger vacuum. The rotary flask was then rinsed with 2 mL dichloromethane (J. T. Baker), the rinsings were filtered, and the entire extract was evaporated to dryness using nitrogen gas (Airgas, Independence, OH) and a water circulating bath (NESLAB endocal RTE-9B, Newington, NH) at 35 °C. This procedure was repeated in triplicate for each of the *Salvia* species, using fresh leaf material for each extraction. All extracts were stored at 2 °C until analysis.

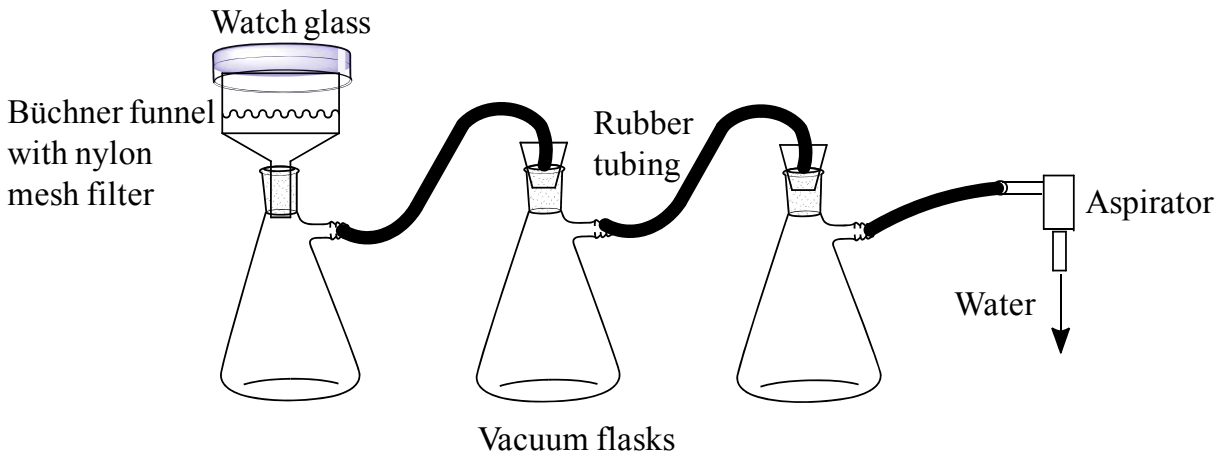

Figure 3.1: Schematic diagram of the filtration system.

3.2.2 Gas Chromatography-Mass Spectrometry Analysis

All extracts were reconstituted in 0.58 mL dichloromethane (J. T. Baker) and progesterone (Sigma Chemical Company, St. Louis, MO) was added as an internal standard (0.00119 M). Extracts were shaken, vortexed, and each was divided into separate autosampler vials (Restek, Bellefonte, PA) with inserts. This allowed for replicate analysis of each extract while minimizing evaporative loss due to puncturing the septum of the vial. Extracts were analyzed using an Agilent 6890N Network GC system (Agilent Technologies, Santa Clara, CA) with a 5973 inert mass selective detector (Agilent Technologies), equipped with ChemStation© software (version E 01.01.335). Extracts were injected via an Agilent 7683 Series autosampler (Agilent Technologies) using a 1 µL injection volume and a split ratio of 50:1. A J&W DB-5MS column (30 m x 0.25 mm x 0.25 µm, Agilent Technologies) was used. The inlet and mass transfer line temperatures were maintained at 300 °C. The carrier gas was ultrahigh purity helium (Airgas, Independence, OH), with a nominal flow rate of 1.0 mL/min. The oven temperature program was as follows: 80 °C for 2 minutes, then 10 °C/min to 340 °C, hold for 8 minutes. Electron ionization (70 eV) was used with a source temperature of 230 °C. The mass analyzer was a single quadrupole and was operated in full scan mode over the range *m/z* 50–550. Spectra of compounds were then collected using ChemStation© and provisional identification of the major compounds in each extract was made by comparison with the NIST MS Search 2.0 library (The NIST Mass Spectral Search Program, Version 2.0 d, Gaithersburg, MD) and various literature sources [1–7].

3.2.3 Data Pretreatment

Each extract was analyzed in triplicate, which resulted in a total of nine TICs for each *Salvia* species. All TICs were pretreated prior to data analysis. As a result of using high oven temperatures, chromatograms had a rising baseline due to column bleed. This would cause problems in subsequent PCA because differences in the baseline abundance would be identified as the major source of variance among the extracts. Therefore, this rising baseline was reduced by subtracting the background using the ChemStation© software. To do this, the mass spectrum at retention time 35 minutes was selected to represent the background spectrum. This was repeated for each TIC. The selected background was then subtracted from each TIC. To further reduce the rising baseline, TICs were truncated at approximately 30.6 minutes since no compounds were identified past this time in any of the species. Truncated TICs were then smoothed to reduce electronic noise. A Savitzky-Golay smooth was performed in OriginPro 7.5 (Version 7.5, OriginLab, Guangzhou, China) using a second order polynomial and a window size of 5 points.

To correct for retention time drift due to instrumental variability, TICs were aligned to an average target chromatogram, which was created using one TIC of each *Salvia* species. A random number generator in Microsoft Excel (Versions 2007 and 2010, Microsoft Corporation, Redmond, WA) was used to select the TICs to create the target. The five TICs were imported into Microsoft Excel and the average abundance at each retention time was calculated to generate the target chromatogram. Total ion chromatograms were firstly aligned to the target chromatogram using a correlation optimized warping (COW) algorithm (LineUp, Version 3.0, InfoMetrix, Inc., Bothell, WA) with a warp of 5 data points and segment size of 65 data points. Then, the TICs were aligned again, using a peak match algorithm available in the literature,

which was performed in Matlab (Version 7.11.0.584 R2010b, The Mathworks, Natick, MA) [8]. For this algorithm, the noise was defined as the standard deviation of 88 data points in the retention time range 29.8 to 30.3 minutes, a peak was defined as two times the standard deviation of the noise, and a window size of 2 was used.

Following alignment, the TICs were normalized in a two-step procedure in Microsoft Excel. Firstly, TICs of each species were subjected to total area normalization. To do this, the area under each TIC was determined, as was the average area of the TICs of that species. Then, the abundance at each retention time in a TIC was divided by the area of that TIC and multiplied by the average area of the TICs of the same species. Following total area normalization, the TICs were subjected to internal standard normalization. To do this, the abundance at each retention time was divided by the abundance of the internal standard in the TIC, and then multiplied by the average abundance of the internal standard in all TICs of that species.

3.2.4 Data Analysis

PPMC coefficients were calculated in Matlab on pretreated TICs for all pair-wise comparisons of replicate injections of the same extract and replicates of each species to assess the precision of the extraction and analysis procedures. Principal components analysis was then performed on the pretreated data using Matlab. Scores and loadings plots were generated in Microsoft Excel. To statistically assess discrimination of *S. divinorum* from the four *Salvia* species, Student's *t*-tests were performed in Microsoft Excel on the mean scores of each species on both principal components (PCs) 1 and 2 at various confidence levels. To further assess association and discrimination of *S. divinorum* from the four *Salvia* species, PPMC coefficients were also calculated on TICs for all pair-wise comparisons of the species.

3.3 Results and Discussion

3.3.1 Provisional Identification of Compounds Present in Salvia Species

Representative TICs of the five *Salvia* species are shown in Figures 3.2–3.6. Provisional identification of compounds present in each species, along with molecular masses and major fragment ions, are listed in Appendix A. *S. nemorosa* and *S. officinialis* belong to *Salvia* clade I, while *S. divinorum, S. guaranitica,* and *S. splendens* belong to *Salvia* clade II [2, 9, 10]. Species in different clades contained common compounds; therefore, classification of the species into clades based on the volatile compounds was difficult. For example, all five *Salvia* species contained hentriacontane, tetratriacontane, and sitosterol [4]. *S. guaranitica, S. nemorosa, S. officinalis,* and *S. splendens* all contained vitamin E. Squalene, nonacosane, and heptacosane were also present in two or more species.

While there were many common compounds, certain compounds were present in only one species and, hence, could be used to distinguish species. For example, *S. divinorum* was the only species to contain salvinorins A, B, D, an acetate compound, and pentatriacontane [1, 2]. *S. officinalis* contained many early eluting compounds, such as thujone, camphor, and epimanool, while the other species had few or no early eluting compounds [11]. *S. guaranitica* was the only species to contain a phthalide. *S. nemorosa* was the only species to contain β-caryophyllene and *S. splendens* was the only species to contain cardenolide. Although these five *Salvia* species could be differentiated based on visual assessment of the TICs, there are over 900 *Salvia* species, some of which may be more difficult to distinguish from *S. divinorum* based only on visual assessment of TICs. Therefore, statistics were applied to this research as a proof-of-concept study.

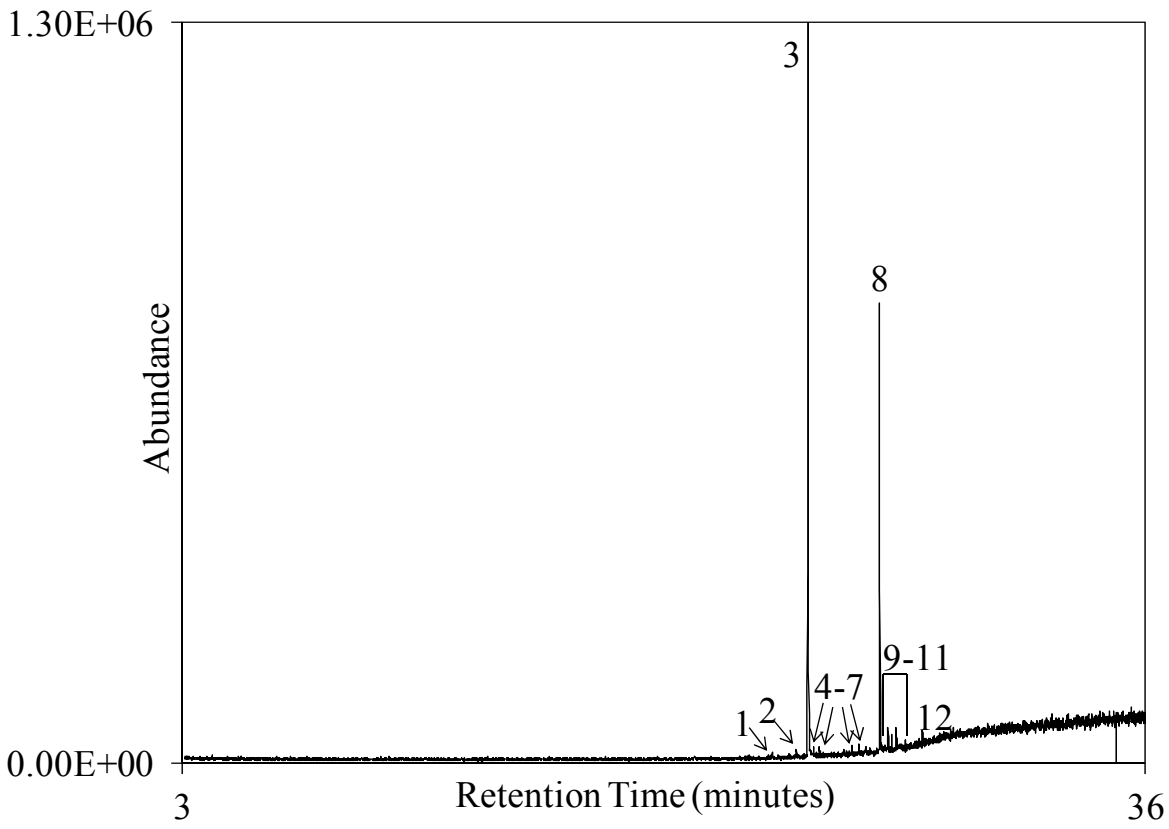

Figure 3.2: Total ion chromatogram of *Salvia divinorum* extracted in dichloromethane and analyzed using gas chromatography-mass spectrometry. Provisional identifications based on library searches and literature sources are as follows: 1=heptacosane, 2=squalene, 3=progesterone, 4=nonacosane, 5=an acetate, 6=hentriacontane, 7=salvinorin B, 8=salvinorin A, 9=tetratriacontane, 10=salvinorin D, 11=sitosterol, and 12=pentatriacontane [1, 2]. See Appendix A, Table A.1 for more detailed information.

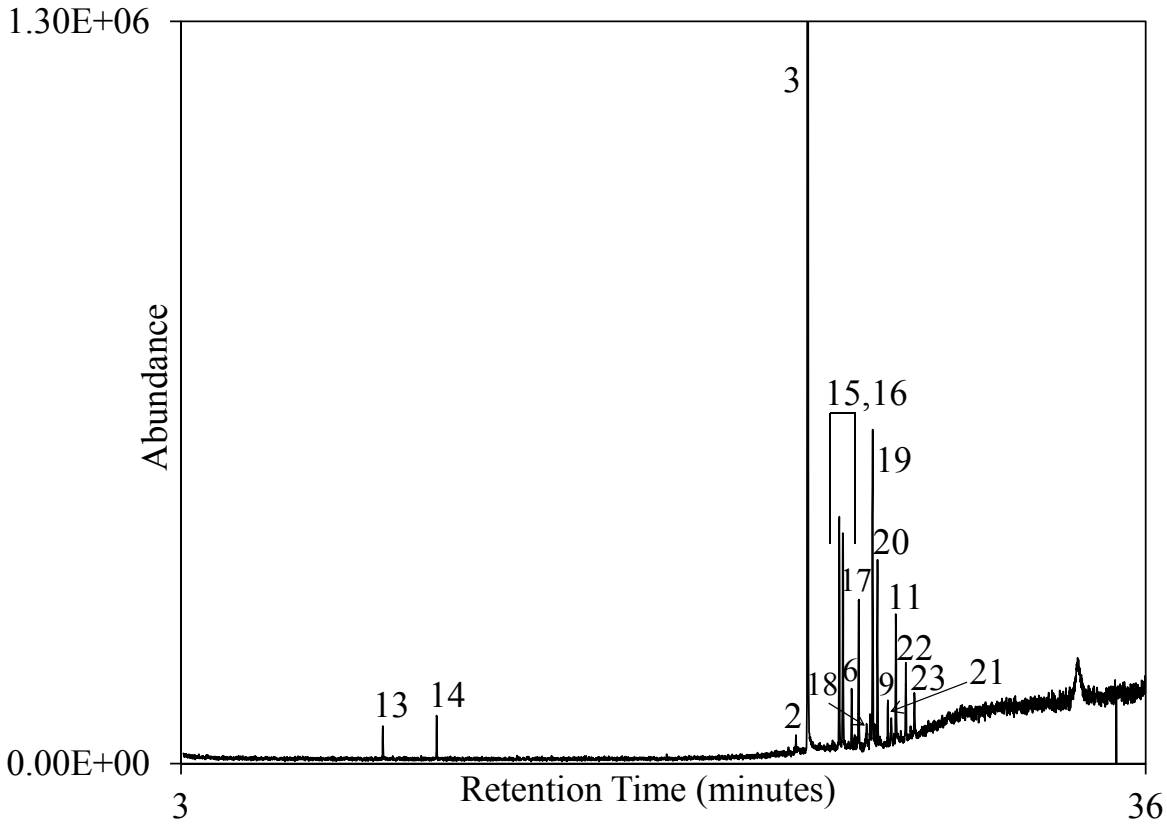

Figure 3.3: Total ion chromatogram of *Salvia guaranitica* extracted in dichloromethane and analyzed using gas chromatography-mass spectrometry. Provisional identifications based on library searches and literature sources are as follows: 2= squalene, 3=progesterone, 6= hentriacontane, 9=tetratriacontane, 11=sitosterol, 13=phthalide, 14=β-cubebene, 15= an ester, 16= an alkene, 17=vitamin E, 18=dione, 19=dione, 20=nitrolepidine, 21=benzenesulfonamide, 22=β-amyrin, and 23=α-amyrin [5, 6]. See Appendix A, Table A.2 for more detailed information.

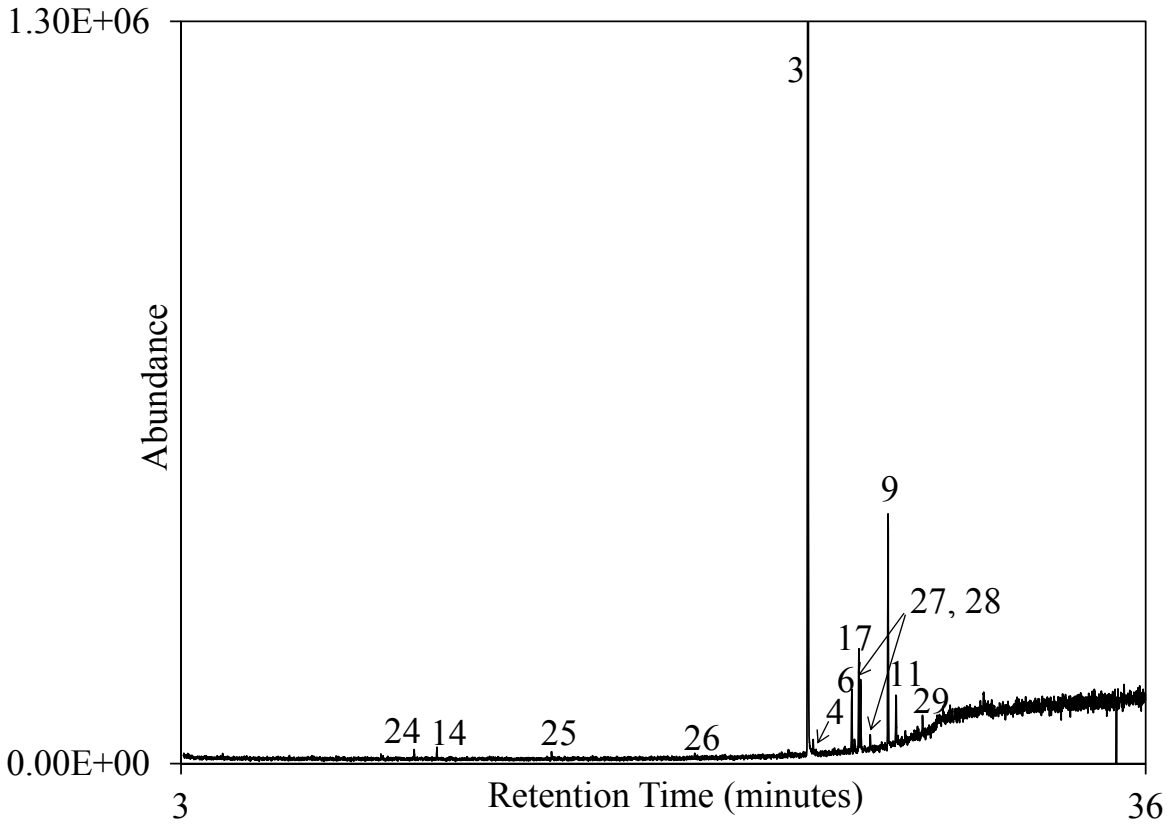

Figure 3.4: Total ion chromatogram of *Salvia nemorosa* extracted in dichloromethane and analyzed using gas chromatography-mass spectrometry. Provisional identifications based on library searches and literature sources are as follows: 3=progesterone, 4=nonacosane, 6=hentriacontane, 9=tetratriacontane, 11=sitosterol, 14=β-cubebene, 17=vitamin E, 24=β-caryophyllene, 25=alkene, 26=amide, 27=phenyl, 28=11-n-decyltetracosane, and 29=hextriacontane [4]. See Appendix A, Table A.3 for more detailed information.

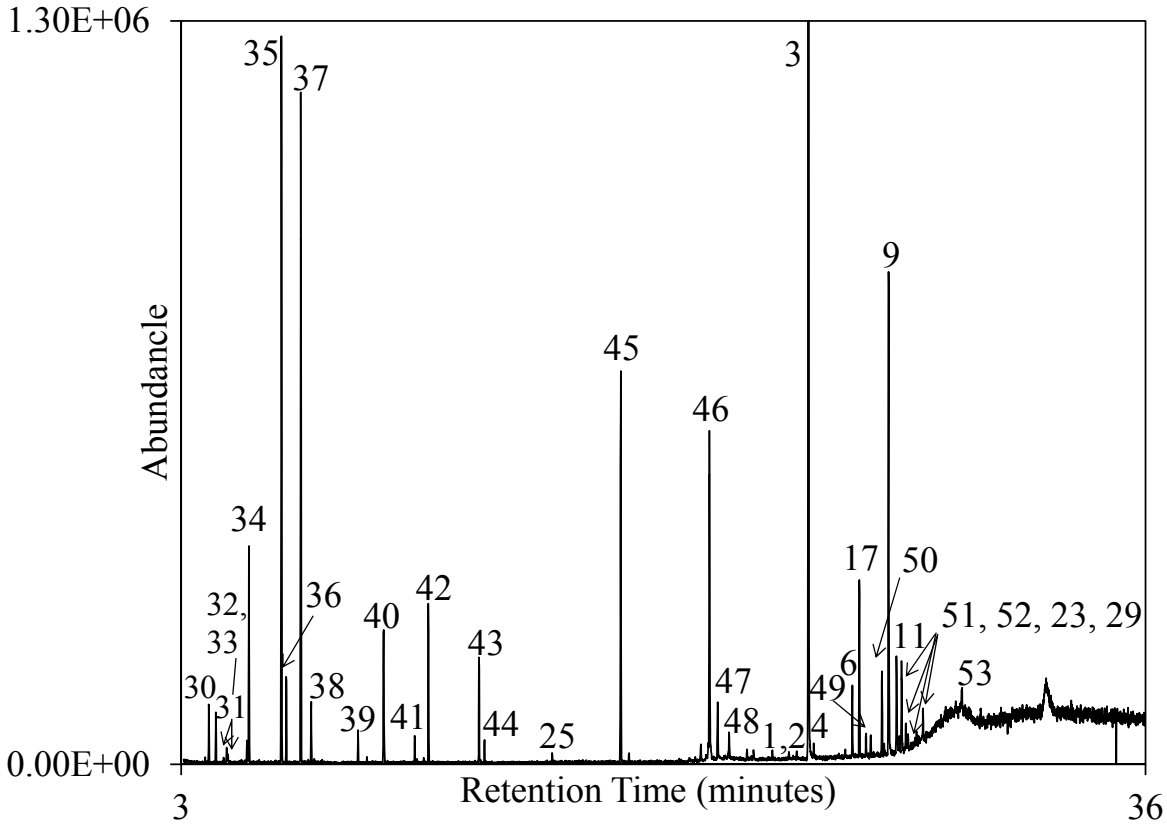

Figure 3.5: Total ion chromatogram of *Salvia officinalis* extracted in dichloromethane and analyzed using gas chromatography-mass spectrometry. Provisional identifications based on library searches and literature sources are as follows: 1=heptacosane, 2=squalene, 3=progesterone, 4=nonacosane, 6=hentriacontane, 9=tetratriacontane, 11=sitosterol, 17=vitamin E, 23=α-amyrin, 25=alkene, 29=hexatriacontane, 30=α-pinene, 31=camphene, 32=β-pinene, 33=β-myrcene, 34=eucalyptol, 35=thujone, 36=thujone, 37=camphor, 38=borneol, 39=bornyl acetate and/or isoborynyl acetate, 40=bicyclic compound, 41=a cyclopropane compound, 42=α-caryophyllene, 43=epiglobulol, 44=diene, 45=epimanool, 46=isocarnosol, 47=phenyl, 48=dione, 49=squalane, 50=triacontane, 51=dotriacontane, 52=9-n-octylhexacosane, and 53=oleanolic acid [3]. See Appendix A, Table A.4 for more detailed information.

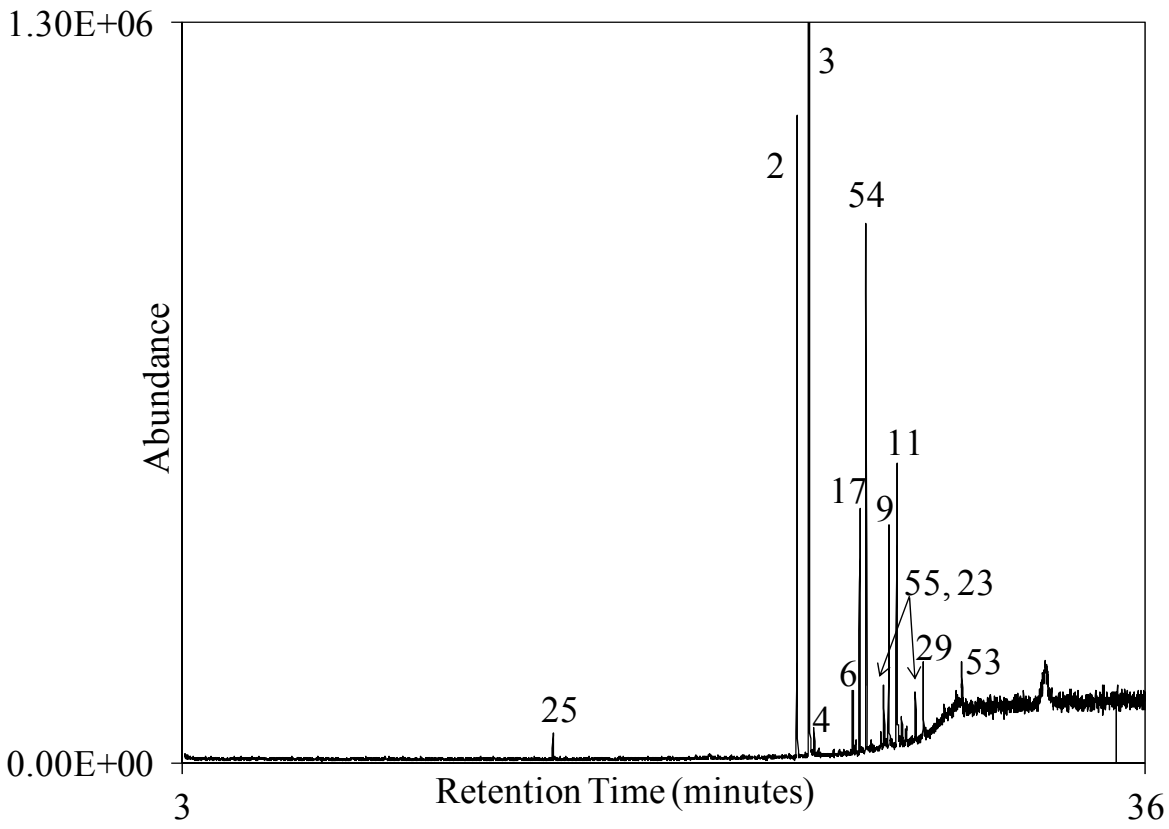

Figure 3.6: Total ion chromatogram of *Salvia splendens* extracted in dichloromethane and analyzed using gas chromatography-mass spectrometry. Provisional identifications based on library searches and literature sources are as follows: 2=squalene, 3=progesterone, 4=nonacosane, 6=hentriacontane, 9=tetratriacontane, 11=sitosterol, 17=vitamin E, 23=α-amyrin, 25=alkene, 29=hextriacontane, 53=oleanolic acid, 54=8-episalviarin, and 55=cardenolide [7]. See Appendix A, Table A.5 for more detailed information.

3.3.2 Investigation of Extraction and Analysis Precision Using Pearson Product Moment Correlation Coefficients

The mean, standard deviation, and range of PPMC coefficients were calculated for replicate injections (n=3) of each extract to demonstrate precision in the instrumental analysis procedure. However, since data pretreatment was performed prior to calculating PPMC coefficients, precision of these procedures was also reflected in the coefficients.

Ideally, replicates are expected to have a PPMC coefficient of 1.00. With the exception of one extract of *S. nemorosa*, all mean PPMC coefficients for injection replicates and the ranges of coefficients were greater than 0.9, which indicated strong correlation (Table 3.6). The lower mean PPMC coefficient (0.8583 ± 0.1202) and the range of coefficients (0.7366 – 0.9769) for one extract of *S. nemorosa* was due to retention time misalignment of the internal standard peak among replicates. All standard deviations were relatively low (≤ 0.06) except for the afore-mentioned extract of *S. nemorosa*. Thus, overall, there was acceptable precision in the instrumental analysis and data pretreatment procedures.

To demonstrate precision in the extraction procedure, in addition to the instrumental analysis and data pretreatment procedures, the mean, standard deviation, and range of PPMC coefficients were also calculated for all extracts (n=36) of each species. The mean PPMC coefficients were greater than 0.9, which indicated strong correlation between extracts of the same species (Table 3.7). Lower coefficients were expected since different leaf material was used for each extraction. The ranges of coefficients were greater than 0.9 with the exception of *S. nemorosa,* which had a broader range of coefficients (0.7366 – 0.9981) due to the retention time misalignment discussed previously. All standard deviation were relatively low; however, *S. nemorosa* had a higher standard deviation (0.0563) than the replicates of the other species, again

Table 3.6: Mean, standard deviation, and range of Pearson product moment correlation (PPMC) coefficients for replicate injections of the same extract

Mean PPMC coefficient ± standard deviation				Range of PPMC coefficients		
Species	Extract 1 (n=3)	Extract 2 (n=3)	Extract 3 (n=3)	Extract 1 (n=3)	Extract 2 (n=3)	Extract 3 (n=3)
S. divinorum	0.9550 ± 0.0304	0.9886 ± 0.0071	0.9953 ± 0.0033	0.9279 – 0.9879	0.9830 – 0.9966	0.9922 – 0.9988
S. guaranitica	0.9968 ± 0.0013	0.9782 ± 0.0176	0.9872 ± 0.0102	0.9959 – 0.9983	0.9659 – 0.9985	0.9767 – 0.9971
S. nemorosa	0.8583 ± 0.1202	0.9958 ± 0.0018	0.9890 ± 0.0062	0.7366 – 0.9769	0.9937 – 0.9969	0.9819 – 0.9934
S. officinalis	0.9766 ± 0.0060	0.9869 ± 0.0053	0.9881 ± 0.0035	0.9728 – 0.9834	0.9811 – 0.9912	0.9853 – 0.9920
S. splendens	0.9869 ± 0.0097	0.9844 ± 0.0113	0.9873 ± 0.0084	0.9770 – 0.9962	0.9715 – 0.9930	0.9776 – 0.9926

Table 3.7: Mean, standard deviation, and range of Pearson product moment correlation (PPMC) coefficients for all extracts for each species

Species	Mean PPMC coefficient ± standard deviation (n=36)	Range of PPMC coefficients (n=36)
S. divinorum	0.9696 ± 0.0243	0.9156 – 0.9988
S. guaranitica	0.9828 ± 0.0163	0.9368 – 0.9985
S. nemorosa	0.9575 ± 0.0563	0.7366 – 0.9981
S. officinalis	0.9703 ± 0.0152	0.9370 – 0.9920
S. splendens	0.9874 ± 0.0086	0.9670 – 0.9971

due to the retention time misalignment. Overall, the high PPMC coefficients indicated acceptable precision in the extraction, instrumental analysis, and data pretreatment procedures.

3.3.3 Association and Discrimination of S. divinorum from Other Salvia Species

The scores plot (Figure 3.8) of the first and second principal components (PC1 and PC2, respectively) accounted for 77% of the total variance among the five *Salvia* species. *S. officinalis* was the only species positioned positively on PC1. On PC2, *S. splendens* was positioned positively, *S. divinorum* was positioned negatively, and *S. guaranitica, S. nemorosa,* and *S. officinalis* were positioned around zero. Replicates of *S. splendens*, *S. divinorum*, and *S. guaranitica* were clustered closely; however, replicates of *S. officinalis* were spread in PC1 and PC2, while replicates of *S. nemorosa* were spread in PC2.

The positioning of each species, and the spread observed among replicates, can be explained with reference to the loadings plots for PC1 and PC2, which contain the compounds contributing most to the variance among the five species (Figure 3.9).

The majority of the compounds weighted positively on PC1 in the loadings plot (*e.g.*, α-pinene, camphene, eucalyptol, thujone, camphor, epiglobulol, epimanool, isocarnosol) were compounds present in *S. officinalis* (Figure 3.9A). While some compounds present in *S. officinalis* were weighted negatively on this PC, their contribution was minimal in comparison to the positively- weighted compounds. Therefore, *S. officinalis* was positioned positively on PC1 in the scores plot. Spread observed among replicates of *S. officinalis* on PC1 in the scores plot was primarily due to differences in the abundance of the more volatile compounds, thujone, camphor, and epimanool, which were weighted positively on this PC.

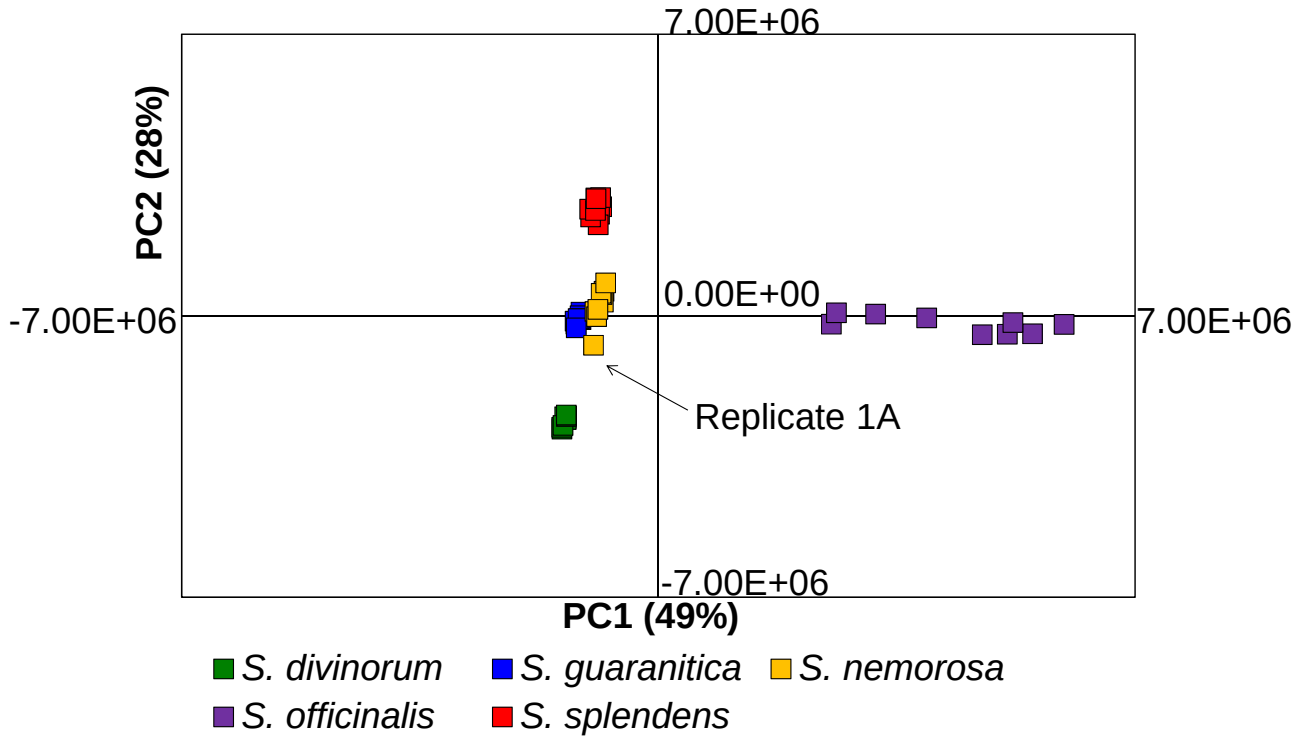

Figure 3.8: Scores plot of principal component 1 versus principal component 2 showing discrimination of *S. divinorum* from the other four *Salvia* species based on chemical profiles obtained by gas chromatography-mass spectrometry.

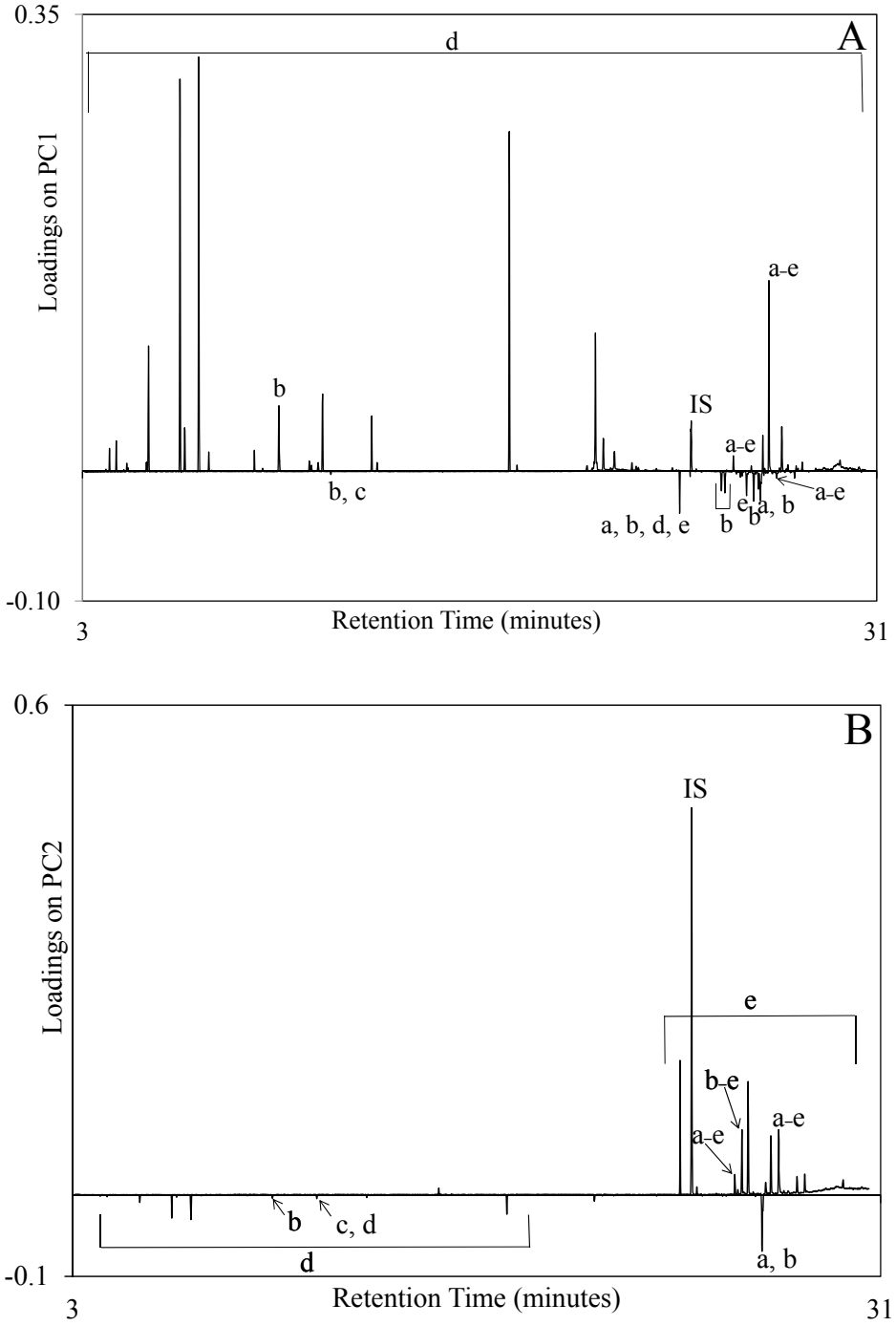

Figure 3.9: Loading plots for (A) principal component 1 (PC1) and (B) principal component 2 (PC2) based on principal components analysis of the five *Salvia* species analyzed by gas chromatography-mass spectrometry. Selected compounds are labeled based on species as follows: a=*Salvia divinorum,* b=*Salvia guaranitica,* c=*Salvia nemorosa,* d=*Salvia officinialis*, e=*Salvia splendens*. IS=internal standard. See Appendix A for provisional identifications based on library searches and literature sources.

S. officinalis was positioned around zero on PC2 in the scores plot: two replicates of this species were positioned slightly positively, while the other seven replicates were positioned slightly negatively. This difference in positioning was due to differences in abundance of compounds such as hentriacontane, vitamin E, sitosterol, which were weighted positively on PC2, and thujone and camphor, which were weighted negatively on PC2. For example, the replicate of *S. officinialis* positioned most positively on PC2 in the scores plot contained the highest abundance of hentriacontane, vitamin E, and sitosterol, and the lowest abundance of thujone and camphor, while the replicate positioned most negatively on PC2 contained the lowest abundance of hentriacontane, vitamin E, and sitosterol.

In addition to the above problems, a retention time misalignment of the internal standard peak in the TICs also contributed to the positioning and spread among *S. officinalis* replicates on PC2 of the scores plot. Figure 3.10 shows this misalignment among the TICs of nine replicates of *S. officinalis* and one replicate of *S. nemorosa*. The effect of the misalignment was observed when the TICs were mean centered, which was necessary for PCA. Figure 3.11A shows the mean-centered data for one replicate of *S. officinalis.* Only one peak should be apparent for each compound; however, due to the misalignment of the internal standard previously mentioned, two peaks were observed corresponding to the internal standard. One peak had a retention time (t_R) of 24.44 min, while the other peak was at t_R 24.46 min.

In order to mean center a data set, the average abundance at each retention time for all sample TICs was calculated. This average was then subtracted from the abundance at each retention time in each TIC to generate the mean-centered data. The replicate of *S. officinalis* illustrated in Figure 3.11 contained the highest abundance of internal standard among the replicates of this species at 24.44 min and, therefore, when the average abundance of all TICs

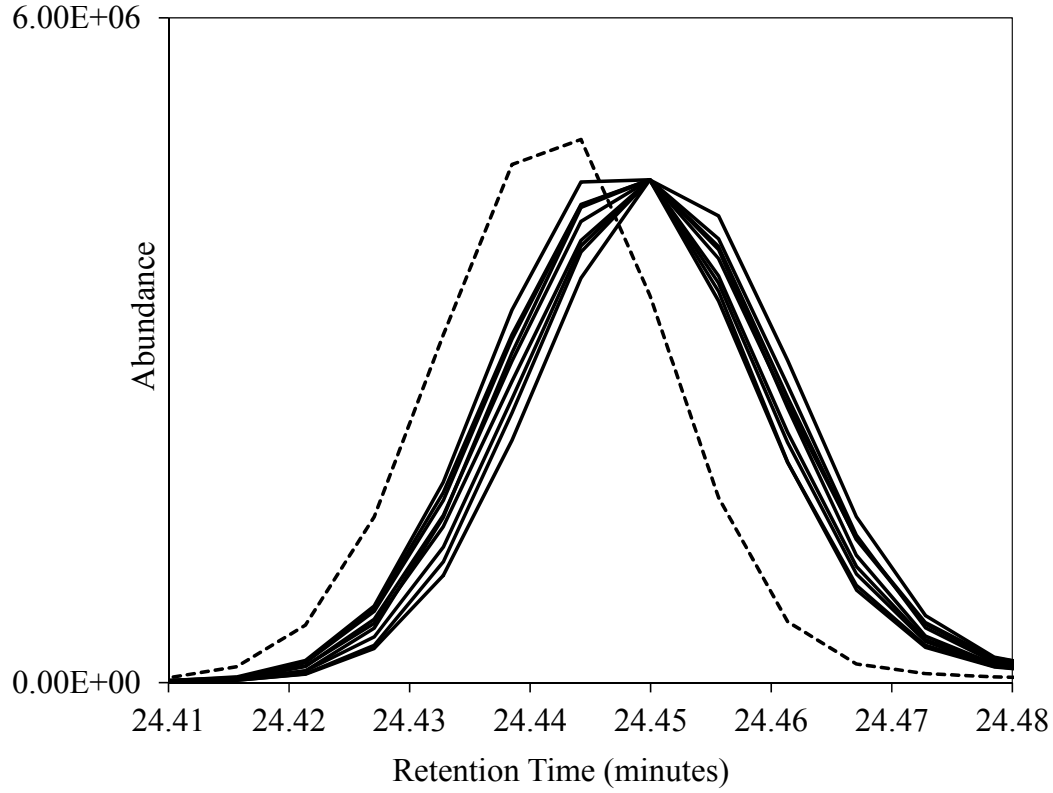

Figure 3.10: Magnified version of nine total ion chromatograms of *S. officinalis* (solid line) and one total ion chromatogram of *S. nemorosa* (dashed line) showing the misalignment of the internal standard peak.

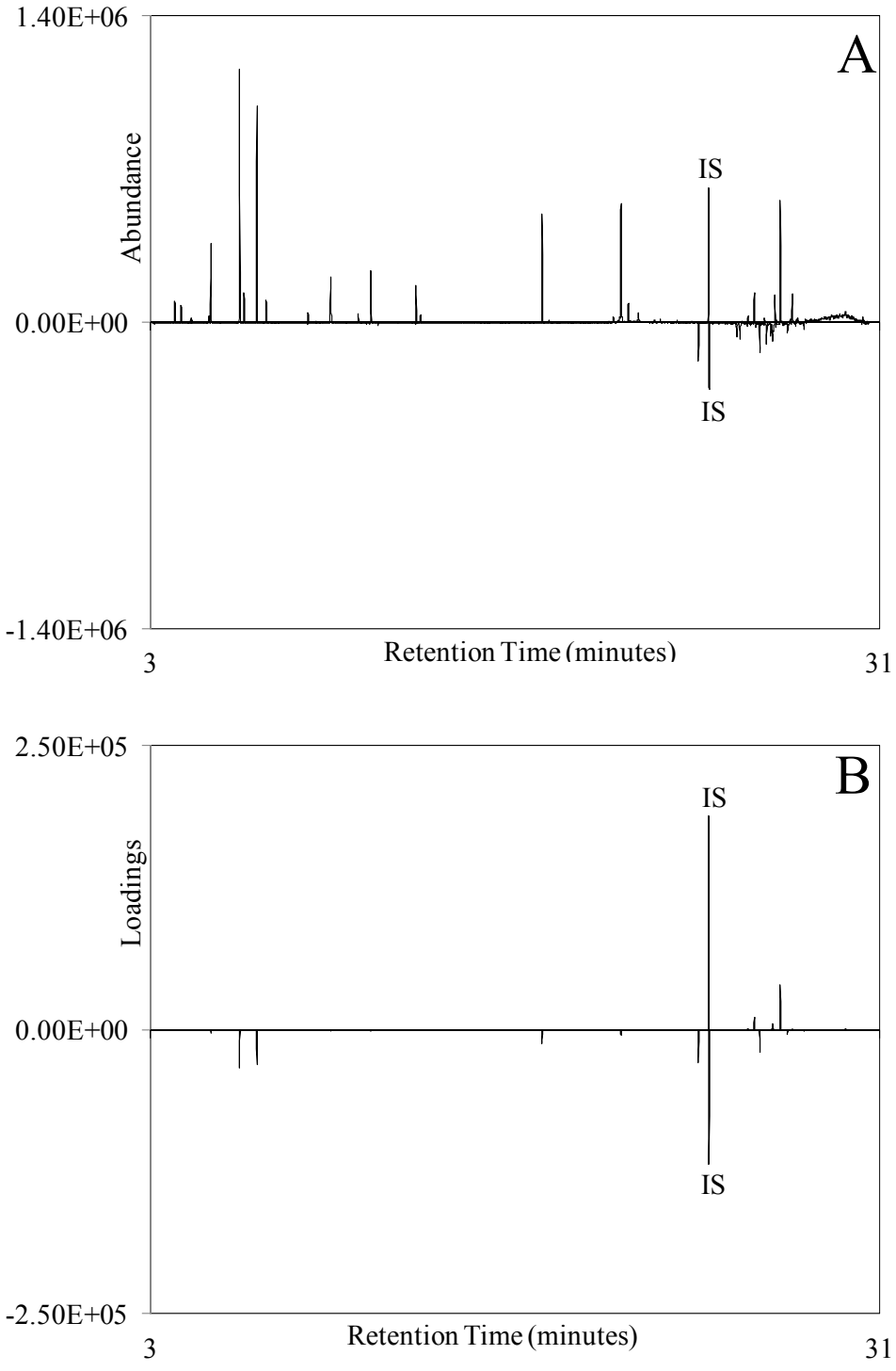

Figure 3.11: (A) Mean-centered data of one replicate of *S. officinalis* and (B) loadings for the same replicate of *S. officinalis* on principal component 2 based on chemical profiles obtained by gas chromatography-mass spectrometry with progesterone added as the internal standard (IS).

was subtracted, the result was a positive contribution at this retention time in the mean-centered data. However, this same replicate had the lowest abundance of internal standard among the replicates at 24.46 min in the TIC. Therefore, when the average abundance of all TICs at this retention time was subtracted, the result was a negative contribution at this retention time in the mean-centered data.

The mean-centered data were then multiplied by the eigenvector for PC2, which generated the loadings, and every data point was summed to generate the score for the sample on PC2. The eigenvector was positive at these retention times (Figure 3.9B) such that, when multiplied by the mean-centered data, the product was positive at 24.44 min and negative at 24.46 min (Figure 3.10B). The internal standard was weighted heavily. For this replicate, the positive loading of the internal standard outweighed the negative loading such that this replicate was positioned positively on PC2 in the scores plot. All replicates of *S. officinalis* contained two internal standard peaks in their PC2 loadings. Therefore, a combination of differences in abundance of various compounds including the internal standard, as well as a misalignment of the internal standard, contributed to the positioning of *S. officinalis* replicates on PC2.

S. divinorum was positioned negatively on both PC1 and PC2 in the scores plot. From the corresponding loadings plots, salvinorin A, the main compound in *S. divinorum*, was weighted negatively on both PCs, which resulted in the observed positioning in the scores plot.

S. guaranitica was positioned negatively on PC1 in the scores plot. Some compounds present in *S. guaranitica* (*i.e.* phthalide, hentriacontane, tetratriacontane, and β-amyrin) were weighted positively on PC1 (Figure 3.9A). However, many more compounds in this species (*e.g.*, β-cubebene, squalene, sitosterol, and α-amyrin) were weighted negatively. Overall, *S. guaranitica* was positioned negatively on PC1 in the scores plot.

S. guaranitica was positioned around zero on PC2. In this case, there were a similar number of compounds weighted positively as weighted negatively (Figure 3.9B). Therefore, the positive and negative weighting compounds effectively canceled out, which resulted in *S. guaranitica* being positioned around zero on PC2.

S. nemorosa was positioned negatively on PC1. Some compounds from *S. nemorosa* (*i.e.* β-caryophyllene, hentriacontane, 11-n-decyltetracosane, tetratriacontane, and hextriacontane) were weighted positively, while other compounds (*i.e.* β-cubebene, an alkene, vitamin E, a phenyl, and sitosterol) were weighted negatively on PC1. To explain the overall negative positioning of this species on PC1, the mean-centered data were considered (Figure 3.12).

Due to the process of mean centering, compounds from other *Salvia* species were introduced. For the replicate of *S. nemorosa* in Figure 3.12, compounds from *S. officinalis* (*e.g.*, thujone, camphor, and epimanool) had a negative contribution in the mean-centered data. When multiplied by the positive weighting compounds at these retention times in the eigenvector for PC1, the product was negative. Since there were more negative loading compounds than positive, *S. nemorosa* was positioned negatively on PC1 in the scores plot.

S. nemorosa was positioned around zero on PC2. β-caryophyllene was weighted negatively while the other compounds present in *S. nemorosa* were weighted positively on PC2. The positioning of *S. nemorosa* could be explained again with reference to the mean-centered data. The abundance of all compounds was minimal in comparison to the abundance of the internal standard, which, as previously mentioned, had both positive and negative contributions in all but one replicate of *S. nemorosa*. With the exception of this replicate, the internal standard was loaded both positively and negatively when multiplied by the eigenvector for PC2. In

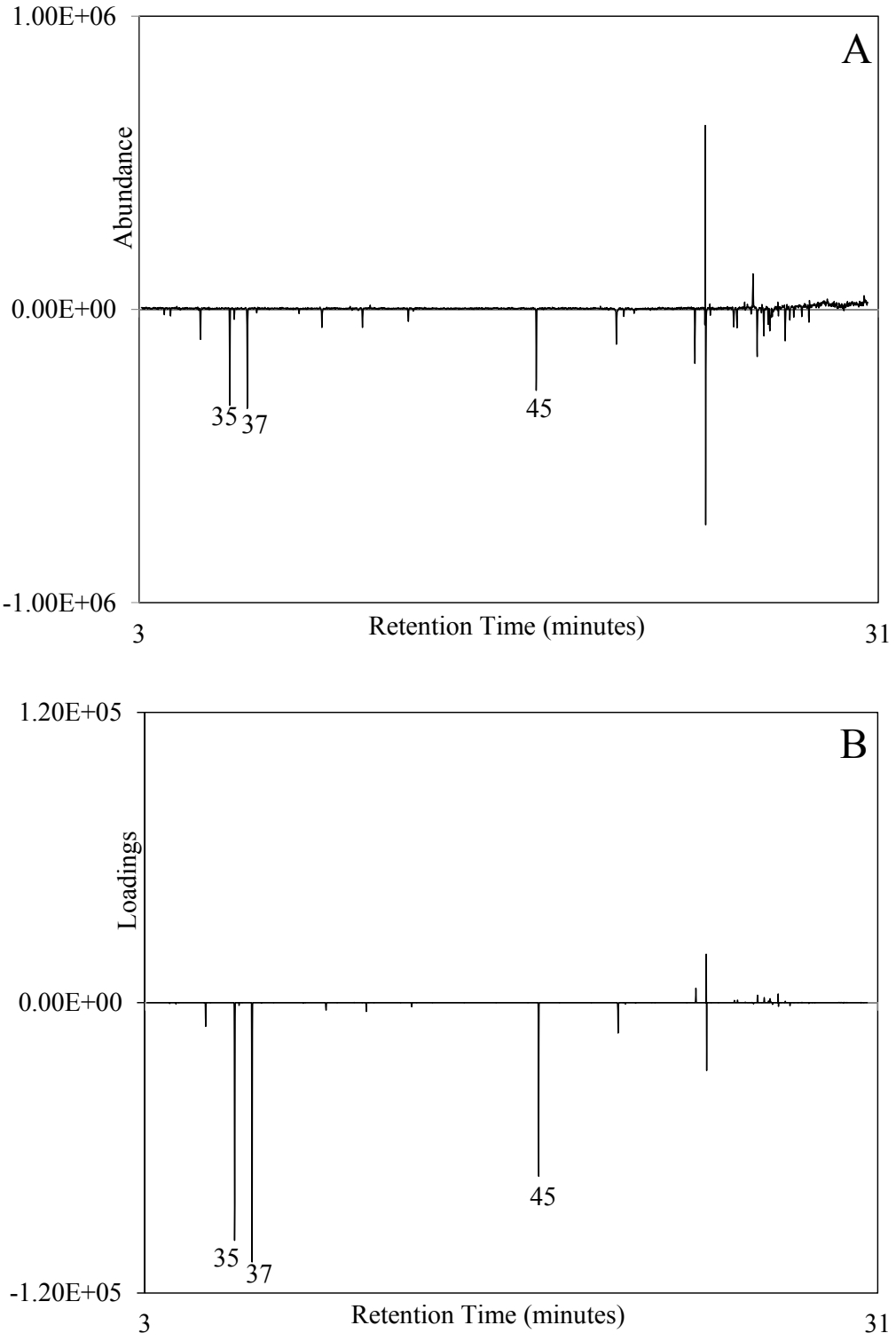

Figure 3.12: (A) Mean-centered total ion chromatogram of one replicate of *S. nemorosa* and (B) loadings for the same replicate of *S. nemorosa* on principal component 1 based on chemical profiles obtained by gas chromatography-mass spectrometry with 35=thujone, 37=camphor, and 45=epimanool.

general, the positive and negative loadings canceled out, which resulted in *S. nemorosa* being positioned around zero on PC2.

The replicate of *S. nemorosa* discussed in Figure 3.8, labeled replicate 1A, was positioned more negatively than the others in the scores plot. After alignment was performed, the internal standard in this replicate displayed drift in retention time. Therefore, this replicate had a lower than average abundance of the internal standard at 24.46 min such that, on mean centering, there was a negative contribution at this retention time. When multiplied by the eigenvector for PC2, this peak was loaded negatively and had a greater contribution in this replicate than in the other replicates of *S. nemorosa*. As a result, this replicate was positioned most negatively among the other replicates of *S. nemorosa* on PC2 in the scores plot. Therefore, the spread observed in PC2 for *S. nemorosa* was due to the misalignment of the internal standard.

S. splendens was positioned negatively on PC1. Several compounds from *S. splendens* (*i.e.* nonacosane, hentriacontane, tetratriacontane, hexatriacontane, and oleanolic acid) were weighted positively, while other compounds from *S. splendens* (*i.e.* an alkene, squalene, vitamin E, 8-epi-salviarin, a cardenolide, sitosterol, and α-amyrin) were weighted negatively on PC1 (Figure 3.9A). To better explain the negative positioning of *S. splendens* on PC1, the mean-centered data were investigated. As previously described, compounds from other species (*i.e.* thujone, camphor, and epimanool from *S. officinalis*) were introduced during the process of mean centering. In the mean-centered data for *S. splendens*, these compounds had a negative contribution. When multiplied by the eigenvector for PC1, these compounds were loaded negatively. Overall, the negatively loading compounds contributed more than those loading positively such that *S. splendens* was positioned negatively on PC1. *S. splendens* was positioned

positively on PC2 in the scores plot since all compounds from this species were weighted positively in the loadings plot for PC2 (Figure 3.9B).

To statistically assess the positioning of *S. divinorum* compared to the other four *Salvia* species in the scores plot, two-tailed Student's *t*-tests were performed on the mean scores of all species for both PC1 and PC2, at the 99% confidence level. The unequal variance Student's *t*-test was used. The null hypothesis, H_0, tested was that the mean scores of the two species tested were equal, while the alternative hypothesis, H_1, stated that the mean scores of the two species were statistically different. Table 3.8 shows the absolute value of the calculated *t* values and the critical *t* values for comparisons of *S. divinorum* and the other four *Salvia* species on both PC1 and PC2 [12]. For all comparisons, the absolute value of $t_{calculated}$ was greater than $t_{critical}$. Therefore, *S. divinorum* was statistically differentiated from the remaining four species on both PC1 and PC2, at the 99% confidence level.

Due to the close proximity of *S. guaranitica, S. nemorosa,* and *S. splendens* in the scores plot, Student's *t*-tests were also performed on pair-wise comparisons of these species to determine if they could be statistically differentiated (Table 3.9). *S. guaranitica* was statistically differentiated from *S. splendens* on both PC1 and PC2 at the 99% confidence level. *S. nemorosa* and *S. guaranitica* were statistically differentiated on PC1 at the 99% confidence level. These species were not statistically differentiated on PC2 at the 99% or the 95% confidence level, but were differentiated at the 90% confidence level. *S. nemorosa* and *S. splendens* were statistically differentiated on PC2 at the 99% confidence level; however, these species were not statistically differentiated on PC1 at the 99%, 95%, and 90% confidence level. *S. nemorosa* and *S. splendens* were positioned closely on the scores plot due to their chemical similarity. *S. nemorosa* and *S.*

Table 3.8: Absolute value of $t_{calculated}$ and $t_{critical}$ values calculated using two-tailed Student's *t*-tests performed on the mean scores of principal component 1 (PC1) and principal component 2 (PC2) for comparisons (n=18) of *S. divinorum* to the other *Salvia* species*

Comparisons of *S. divinorum* to	PC1			PC2		
	$\lvert t_{calculated} \rvert$	$t_{critical}$	Significant Difference	$\lvert t_{calculated} \rvert$	$t_{critical}$	Significant Difference
S. guaranitica	13.53	2.947	Yes	43.14	2.947	Yes
S. nemorosa	22.04	3.106	Yes	17.94	3.250	Yes
S. officinalis	13.29	3.355	Yes	28.81	3.012	Yes
S. splendens	20.28	3.106	Yes	58.80	3.055	Yes

*99% confidence interval [12]

Table 3.9: Absolute value of $t_{calculated}$ and $t_{critical}$ values calculated using two-tailed Student's *t*-tests performed on the mean scores of principal component 1 (PC1) and principal component 2 (PC2) for comparisons (n=18) between *S. guaranitica, S. nemorosa,* and *S. splendens*

Comparisons	PC1			PC2		
	$\lvert t_{calculated} \rvert$	$t_{critical}$	Significant Difference	$\lvert t_{calculated} \rvert$	$t_{critical}$	Significant Difference
S. nemorosa to *S. splendens*	1.728	1.753*	No	13.97	3.106**	Yes
S. nemorosa to *S. guaranitica*	13.31	3.106**	Yes	2.272	1.860*	Yes
S. guaranitica to *S. splendens*	11.40	3.055**	Yes	31.93	3.106**	Yes

*90% confidence interval [12]
**99% confidence interval [12]

splendens both contained an alkene, nonacosane, hentriacontane, vitamin E, tetratriacontane, sitosterol, and hextriacontane.

Although Student's *t*-tests demonstrated that *S. divinorum* could be differentiated from four related *Salvia* species on PC1 and PC2 at the 99% confidence level, the scores plot only accounted for 77% of the total variance. Therefore, PPMC coefficients, which assess similarity between entire TICs based on a point-by-point comparison, were calculated between each *Salvia* species, as well as between replicates of each species.

Mean PPMC coefficients calculated between *S. divinorum* and the remaining four *Salvia* species were greater than 0.8, which indicated strong correlation (Table 3.10). However, the range of PPMC coefficients between *S. divinorum* and the other four species indicated moderate to strong correlation, except for the comparison of *S. divinorum* to *S. guaranitica* and to *S. splendens*, which indicated strong correlation. These correlations were due to chemical similarities between *S. divinorum* and the other four species. For example, all five species contained hentriacontane, tetratriacontane, and sitosterol. *S. divinorum*, *S. guaranitica,* and *S. splendens* all contained squalene. In addition, *S. divinorum* and *S. splendens* both contained nonacosane. Furthermore, the internal standard added to all five species increased the correlation between the species. This issue needs to be further investigated in the future.

S. divinorum and *S. nemorosa* contained some common compounds (*i.e.* hentriacontane, tetratriacontane, sitosterol, and nonacosane). However, the moderate correlation between these species was due to a misaligned internal standard peak previously mentioned.

S. divinorum and *S. officinalis* also contained common compounds (*i.e.* hentriacontane, tetratriacontane, sitosterol, nonacosane, squalene, and heptacosane). *S. officinalis* contained

Table 3.10: Mean, standard deviation, and range of Pearson product moment correlation (PPMC) coefficients for comparisons (n=81) between *Salvia* species

Comparisons	Mean PPMC coefficient ± standard deviation	Range of PPMC coefficient
S. divinorum to *S. guaranitica*	0.9537 ± 0.0188	0.9077 –0.9821
S. divinorum to *S. nemorosa*	0.9490 ± 0.0460	0.7295 –0.9856
S. divinorum to *S. officinalis*	0.8100 ± 0.0508	0.7382 –0.8938
S. divinorum to *S. splendens*	0.9477 ± 0.0157	0.8999 –0.9759
S. guaranitica to *S. nemorosa*	0.9491 ± 0.0487	0.7283–0.9845
S. guaranitica to *S.officinalis*	0.8128 ± 0.0516	0.7402–0.8964
S. guaranitica to *S. splendens*	0.9529 ± 0.0166	0.8984–0.9777
S. nemorosa to *S. officinalis*	0.8255 ± 0.0621	0.6624–0.9206
S. nemorosa to *S. splendens*	0.9500 ± 0.0420	0.7925–0.9848
S. officinalis to *S. splendens*	0.8202 ± 0.0521	0.7454–0.9072

many early-eluting compounds that were not present in *S. divinorum*, which resulted in the moderate correlation.

In addition, mean PPMC coefficients and ranges of PPMC coefficients were calculated between pair-wise comparisons of replicates of *S. guaranitica, S. nemorosa,* and *S. splendens* since these species were closely positioned on the scores plot. All mean PPMC coefficients were greater than 0.9, which indicated strong correlation that was due to chemical similarities among these species. However, the range of PPMC coefficients indicated a moderate to strong correlation for *S. nemorosa* to *S. guaranitica* and to *S. splendens*, which was again due to the misalignment of the internal standard peak in *S. nemorosa*.

S. divinorum was visually discriminated from the four *Salvia* species based on TICs. However, PPMC coefficients showed that *S. divinorum* was moderately to strongly correlated to the other four *Salvia* species; however, the internal standard added to all species artificially increased the correlation coefficients. Using PCA and Student's *t*-tests, *S. divinorum* was discriminated from the other four species. This apparent contradiction can also be explained by considering how PPMC coefficients and the PCA scores plot were calculated. Both PPMC coefficients and the PCA scores plot were calculated using entire TICs; however, PPMC coefficients were calculated between two TICs at a time based on a point-by-point comparison of all data points. PCA magnified the differences between species and was used to identify the most variable compounds among the species. Therefore, the PCA scores plot only accounted for a certain amount of variance among the samples and as little similarity as possible. In addition, PPMC coefficients were insensitive to relative abundance differences, while PCA was sensitive to abundance differences. Therefore, the positioning of the species on the scores plot was affected by relative abundance differences; however, the calculation of the PPMC coefficients

was not. Overall, this research showed that PCA was helpful in identifying the discriminating factors among the entire data set and PCA along with Student's *t*-tests discriminated *S. divinorum* from the other four species despite all the similarities among the TICs of the species, which was shown with PPMC coefficients.

3.4 Summary

S. divinorum and four related *Salvia* species were extracted by rotary agitation for 16 hours with dichloromethane and analyzed by GC-MS. Following analysis, TICs were background subtracted, truncated, smoothed, retention time aligned, and normalized. PPMC coefficients were calculated for pair-wise comparisons of replicate injections of each extract and for all replicates of each species. In general, moderate to strong correlation among injection and extraction replicates indicated acceptable precision in the extraction, instrumental analysis, and data pretreatment procedures.

Using TICs, *S. divinorum* was visually discriminated from the four *Salvia* species. Using PCA, each *Salvia* species was visually discriminated from *S. divinorum* on the scores plot. Student's *t*-tests, calculated on the mean scores on both PCs, showed *S. divinorum* was statistically differentiated from the four other *Salvia* species at the 99% confidence level. PPMC coefficients were also calculated between *S. divinorum* and the other *Salvia* species. Moderate to strong correlations were observed, which indicated chemical similarity between *S. divinorum* and the other four *Salvia* species. However, the internal standard added to each species increased the correlation between species. Therefore, *S. divinorum* was discriminate from the other *Salvia* species using TICs, PCA, and Student's *t*-tests.

References

References

[1] Jermain JD, Evans HK. 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.

[2] Bodnar Willard MA. Forensic analysis of *Salvia divinorum* and related *Salvia* species using chemometric procedures. MS Thesis (2010).

[3] Radulescu V, Chiliment S, Oprea E. Capillary gas chromatography-mass spectrometry of volatile and semi-volatile compounds of *Salvia officinalis.* Journal of Chromatography A 2004; 1027:121–6.

[4] Ulubelen A, Topcu G, Sönmez U, Eriş C. Terpenoids from *Salvia nemorosa.* Phytochemistry 1994; 35:1065–7.

[5] Carrer RP, Vanderlinde R, Dutra S, Marcon A, Echeverrigaray S. Essential oil variation among Brazilian accessions of *Salvia guaranitica L*. Flavour Fragr. J. 2007; 22:430–4.

[6] Vallverdú C, Vila R, Lorenzo D, Paz D, Dellacassa E, Davies P, Villamil J, Tomi F, Casanova J, Cañigueral S. Composition of the essential oil of cultivated *Salvia guaranitica* from Uruguay 2005; 20:421–4.

[7] Fontana G, Savona G, Rodríguez B, Dersch C, Rothman R, Prisinzano T. Synthetic studies of neoclerodane diterpenoids from *Salvia splendens* and evaluation of opioid receptor affinity. Tetrahedron 2008; 64:10041–8.

[8] Johnson KJ, Wright BW, Jarman KH, Synovec RE. High-speed peak matching algorithm for retention time alignment of gas chromatographic data for chemometric analysis. Journal of Chromatography A 2003; 996:141–55.

[9] Walker JB, Sytsma KJ, Treutlein J, Wink M. *Salvia* (Laminaceae) is not monophyletic: implications for the systematics, radiation, and ecological specializations of *Salvia* and tribe Mentheae. American Journal of Botany 2004; 91:1115–25.

[10] Tychonievich, J. *Salvia* clades. Personal Communication. 2009.

[11] 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; 20:131–6.

[12] Adam C. Essential mathematics and statistics for forensic science. Hobken, NJ: Wiley-Blackwell 2010.

Chapter 4: Association and Discrimination of *Salvia divinorum* from Related *Salvia* Species Based on Chemical Fingerprints Generated Using Liquid Chromatography-Mass Spectrometry

4.1 Introduction

Gas chromatography-mass spectrometry (GC-MS) is limited to the analysis of volatile compounds that are thermally stable within normal operating temperatures of the instrument. A chemical fingerprint of *Salvia divinorum* generated using GC-MS would then exclude several non-volatile compounds. Hence, in this research, liquid chromatography-mass spectrometry (LC-MS) was used to analyze the non-volatile compounds in *S. divinorum* and related *Salvia* species. Chemical fingerprints generated using LC-MS are expected to be complimentary to those generated using GC-MS. However, LC-MS may potentially offer more discriminating information because some plant species are known to contain several non-volatile compounds [1].

This chapter describes the discrimination of *S. divinorum* from the other *Salvia* species based on their polar, non-volatile chemical compounds using LC-MS. Firstly, extraction solvents and ionization modes were investigated using *S. divinorum* as the model plant species. Then, *S. divinorum, S. guaranitica, S. nemorosa, S. officinalis,* and *S. splendens* were extracted using a rotary agitator for 16 hours and analyzed using LC-MS and LC-tandem mass spectrometry (MS^2). Resulting total ion chromatograms (TICs) were truncated, retention time aligned, and normalized. Mean Pearson product moment correlation (PPMC) coefficients were then used to evaluate the precision of extraction and analysis procedures and to evaluate the association of *S. divinorum* extracts and discrimination of *S. divinorum* from the other four *Salvia* species. Principal components analysis (PCA) was used to evaluate the discrimination of *S. divinorum* from the other four *Salvia* species in the scores plot.

Loadings plots were used to identify the compounds contributing most to the variance among the species. Student's *t*-tests were used to statistically evaluate the association and discrimination in the PCA scores plot.

4.2 Material and Methods

4.2.1 Investigation of Extraction Solvent

Using leaves of *S. divinorum*, a preliminary study was conducted comparing two extraction solvents: 100% acetonitrile (HPLC grade, EMD Chemical Inc., Darmstadt, Germany) and a 3:3:2, by volume, mixture of acetonitrile (EMD Chemical Inc.), 2-propanol (ACS grade, Sigma-Aldrich, St. Louis, MO), and water (HPLC grade, J. T. Baker, Phillipsburg, NJ). Acetonitrile, or mixtures containing acetonitrile, have been used previously for the extraction of *S. divinorum* and other plant materials [2-5]. In addition, acetonitrile is more compatible with LC systems than other solvents used for GC analysis such as dichloromethane.

All glassware used for the extraction was acid-washed following standard procedures. Dried *Salvia divinorum* leaves were purchased from a local supplier (Frivolity Kingdom, Jackson, MI). Leaves were stored in their original packaging at room temperature in complete darkness until needed for analysis. A rapid and simple extraction was performed since these data were only qualitatively analyzed to investigate extraction solvent. Approximately 0.2 g dried *S. divinorum* leaves were placed in a beaker and 10 mL acetonitrile were added. After 5 minutes, the extract was filtered using the experimental set-up described in Chapter 3 (Figure 3.1). The beaker containing the extract was then rinsed with 5 mL acetonitrile and filtered. The extract and rinses were collected, and then evaporated to dryness using nitrogen gas (Airgas, Independence,

OH) and a water circulating bath (NESLAB endocal RTE-9B, Newington, NH) at 35 °C. The resulting extract was then stored at 2 °C until analysis. This procedure was repeated with new *S. divinorum* leaves using the acetonitrile/2-propanol/water mixture as the extraction solvent.

4.2.2 Investigation of Liquid Chromatography-Mass Spectrometry Ionization Mode

Prior to LC-MS analysis, extracts were reconstituted in 0.58 mL of the appropriate solvent (acetonitrile or acetonitrile/2-propanol/water). To assess injection precision, progesterone (Sigma Chemical Company, St. Louis, MO) was added as an internal standard (0.00119 M) only to the sample extracted using acetonitrile. However, due to the fixed volume injection loop of the instrument, the investigation of injection precision by this method was later deemed unnecessary; therefore, the addition of the internal standard to subsequent extracts was discontinued. The two extracts were analyzed using an Applied Biosystems Sciex 3200 Q TRAP LC-MS system (Foster City, CA) equipped with Analyst Software (version 1.4.2) and introduced via a Shimadzu SILHTC autosampler (Columbia, MD), maintained at 15 °C, with a 5 µL fixed volume injection loop. An Ascentis Express C-18 column (5.0 cm x 2.1 mm x 2.7 µm, Sigma-Aldrich, St Louis, MO) was used and was maintained at 50 °C during analysis. The mobile phase was composed of 10 mM ammonium acetate (ACS reagent, J. T. Baker, Phillipsburg, NJ) in water (Milli-Q, Billerica, MA), and methanol (HPLC grade, Sigma-Aldrich, St. Louis, MO), with a constant flow rate of 0.4 mL/min. The methanol gradient was 1% at 1 minute, increasing linearly to 99% until 20 minutes. From 20–27 minutes, the methanol gradient was held at 99%. Immediately after, the methanol gradient was reduced back to 1% and held for approximately 3 minutes.

Following separation, compounds were ionized using electrospray ionization (ESI) with a source temperature of 550 °C. Both positive and negative ionization modes were investigated.

The ion spray voltage was 5500 V for positive mode and -4500 V for negative mode. The declustering potential was 45 V for positive mode and -45 V for negative mode. The mass analyzer was a hybrid triple quadrupole/linear ion trap. Nitrogen was used as the collision gas, set to an approximate pressure of 3.3×10^{-5} torr. A dynamic fill time was used to enhance the accuracy of the mass spectra. An enhanced mass spectrum (EMS) scan was collected from *m/z* 80–1200 with a scan rate of 1000 amu/s.

4.2.3 Extraction of Salvia Species

Following investigation of the extraction solvent and ionization mode, *S. divinorum* and the four *Salvia* species (*S. guaranitica,* S. *nemorosa, S. officinalis,* and *S. splendens*) were extracted for analysis. *Salvia* leaves were obtained from the same sources and extracted using the procedure described in Chapter 3, section 3.2.1, except using acetonitrile (EMD Chemical Inc.) as the extraction solvent rather than dichloromethane. This extraction procedure was repeated in triplicate for each *Salvia* species, using different leaf material for each extraction. All extracts were stored at 2 °C until analysis.

4.2.4 Liquid Chromatography-Mass Spectrometry Analysis

All extracts were reconstituted in 0.58 mL acetonitrile (EMD Chemical Inc.), shaken and vortexed, and each extract was divided into three autosampler vials with inserts to allow for replicate injections. This was necessary to minimize the risk of evaporative loss due to puncturing the septum of the vial during analysis of replicates. Instrument parameters were similar to those described in section 4.2.2 with some exceptions. The autosampler was maintained at 5 °C during analysis to prevent sample degradation. The ion spray voltage was

5500 V and the declustering potential was 30 V. An EMS scan was collected in positive mode over the range *m/z* 80–1000 with a scan rate of 4000 amu/s.

The *Salvia* extracts were also analyzed by MS^2 to aid in compound identification. An information dependent acquisition (IDA) with an EMS scan and enhanced product ion (EPI) scan was performed. The IDA was programmed to select the two most abundant fragments from the EMS scan and perform the EPI scan. The parameters in section 4.2.2 were used with some additions. Due to the presence of masses at *m/z* 87.000, 141.000, 158.000, and 214.1000 in the background, these ions were excluded from tandem experiments. For the EMS scan, the collision energy was set at 10 V with a collision energy spread of 0 V. For the EPI scans, the collision energy was set at 45 V with a collision energy spread of 15 V.

4.2.5 Data Pretreatment

Each extract was analyzed in triplicate, which resulted in a total of nine TICs for each *Salvia* species. All TICs were pretreated prior to data analysis. As a result of using a dynamic fill time during mass analysis, TICs had different numbers of data points. This would cause problems in subsequent PCA because the difference would be identified as the major source of variance among the extracts. As a result, it was necessary to standardize the number of data points across all TICs. To do this, TICs were first truncated to the smallest number of data points (1511) among all *Salvia* species. An acetonitrile TIC with a similar number of data points (1512) was then selected as the target chromatogram and truncated to 1511 data points. Using a correlation optimized warping (COW) algorithm in LineUp™ (Version 3.0, InfoMetrix, Inc., Bothell, WA), truncated TICs were first aligned to the acetonitrile target TIC with a warp of 6 data points and segment size of 250 data points. At this stage, all TICs had the same number of

data points and these data were used in subsequent pretreatment procedures.

For retention time alignment, an average target chromatogram was created in Microsoft Excel (Versions 2007 and 2010, Microsoft Corporation, Redmond, WA). Chapter 3, section 3.2.3, explains how the average target chromatogram was created. Total ion chromatograms were aligned to the target chromatogram using the COW algorithm with a warp of 6 data points and segment size of 65 data points. Following alignment, TICs were normalized using total area normalization by species.

4.2.6 Data Analysis

Mean PPMC coefficients were calculated in Matlab (Version 7.11.0.584 R2010b, The Mathworks, Natick, MA) on pretreated TICs for all pair-wise combinations of replicate injections of the same extract and for replicates of each species to assess the precision of the extraction and analysis procedures. Principal components analysis was then performed on the aligned and normalized data using Matlab. Scores and loadings plots were generated in Microsoft Excel. To statistically assess discrimination of *Salvia divinorum* from the four other *Salvia* species, Student's *t*-tests were performed in Microsoft Excel on the mean scores of each species on both principal component 1 (PC1) and principal component 2 (PC2) at the 99% confidence level. To further assess association and discrimination of *S. divinorum* from the four *Salvia* species, PPMC coefficients were calculated on TICs for all pair-wise comparisons of species.

4.3 Results and Discussion

4.3.1 Investigation of Extraction Solvent and Ionization Mode

Total ion chromatograms of *S. divinorum* extracted using acetonitrile and acetonitrile/2-propanol/water, analyzed using both positive and negative ionization were visually compared. Regardless of the extraction solvent used, *S. divinorum* analyzed in positive mode contained compounds at abundances approximately an order of magnitude greater than those analyzed in negative mode. Additionally, more compounds (*i.e.* salvinorins A–D and F) were identified in positive mode than in negative mode [2, 3, 6]. Therefore, for all subsequent analyses, positive ionization mode was used.

To determine the best extraction solvent, TICs of *S. divinorum* analyzed in positive mode following an extraction using acetonitrile and acetonitrile/2-propanol/water were compared (Figure 4.1). Salvinorins A, B, C, D, F, as well as five unidentified compounds (A–E), were extracted at relatively similar abundances by both solvents [2, 3, 6]. Since no obvious differences in compounds extracted were apparent between the two solvents, practicality of the extraction procedure was considered. In a forensic laboratory, efficient analyses are necessary to increase sample throughput. Due to its polarity, water is difficult to evaporate and therefore the acetonitrile/2-propanol/water was not considered further. For all subsequent extractions, acetonitrile was used as the extraction solvent and all extracts were analyzed in positive mode.

4.3.2 Provisional Identification of Compounds Present in Each Species

Representative TICs of the five *Salvia* species extracted in acetonitrile and analyzed in positive mode are shown in Figures 4.2–4.6. Provisional identifications for compounds present

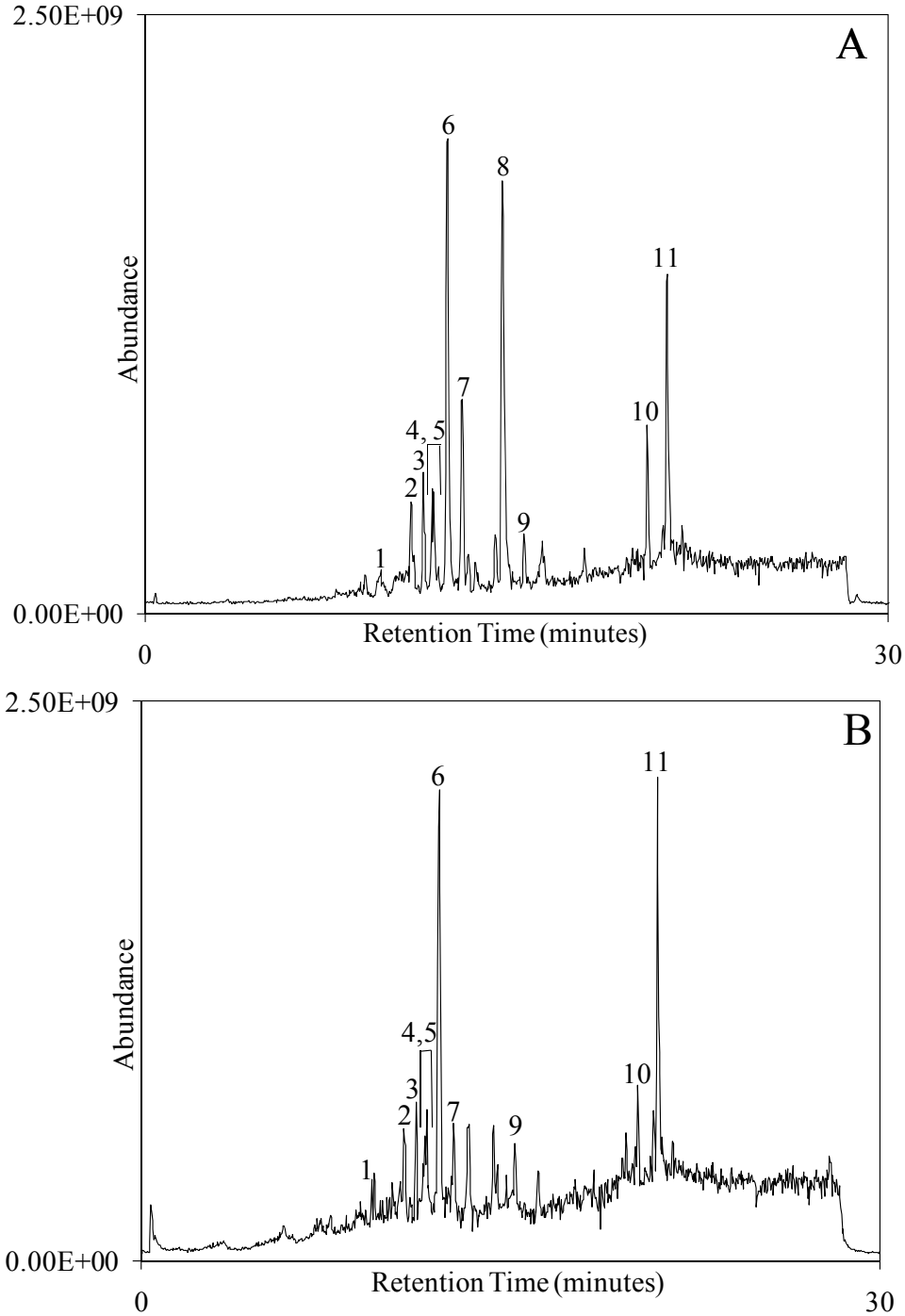

Figure 4.1: Total ion chromatograms of *S. divinorum* extracted in (A) acetonitrile with progesterone added as an internal standard (IS) and (B) acetonitrile/2-propanol/water. Both extracts were analyzed using positive ionization mode. Provisional identifications, based on literature sources, are as follows: 1=salvinorin B, 2=salvinorin D, 3=salvinorin F, 4=unidentified compound A, 5=unidentified compound B, 6=salvinorin A, 7=salvinorin C, 8=progesterone (IS), 9=unidentified compound C, 10=unidentified compound D, and 11=unidentified compound E [2, 3, 6, 7].

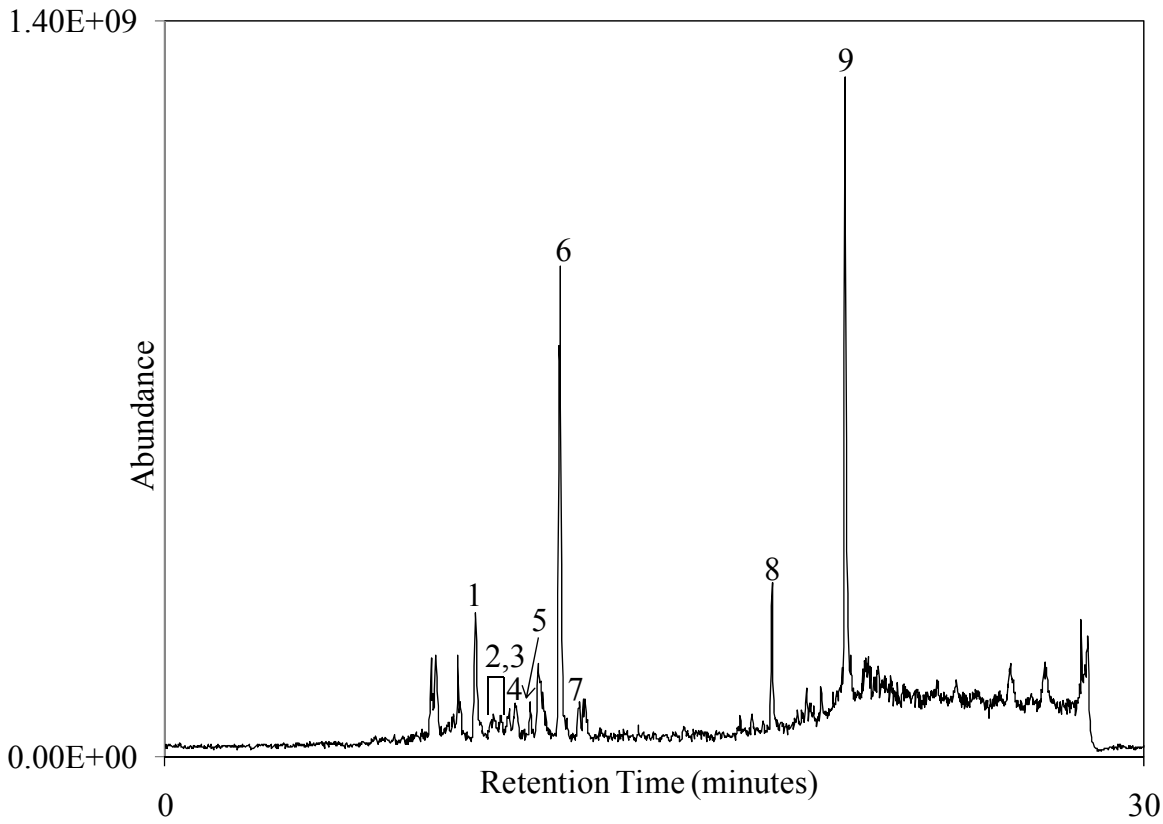

Figure 4.2: Total ion chromatogram of *Salvia divinorum* extracted in acetonitrile and analyzed in positive mode using liquid chromatography-mass spectrometry. Provisional identifications based on literature sources are as follows: 1=salvinorin B, 2=terpenoid (molecular mass= *m/z* 406), 3=terpenoid (molecular mass= *m/z* 378), 4=salvinorin D, 5=salvinorin F, 6=salvinorin A, 7=salvinorin C, 8=sesquiterpenoid, and 9=unidentified compound (molecular mass= *m/z* 426) [2, 3, 6]. See Appendix B, Table B.1, for more detailed information.

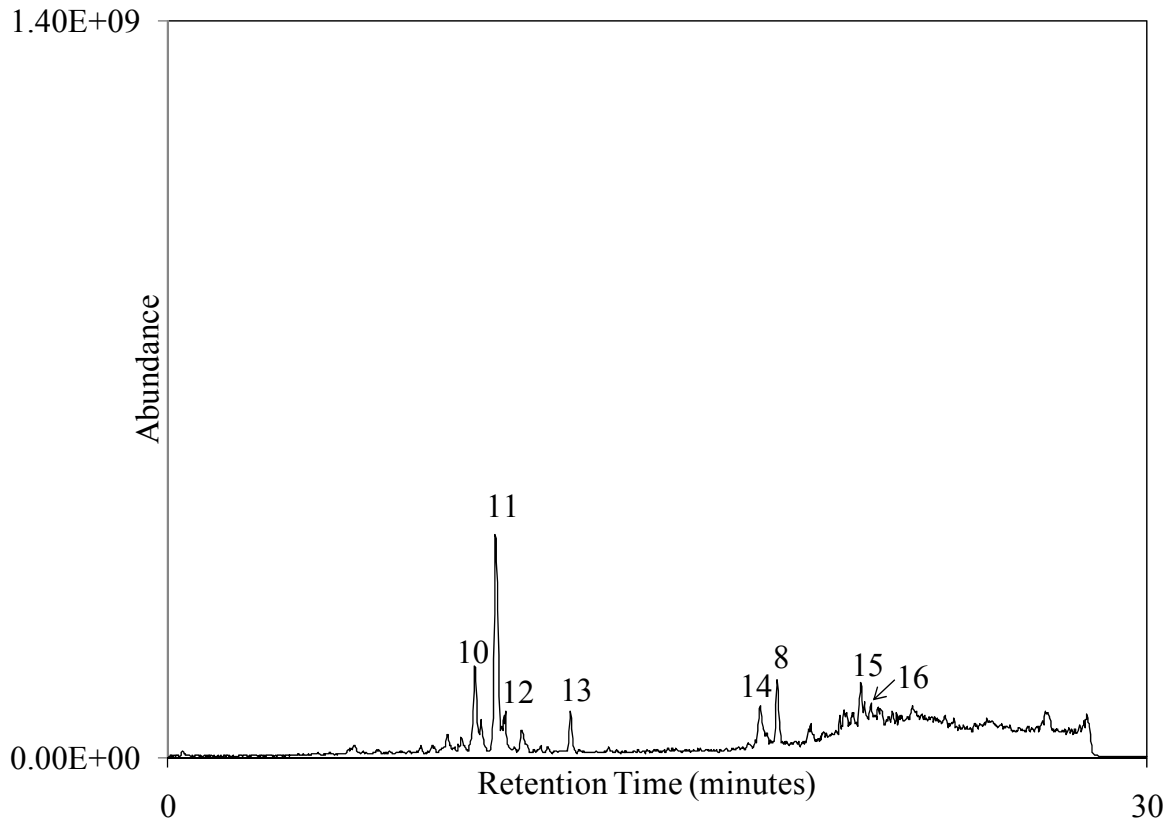

Figure 4.3: Total ion chromatogram of *Salvia guaranitica* extracted in acetonitrile and analyzed in positive mode using liquid chromatography-mass spectrometry. Provisional identifications based on literature sources are as follows: 8=sesquiterpenoid, 10=diterpenoid (molecular mass=*m/z* 340), 11=terpenoid (molecular mass=*m/z* 398), 12= terpenoid (molecular masses=*m/z* 340, 398), 13=terpenoid (molecular mass=*m/z* 398), 14=oleanolic/ursolic acid, 15=vitamin E, and 16=unidentified compound [8–12]. See Appendix B, Table B.2, for more detailed information.

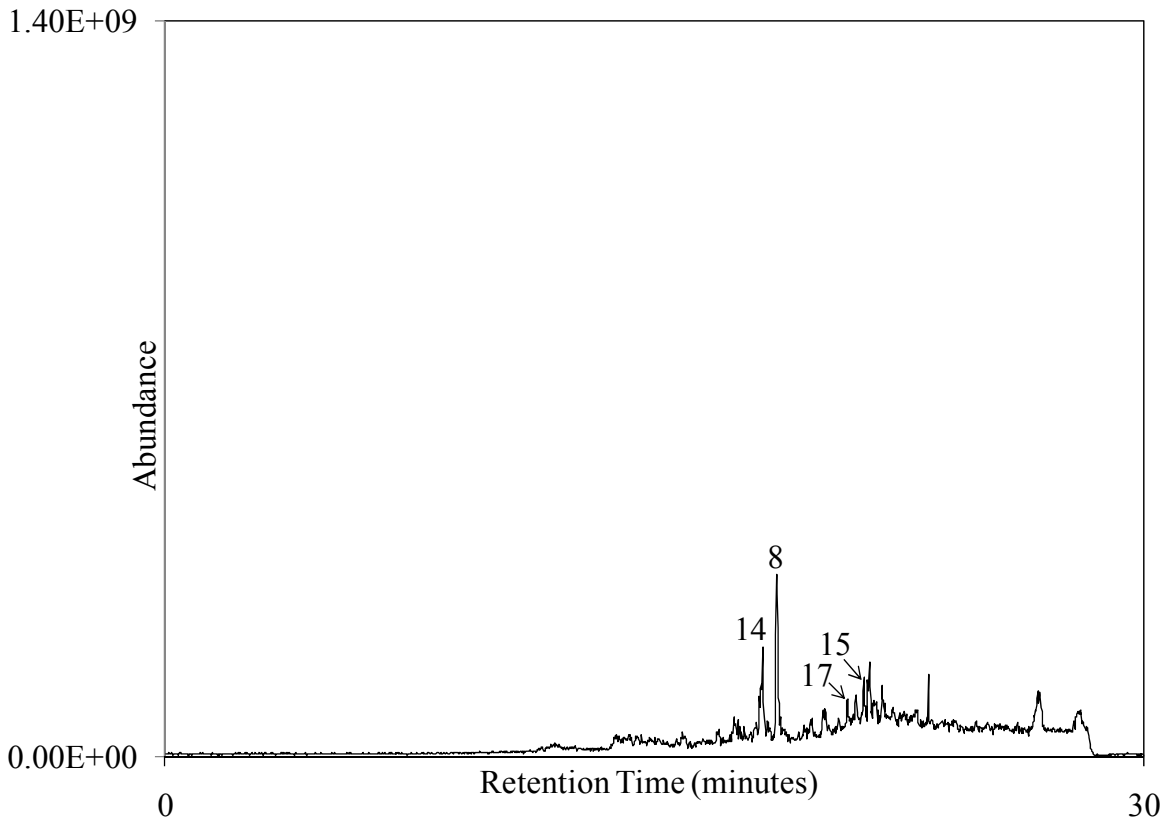

Figure 4.4: Total ion chromatogram of *Salvia nemorosa* extracted in acetonitrile and analyzed in positive mode using liquid chromatography-mass spectrometry. Provisional identifications based on literature sources are as follows: 8=sesquiterpenoid, 14=oleanolic/ursolic acid, 15=vitamin E, and 17=unidentified compound [8–12]. See Appendix B, Table B.3, for more detailed information.

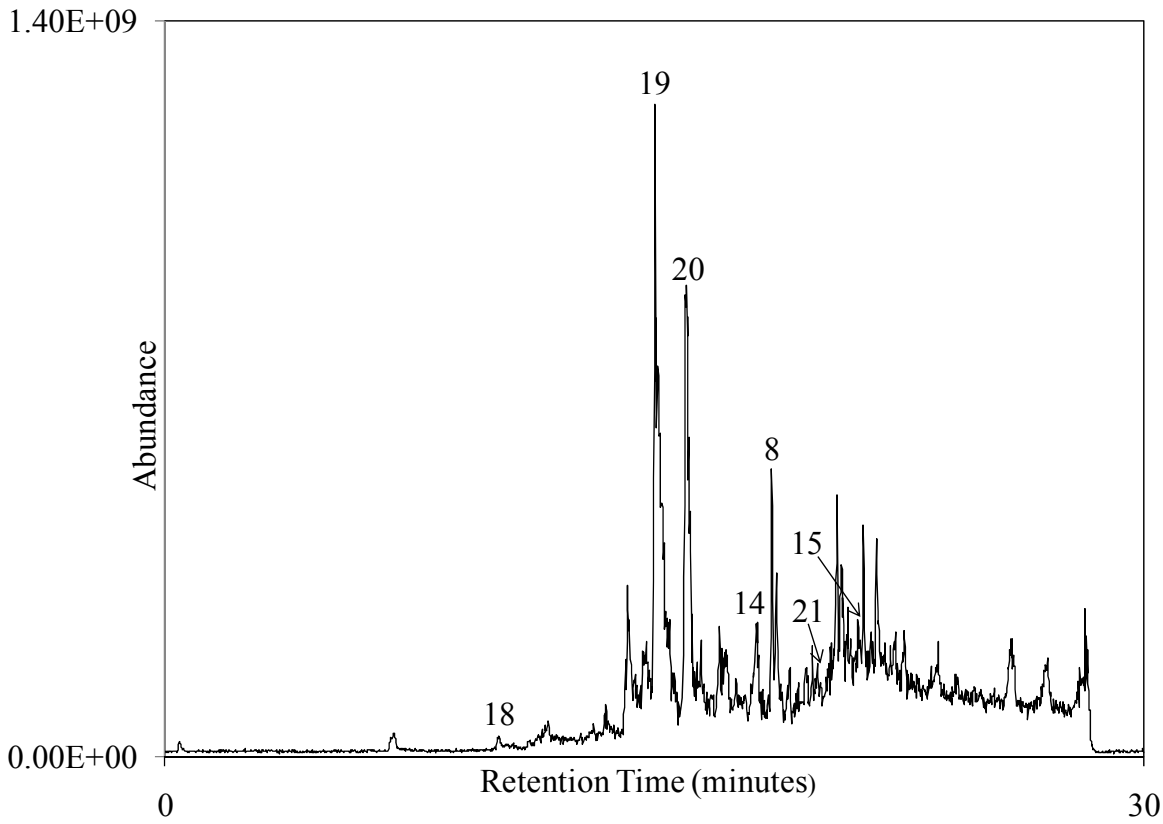

Figure 4.5: Total ion chromatogram of *Salvia officinalis* extracted in acetonitrile and analyzed in positive mode using liquid chromatography-mass spectrometry. Provisional identifications based on literature sources are as follows: 8=sesquiterpenoid, 14=oleanolic/ursolic acid, 15=vitamin E, 18=3-O-methylkaempferol, 19=carnosol and carnosic acid, 20=methyl carnosate/rosmanol/ epirosmanol, and 21=unidentified compound [8–12, 13–17]. See Appendix B, Table B.4, for more detailed information.

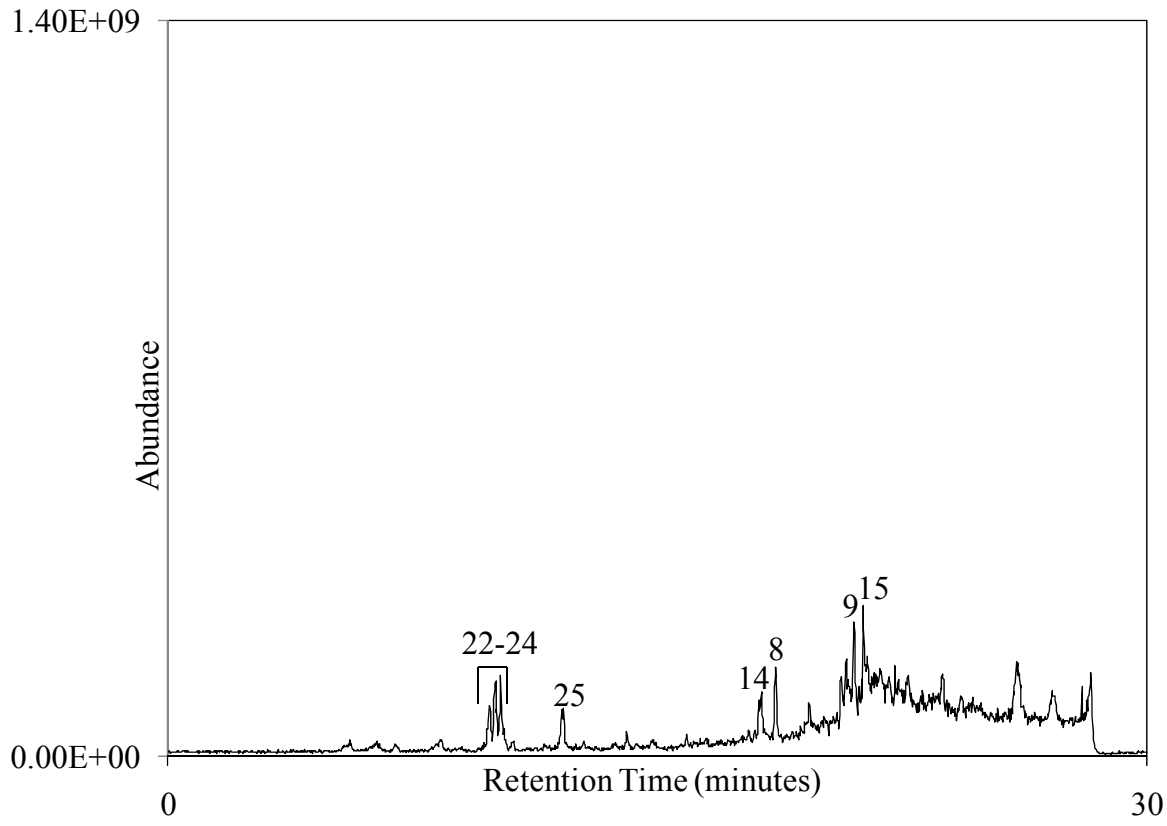

Figure 4.6: Total ion chromatogram of *Salvia splendens* extracted in acetonitrile and analyzed in positive mode using liquid chromatography-mass spectrometry. Provisional identifications based on literature sources are as follows: 8=sesquiterpenoid, 9=unidentified compound (molecular mass=*m/z* 426),14=oleanolic/ursolic acid, 15=vitamin E, 22=diterpenoid (molecular mass=*m/z* 400), 23=diterpenoid (molecular mass=*m/z* 342), 24=diterpenoid (molecular mass=*m/z* 362), and 25=phenolic methyl ether and an unidentified compound (molecular masses=*m/z* 406, 548) [8–12]. See Appendix B, Table B.5, for more detailed information.

in each species with molecular masses and major fragmentation ions are listed in Appendix B. Some peaks are identified as more than one compound, due to compounds being isomers (*e.g.*, oleanolic and ursolic acid, as well as rosmanol and epirosmanol [15]), compounds having the same molecular ion and major fragment ions (*e.g.*, rosmanol, epirosmanol, and methyl carsonate [17]), or compounds being poorly resolved (*e.g.*, compounds 19 and 25).

S. nemorosa and *S. officinialis* belong to *Salvia* clade I, while *S. divinorum, S. guaranitica,* and *S. splendens* belong to *Salvia* clade II [18–20]. Species in different clades contained common compounds; therefore, classification of the species into clades based on the non-volatile compounds was also difficult. For example, all five *Salvia* species contained a sesquiterpenoid with a molecular mass of *m/z* 264. *S. guaranitica, S. nemorosa*, *S. officinalis*, and *S. splendens* all contained oleanolic and/or its isomer, ursolic acid and vitamin E [8–12]. In addition, compounds were common among some species in the same clade but not all species in that clade. For example, *S. divinorum* and *S. splendens,* which are both in *Salvia* clade II, contained an unidentified compound with a molecular mass of m/z 426; however, *S. guaranitica*, which also belongs to *Salvia* clade II, did not contain this compound [18]. Furthermore, it is noteworthy that *S. divinorum* was the only species to contain salvinorins A, B, C, D, and F and TICs of *S. divinorum* were visually discriminated from the other four *Salvia* species [2, 3, 6].

4.3.3 Investigation of Extraction and Analysis Precision Using Pearson Product Moment Correlation Coefficients

The mean, standard deviation, and range of PPMC coefficients were calculated for replicate injections (n=3) of each extract to demonstrate precision in the instrumental analysis procedure (Table 4.3). However, since data pretreatment was performed prior to calculating

Table 4.3: Mean, standard deviation, and range of Pearson product moment correlation (PPMC) coefficients for replicate injections of the same extract

Mean PPMC coefficients ± standard deviation				Range of PPMC coefficients		
Species	Extract 1 (n=3)	Extract 2 (n=3)	Extract 3 (n=3)	Extract 1 (n=3)	Extract 2 (n=3)	Extract 3 (n=3)
S. divinorum	0.9577 ± 0.0192	0.9801 ± 0.0049	0.9563 ± 0.0007	0.9537–0.9786	0.9772–0.9858	0.9558–0.9571
S. guaranitica	0.9641 ± 0.0053	0.9379 ± 0.0254	0.9530 ± 0.0147	0.9585–0.9691	0.9171–0.9663	0.9423–0.9698
S. nemorosa	0.7601 ± 0.0321	0.9496 ± 0.0229	0.8281 ± 0.1086	0.7230–0.7794	0.9257–0.9714	0.7647–0.9535
S. officinalis	0.9718 ± 0.0014	0.9785 ± 0.0046	0.9676 ± 0.0097	0.9705–0.9733	0.9747–0.9836	0.9564–0.9737
S. splendens	0.8679 ± 0.0350	0.9589 ± 0.0153	0.9014 ± 0.0393	0.8293–0.8976	0.9424–0.9727	0.8761–0.9467

PPMC coefficients, precision of these procedures was also reflected in the coefficients.

All mean PPMC coefficients for injection replicates were 0.9 or greater, with the exception of Extracts 1 and 3 of *S. nemorosa* and Extract 1 of *S. splendens*. The ranges of coefficients were all greater than 0.8 except Extract 1 and 3 of *S. nemorosa*. Ideally replicates are expected to have a PPMC coefficient of 1.00. However, the retention time alignment of *S. nemorosa* and *S. splendens* was compromised to improve the alignment of the other species. This resulted in retention time misalignments of the sesquiterpenoid and oleanolic and/or ursolic acid in *S. nemorosa* and *S. splendens,* which lowered the mean PPMC coefficients among replicates of these species. All standard deviations were low (≤ 0.04) except for Extract 3 of *S. nemorosa,* for which the standard deviation was 0.1086. This higher standard deviation was due to misalignments of the sesquiterpenoid and oleanolic and/or ursolic acid as previously discussed. Thus, overall, there was acceptable precision in the instrumental analysis and data pretreatment procedures.

To demonstrate precision in the extraction procedure, in addition to the instrumental analysis and data pretreatment procedures, the mean, standard deviation, and range of PPMC coefficients were also calculated for all extracts of each species (n=36). All mean PPMC coefficients were greater than 0.8, which indicated strong correlation between extracts of the same species (Table 4.4). The ranges of the coefficients were greater than 0.8 with the exception of *S. nemorosa,* which also had a higher standard deviation (0.0874) than the replicates of the other species due to the retention time misalignments discussed previously. Overall, and despite these misalignments, PPMC coefficients indicated acceptable precision in the extraction, instrumental analysis, and data pretreatment procedures.

Table 4.4: Mean, standard deviation, and range of Pearson product moment correlation (PPMC) coefficients for all extracts for each species

Species	Mean PPMC coefficient ± standard deviation (n=36)	Range of PPMC coefficients (n=36)
S. divinorum	0.9498 ± 0.0165	0.9210–0.9858
S. guaranitica	0.9402 ± 0.0199	0.9017–0.9698
S. nemorosa	0.8300 ± 0.0874	0.7214–0.9714
S. officinalis	0.9441 ± 0.0282	0.8853–0.9836
S. splendens	0.9041 ± 0.0377	0.8293–0.9727

4.3.4 Association and Discrimination of Salvia divinorum from Other Salvia Species

The scores plot (Figure 4.7) of PC1 and PC2 accounted for 82% of the total variance among the *Salvia* species. *S. officinalis* was the only species positioned positively on PC1, while *S. divinorum* was the only species positioned positively on PC2. *S. guaranitica, S. nemorosa,* and *S. splendens* were positioned closely, negatively on both PC1 and PC2. Replicates of each species were generally closely clustered in the scores plot on both PC1 and PC2, with the exception of *S. officinalis* for which spread among replicates was observed. The positioning of all species, and spread among replicates of *S. officinalis,* in the scores plot could be explained based on the loadings plots for PC1 and PC2 (Figure 4.8).

S. officinalis was positioned positively on PC1 because it contained all of the compounds that were weighted positively on this PC (*e.g.* phenolic diterpenes, sesquiterpenoid, oleanolic and/or ursolic acid, vitamin E, and an unidentified compound). While one compound present in *S. officinalis* (3-O-methylkaempferol), was weighted negatively on PC1, its contribution was minimal in comparison to the positively weighted compounds [13]. As a result, *S. officinalis* was positioned positively on PC1 in the scores plot.

Compounds present in *S. officinalis* were weighted both positively (*e.g.*, a sesquiterpenoid, vitamin E, and an unidentified compound) and negatively (*e.g.*, 3-O-methylkaempferol, phenolic diterpenes, and oleanolic and/or ursolic acid) on PC2. However, since the contribution of the negatively weighted compounds was greater than that of the positively weighted compounds, *S. officinalis* was positioned negatively on PC2.

The spread observed among replicates of *S. officinalis* was the result of variability in the abundance of carnosol and carnosic acid, and methyl carnosate/rosmanol/epirosmanol in

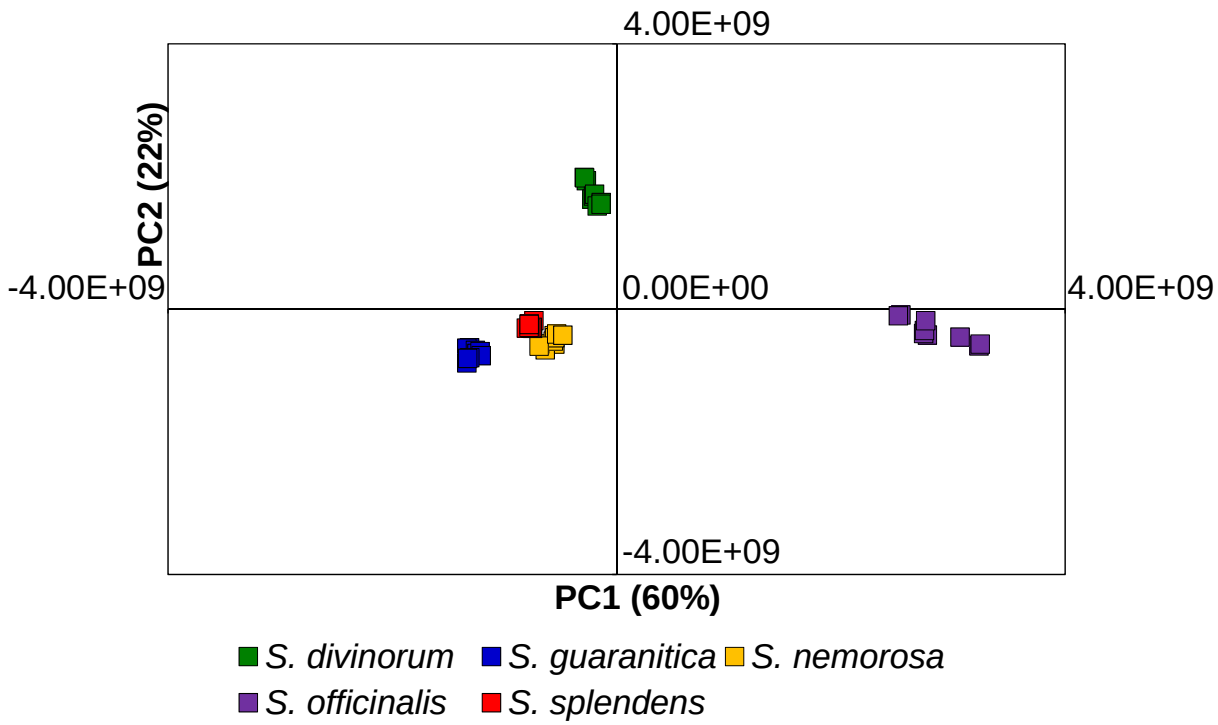

Figure 4.7: Scores plot of principal component 1 versus principal component 2 showing discrimination of *S. divinorum* from the other four *Salvia* species based on chemical profiles obtained by liquid chromatography-mass spectrometry.

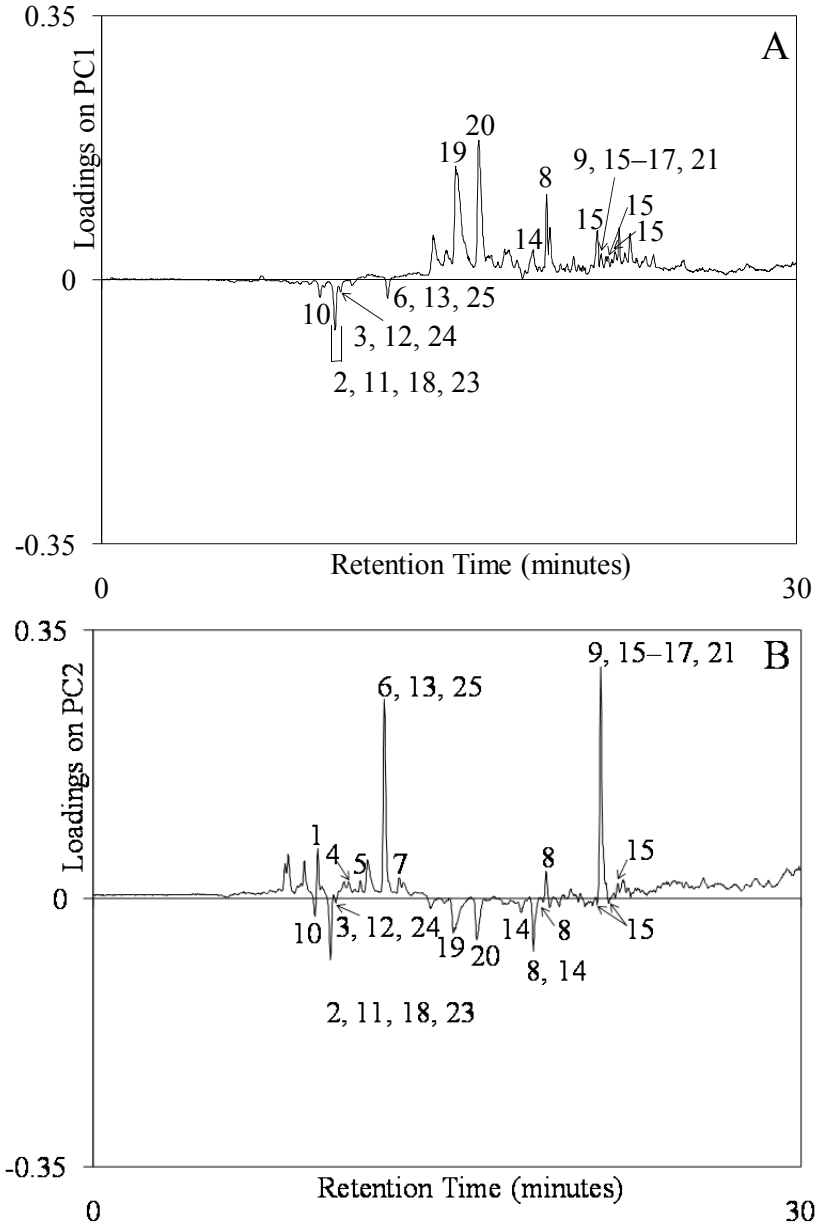

Figure 4.8: Loadings plots for (A) principal component 1 (PC1) and (B) principal component 2 (PC2) based on principal components analysis of the five *Salvia* species analyzed by liquid chromatography-mass spectrometry with 1=salvinorin B, 2=terpenoid , 3=terpenoid, 4=salvinorin D, 5=salvinorin F, 6=salvinorin A, 7=salvinorin C, 8=sesquiterpenoid, 9=unidentified compound, 10=diterpenoid, 11=terpenoid, 12=terpenoid, 13=terpenoid, 14=oleanolic/ursolic acid, 15=vitamin E, 16=unidentified compound, 17=unidentified compound, 18=3-O-methylkaempferol, 19=carnosol and carnosic acid, 20=methyl carnosate/ rosmanol/epirosmanol, 21=unidentified compound, 22=diterpenoid, 23=diterpenoid, 24=diterpenoid, 25=phenolic methyl ether and unidentified compound.

S. officinalis. The amount of carnosic acid and carnosol are known to change as leaves age and therefore, may be the cause of the variability in abundance of these compounds [14–17, 21]. In addition, carnosic acid can decompose into rosmanol, epirosomanol, and/or carnosol in the presence of methanol [22]. As previously stated, methanol was used in the mobile phase to perform the LC analysis. Since these compounds were weighted positively on the loadings plot for PC1 and negatively on the loadings plot for PC2, variability in abundance of these compounds affected the positioning of *S. officinalis* replicates. For instance, replicates with a greater abundance of these compounds were positioned more positively on PC1 and more negatively on PC2 than the other replicates.

S. divinorum was positioned slightly negatively on PC1. This species contained salvinorin A and two terpenoids that were weighted negatively on PC1, as well as a sesquiterpenoid and an unidentified compound that were weighted positively. The overall negative positioning of *S. divinorum* on PC1 in the scores plot was a result of mean centering the TICs, which was necessary for PCA.

To mean center a data set, the first step was to calculate the average abundance of all sample TICs at each retention time. This average was then subtracted from the abundance at each retention time in each TIC to generate the mean-centered data. Figure 4.9 shows the mean-centered data for one replicate of *S. divinorum*.

Due to the process of mean centering, compounds from the other *Salvia* species were introduced. For the replicate of *S. divinorum* in Figure 4.9A, phenolic diterpenes (compounds 19 and 20) from *S. officinalis* had a negative contribution in the mean-centered data. The mean-centered data were then multiplied by the eigenvector for PC1, which generated the loadings, and every data point was summed to generate the score for the sample on PC1. The eigenvector

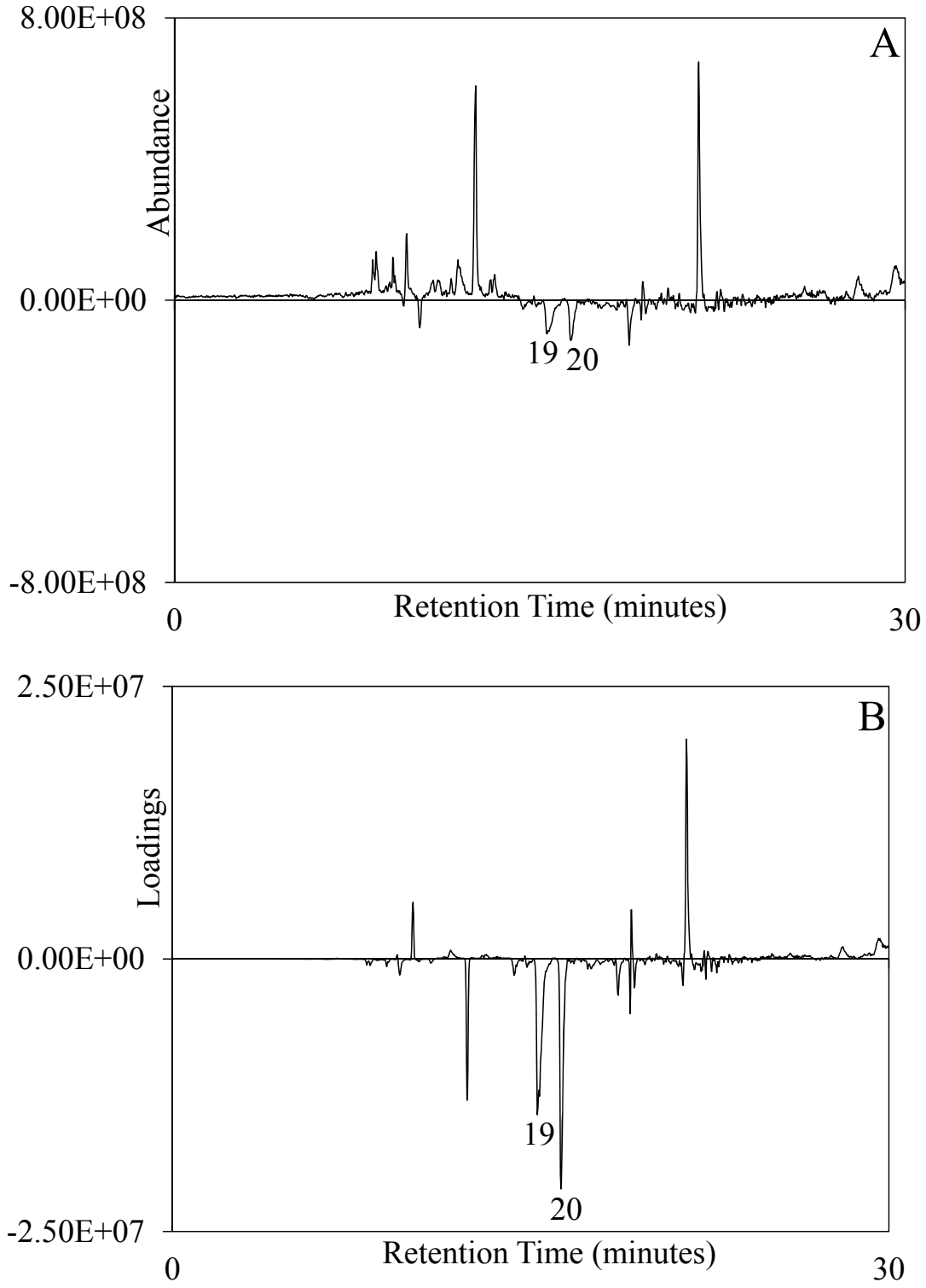

Figure 4.9: (A) Mean-centered total ion chromatogram of *S. divinorum* and (B) loadings for *S. divinorum* on principal component 1 based on chemical profiles obtained by liquid chromatography-mass spectrometry with 19=carnosol and carnosic acid, and 20=methyl carnosate/rosmanol/epirosmanol.

was positive at the retention times corresponding to the phenolic diterpenes. Therefore, when multiplied by the mean-centered data, these compounds were loaded negatively (Figure 4.9B). Overall, the negatively loading compounds contributed more than those loading positively such that *S. divinorum* was positioned slightly negatively on PC1.

S. divinorum was positioned positively on PC2 in the scores plot mainly due to salvinorins A, B, C, D, and F, a sesquiterpenoid, and an unidentified compound that were weighted positively (Figure 4.8B). Two terpenoids from *S. divinorum* were weighted negatively on PC2. However, since many more compounds in this species were weighted positively than negatively, *S. divinorum* was positioned positively on PC2.

S. nemorosa was positioned negatively on PC1 in the scores plot. However, no compounds from *S. nemorosa* were weighted negatively on the PC1 loadings plot. This apparent contradiction in the positioning of *S. nemorosa* could be explained by investigating the mean-centered data. Through the act of mean centering, the phenolic diterpenes from *S. officinalis* were introduced and had a negative contribution. When multiplied by the eigenvector for PC1, these compounds loaded negatively. Overall, the compounds loading negatively contributed more than those loading positively such that *S. nemorosa* was positioned negatively on PC1 in the scores plot.

S. nemorosa was also positioned negatively on PC2 due to the process of mean centering. In addition to the phenolic diterpenes from *S. officinalis*, compounds from other species were introduced into the mean-centered data for *S. nemorosa* (Figure 4.10 A). For example, salvinorin A from *S. divinorum,* a terpenoid from *S. guaranitica*, and a phenolic methyl ether from *S. splendens* contributed negatively in the mean-centered data at approximately 12.3 minutes. When multiplied by the eigenvector for PC2 (Figure 4.10 B), these

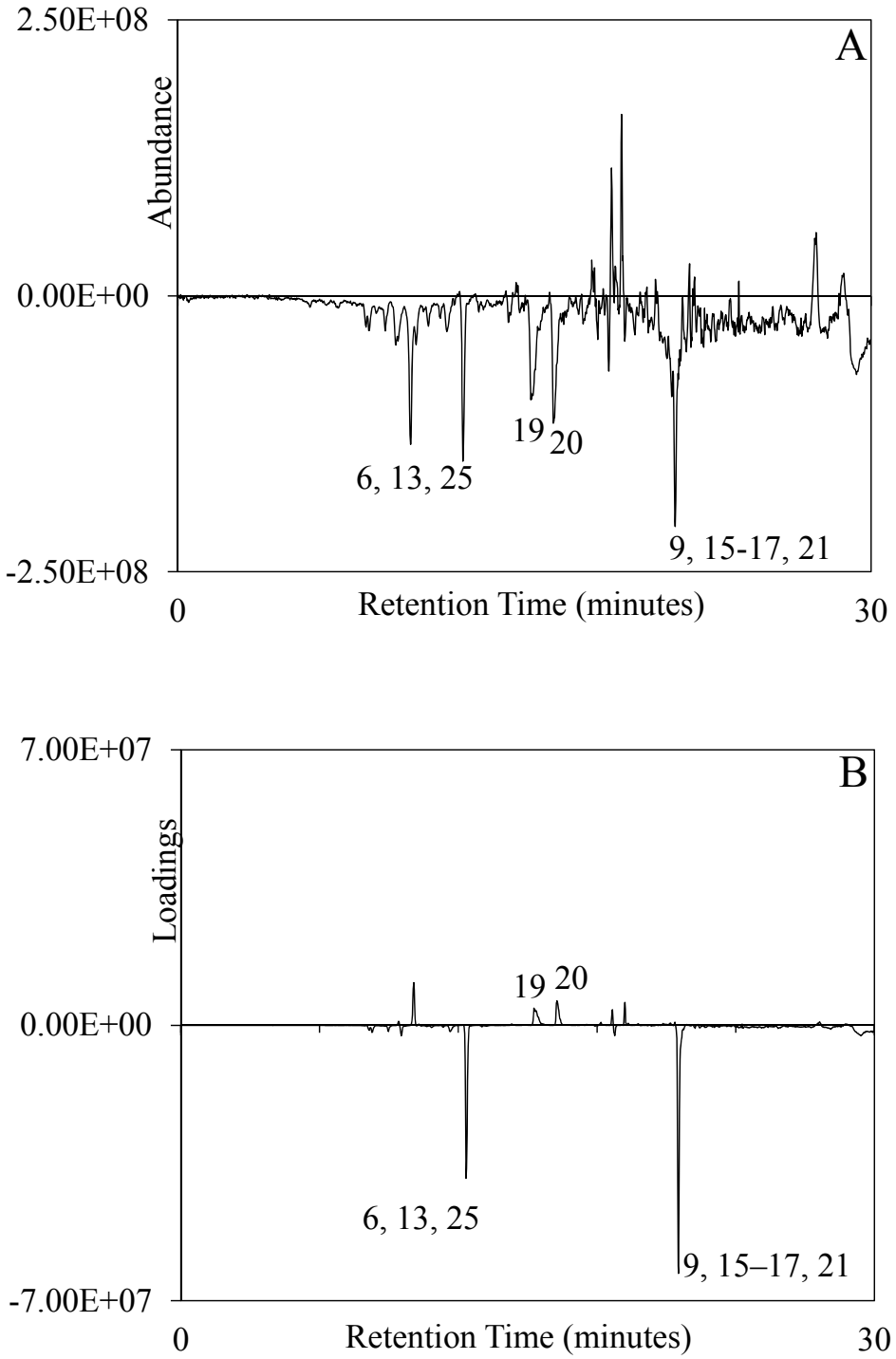

Figure 4.10: (A) Mean-centered total ion chromatogram of *S. nemorosa* and (B) loadings for *S. nemorosa* on principal component 2 based on chemical profiles obtained by liquid chromatography-mass spectrometry with 6=salvinorin A, 9=unidentified compound, 13=terpenoid, 15=vitamin E, 16=unidentified compound, 17=unidentified compound, 19=carnosol and carnosic acid, 20=methyl carnosate/rosmanol/epirosmanol, and 21=unidentified compound.

compounds loaded negatively while the phenolic diterpenes loaded positively. Overall, the negative loading compounds contributed more than those loading positively such that *S. nemorosa* was positioned negatively on PC2.

S. splendens was positioned negatively on PC1 in the scores plot also due to the process of mean centering. In the mean-centered data, phenolic diterpenes from *S. officinalis* were introduced and contributed negatively. When multiplied by the eigenvector for PC1, these compounds loaded negatively, which resulted in the negative positioning of *S. splendens* on PC1.

S. splendens was also positioned negatively on PC2 due to the process of mean centering (Figure 4.11). In the mean-centered data for *S. splendens,* a phenolic methyl ether and unidentified compound from *S. splendens* had a negative contribution at the same retention time as salvinorin A and the unidentified compound from *S. divinorum*, respectively. The compounds in the TIC of *S. divinorum* were more abundant than the compounds at the same retention time in the TIC of *S. splendens.* As a result, the average abundance of all TICs at these retention times was greater than the abundance at these retention times in *S. splendens*, which caused the compounds at these retention times to have a negative contribution in the mean-centered data. These compounds were weighted positively in the eigenvector for PC2 such that, when multiplied by the mean-centered data, these compounds were loaded negatively. This resulted in the negative positioning of *S. splendens* on PC2.

As observed in the scores plot, *S. nemorosa* and *S. splendens* were positioned closely on both PC1 and PC2. This close positioning was due to chemical similarities between these two species. Since *S. nemorosa* and *S. splendens* contained fewer and less abundant compounds than the other three species, compounds from the other species were introduced into the

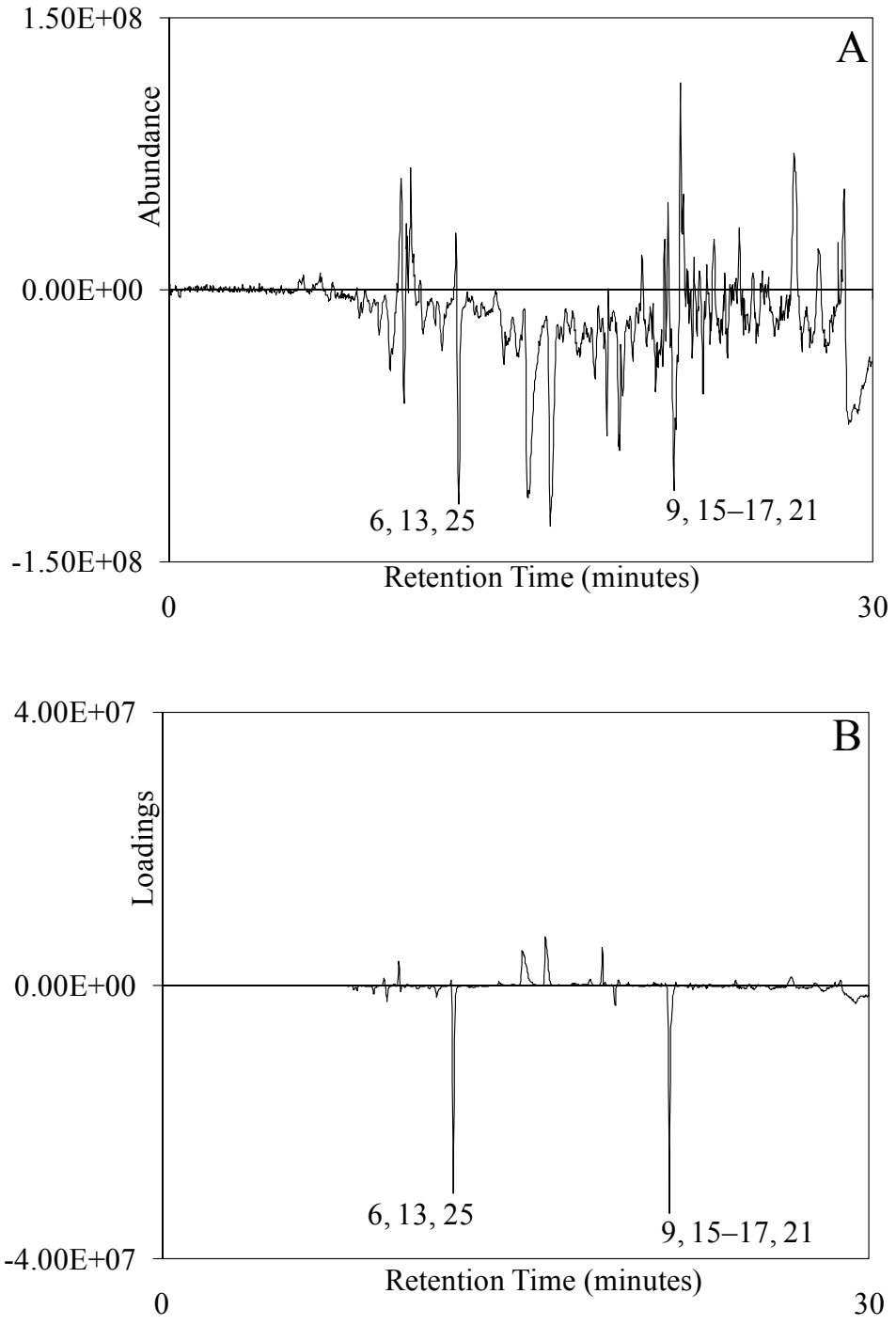

Figure 4.11: (A) Mean-centered total ion chromatogram of *S. splendens*, and (B) loadings for *S. splendens* on principal component 2 based on chemical profiles obtained by liquid chromatography-mass spectrometry with 6=salvinorin A, 9=unidentified compound, 13=terpenoid, 15=vitamin E, 16=unidentified compound, 17=unidentified compound, 21=unidentified compound, and 25=phenolic methyl ether and unidentified compound.

mean-centered data for *S. nemorosa* and *S. splendens*. Therefore, when the mean-centered data for replicates of *S. nemorosa* and for replicates of *S. splendens* were multiplied by the eigenvector for PC1, many of the same compounds were present. For example, both the loadings for *S. nemorosa* and for *S. splendens* contained a terpenoid from *S. guaranitica*, salvinorin A from *S. divinorum*, and phenolic diterpenes from *S. officinalis*. The close positioning of *S. nemorosa* and *S. splendens* on PC2 could be explained in a similar manner.

In addition, there was minimal spread among replicates of *S. nemorosa* and *S. splendens* in the scores plot; however, as previously stated, both these species had lower PPMC coefficients due to misalignments of a sesquiterpenoid, and oleanolic and/or ursolic acid. This contradiction between PPMC coefficients and PCA was reconciled when considering the differences between these two procedures as previously discussed in Chapter 3, section 3.3.3. In addition, PPMC coefficients were low because of the misalignments of oleanolic and/or ursolic acid and the sesquiterpenoid. However, in general, these compounds did not contribute greatly to the positioning of *S. nemorosa* and *S. splendens* on the scores plot.

S. guaranitica was positioned negatively on PC1 in the scores plot due to the process of mean centering. In the mean-centered data for *S. guaranitica*, a terpenoid had a positive contribution while phenolic diterpenes introduced into the mean-centered data from *S. officinalis*, as well as a sesquiterpenoid, had a negative contribution. When the mean-centered data were multiplied by the eigenvector for PC1, all these compounds loaded negatively, which resulted in the negative positioning of *S. guaranitica* on PC1 in the scores plot.

S. guaranitica was also positioned negatively on PC2 in the scores plot. The positioning of *S. guaranitica* could be explained using the mean-centered data. In the mean-centered data for *S. guaranitica*, some compounds from *S. guaranitica* contributed positively while others

contributed negatively. In addition, two phenolic diterpenes from *S. officinalis* were introduced into the mean-centered data and contributed negatively. When the mean-centered data were multiplied by the PC2 eigenvector, the negative loading compounds contributed more than those loading positively, which resulted in the negative positioning of *S. guaranitica* on PC2.

To statistically assess differentiation of *S. divinorum* from the other four *Salvia* species, two-tailed Student's *t*-tests were performed on the mean scores for both PC1 and PC2 at the 99% confidence level. Variances between all comparisons were assumed to be unequal. The null hypothesis, H_0, tested was that the mean scores of any two species were equal while the alternative hypothesis, H_1, stated that the mean scores of any two species were statistically different. Table 4.7 shows the absolute value of the calculated *t* values and critical *t* values for comparison of *S. divinorum* and the other four *Salvia* species on PC1 and PC2 [23]. For all comparisons, the absolute value of $t_{\text{calculated}}$ was greater than t_{critical}. Therefore, Student's *t*-tests showed *S. divinorum* was statistically differentiated from the remaining four species on both PC1 and PC2, at the 99% confidence level.

Due to the close proximity of *S. guaranitica, S. nemorosa,* and *S. splendens* in the scores plot, Student's *t*-tests were also performed on each pair-wise comparison of these species to determine if they could be statistically differentiated (Table 4.8) [23]. For all these comparisons, the absolute value of $t_{\text{calculated}}$ was greater than t_{critical}. Thus, despite close positioning in the scores plot, *S. guaranitica, S. nemorosa,* and *S. splendens* were statistically differentiated from each other on PC1 and PC2 at the 99% confidence level.

Table 4.7: Absolute value of $t_{calculated}$ and $t_{critical}$ values calculated using two-tailed Student's *t*-tests performed on the mean scores of principal component 1 (PC1) and principal component 2 (PC2) for comparisons (n=18) of *S. divinorum* to the other *Salvia* species*

Comparisons of *S. divinorum* to	PC1			PC2		
	$\lvert t_{calculated} \rvert$	$t_{critical}$	Significant Difference	$\lvert t_{calculated} \rvert$	$t_{critical}$	Significant Difference
S. guaranitica	41.29	2.947	Yes	38.47	3.106	Yes
S. nemorosa	11.97	2.947	Yes	34.83	3.106	Yes
S. officinalis	32.89	3.355	Yes	25.47	2.921	Yes
S. splendens	26.85	3.250	Yes	33.95	3.355	Yes

*99% confidence interval [23]

Table 4.8: Absolute value of $t_{calculated}$ and $t_{critical}$ values calculated from Student's *t*-tests performed on the mean scores of principal component 1 (PC1) and principal component 2 (PC2) between *S. guaranitica*, *S. nemorosa,* and *S. splendens* (n=18)*

Comparisons	PC1			PC2		
	$\lvert t_{calculated} \rvert$	$t_{critical}$	Significant Difference	$\lvert t_{calculated} \rvert$	$t_{critical}$	Significant Difference
S. nemorosa to *S. splendens*	7.380	3.250	Yes	7.210	3.106	Yes
S. nemorosa to *S. guaranitica*	23.72	2.977	Yes	5.266	2.947	Yes
S. guaranitica to *S. splendens*	28.07	3.169	Yes	14.61	3.106	Yes

*2-tailed *t*-test calculated at the 99% confidence interval [23]

Therefore, *S. divinorum* was statistically differentiated from the remaining four species in the scores plot on both PC1 and PC2 at the 99% confidence level using both the GC-MS data (Chapter 3, section 3.3.3) and LC-MS data. Using the LC-MS data, *S. guaranitica, S. nemorosa,* and *S. splendens* were all statistically differentiated in the scores plot on both PC1 and PC2 at a 99% confidence level; however, this did not occur when extracts the same species were analyzed by GC-MS. Therefore, the non-volatile compounds analyzed and detected by LC-MS offered greater discrimination of the species than the more volatile compounds analyzed and detected by GC-MS.

Although Student's *t*-tests demonstrated that *S. divinorum* could be differentiated from four related *Salvia* species on PC1 and PC2 at the 99% confidence level, the scores plot only accounted for 82% of the total variance. Therefore, PPMC coefficients, which assess similarity between pairs of samples based on a point-by-point comparison, were calculated between each *Salvia* species.

Mean PPMC coefficients calculated for the comparison of *S. divinorum* to *S. guaranitica* and to *S. officinalis* were less than 0.5, which indicated weak correlation (Table 4.9). For the comparison of *S. divinorum* to *S. nemorosa* and to *S. splendens*, mean PPMC coefficients were between 0.5 and 0.8, which indicated moderate correlation. These correlations were due to chemical differences between *S. divinorum* and the other four species. For example, *S. divinorum* was the only species to contain salvinorins A, B, C, D, and F. In addition, there was only one compound, a sesquiterpenoid, which was common among all five *Salvia* species. *S. divinorum* and *S. splendens* also both contained an unidentified compound, which resulted in the higher PPMC coefficients between these species than between *S. divinorum* and the other three species.

Table 4.9: Mean, standard deviation, and range of Pearson product moment correlation (PPMC) coefficients for comparisons (n=81) between *Salvia* species

Comparisons	Mean PPMC coefficient ± standard deviation	Range of PPMC coefficient
S. divinorum to *S. guaranitica*	0.4068 ± 0.0581	0.2788–0.5098
S. divinorum to *S. nemorosa*	0.5469 ± 0.0762	0.3750–0.6893
S. divinorum to *S. officinalis*	0.3762 ± 0.0721	0.2506–0.5177
S. divinorum to *S. splendens*	0.6432 ± 0.0602	0.5348–0.7561
S. guaranitica to *S. nemorosa*	0.4330 ± 0.0656	0.3119–0.5843
S. guaranitica to *S.officinalis*	0.2534 ± 0.0592	0.1417–0.3739
S. guaranitica to *S. splendens*	0.5937 ± 0.0601	0.4653–0.7119
S. nemorosa to *S. officinalis*	0.6126 ± 0.0880	0.4479–0.7783
S. nemorosa to *S. splendens*	0.7628 ± 0.0418	0.6816–0.8662
S. officinalis to *S. splendens*	0.5048 ± 0.0773	0.3733–0.6237

Mean PPMC coefficients calculated based on the LC-MS data were below 0.8, while mean PPMC coefficients calculated based on the GC-MS data (Chapter 3, section 3.3.3) were greater than 0.8. However, the internal standard, added to all five species analyzed using GC-MS, increased the correlation between the species. Therefore, this issue needs to be further investigated in the future.

In addition, the mean PPMC coefficient and range of PPMC coefficient were calculated between *S. nemorosa* and *S. splendens* since these species were positioned closest to one another on the scores plot. These species had a mean PPMC coefficient of 0.7628 ± 0.0418, which indicated a moderate correlation. However, the range of PPMC coefficients (0.6816–0.8662) between these species indicated a moderate to strong correlation due to chemical similarities. These species both contained a sesquiterpenoid, oleanolic and/or ursolic acid, and vitamin E. Also, these species contained fewer and less abundant additional compounds compared to the other three *Salvia* species.

PPMC coefficients showed that *S. splendens* and *S. nemorosa* were moderate to strongly correlated, which was supported by the close positioning of these species in the scores plot. However, as previously stated, these two species were differentiated on both PCs using Student's *t*-tests. This apparent contradiction was reconciled when considering the differences between these procedures. PPMC coefficients were calculated between two TICs based on a point-by-point comparison of all data points. Likewise, PCA was calculated using entire TICs. However, Student's *t*-tests were calculated on the mean scores of each species on PC1 and PC2; therefore, essentially minimizing the comparison between the two species to a one point comparison on both PCs. Therefore, this research showed that despite similarities among these

species, which was shown with PPMC coefficients and PCA scores plot, Student's *t*-tests were helpful in discriminating these two species.

In conclusion, TICs of *S. divinorum* were visually discriminated from the other four *Salvia* species; however, more salvinorin compounds were identified in the extracts analyzed using LC-MS than GC-MS. PPMC coefficients showed that *S. divinorum* was only weakly to moderately correlated to the other four *Salvia* species based on the non-volatile compounds. These results supported those observed using PCA and Student's *t*-tests, which showed *S. divinorum* was visually and statistically discriminated from the four other *Salvia* species. Thus, it was possible to distinguish *S. divinorum* from the four other *Salvia* species based on the non-volatile compounds analyzed and detected using LC-MS. With the GC-MS data (Chapter 3, section 3.3.3), *S. divinorum* was distinguished from the four other *Salvia* species based on PCA and Student's *t*-tests; however, PPMC coefficients did not support this conclusion due to the presence of the internal standard in all five species. Therefore, based on TICs, PPMC coefficients, PCA, and Student's *t*-tests, the chemical profiles generated using the LC-MS data were more discriminatory than those generated using the GC-MS data.

4.4 Summary

S. divinorum and related *Salvia* species were extracted by rotary agitation for 16 hours with acetonitrile and analyzed by LC-ESI-MS using positive ionization mode. Following analysis, TICs were truncated, retention time aligned, and normalized prior to data analysis. PPMC coefficients were calculated on TICs for replicate injections of the same extract and for all replicates of each species. In general, moderate to strong correlations among injection and extraction replicates indicated acceptable precision in the extraction, instrumental analysis, and data pretreatment procedures.

S. divinorum was visually discriminated from the other four *Salvia* species based on TICs. Using PCA, each *Salvia* species was visually discriminated from *S. divinorum* in the scores plot. Student's *t*-tests, calculated on the mean scores on both PCs, showed *S. divinorum* was statistically differentiated from the four other *Salvia* species at the 99% confidence level. PPMC coefficients were also calculated between *S. divinorum* and the other *Salvia* species. Weak to moderate correlations were observed, which indicated chemical differences among *S. divinorum* and the other four *Salvia* species. Using TICs, PCA, Student's *t*-tests, and PPMC coefficients, it was possible to distinguish *S. divinorum* from the other four *Salvia* species. Therefore, based on TICs, PCA, Student's *t*-tests, and PPMC coefficients, the chemical profiles generated using the LC-MS data were more discriminatory than those generated using the GC-MS data.

References

References

[1] Perrone A, Muzashvili T, Napolitano A, Skhirtladze A, Kemertelidze E, Pizza C, Piacente S. Steroidal glycosides from the leaves of *Ruscus colchicus*: Isolation and structural elucidation based on a preliminary liquid chromatography-electrospray ionization tandem mass spectrometry profiling. Phytochemistry 2009; 70:2078–88.

[2] 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; 20:131–6.

[3] Tsujikawa K, Kuwayama K, Miyaguchi H, Kanamori T, Iwata YT, Yoshida T, Inoue H. Determination of salvinorin A and salvinorin B in *Salvia divinorum*-related products circulated in Japan. Forensic Science International 2008; 180:105–9.

[4] Sana TR, Fischer S, Wohlgemuth G, Katrekar A, Jung K, Ronald PC, Fiehn O. Metabolomic and transcriptomic analysis of the rice response to the bacterial blight pathogen *Xanthomonas oryzae* pv. *oryzae*. Metabolomics 2010; 6:451–65.

[5] Fiehn O, Wohlgemuth G, Scholz M, Kind T, Lee DY, Lu Y, Moon S, Nikolau B. Quality control for plant metabolomics: reporting MSI-complaint studies. The Plant Journal 2008; 53:691–704.

[6] Kennedy JH, Wiseman JM. Direct analysis of *Salvia divinorum* leaves for salvinorin A by thin layer chromatography and desorption electrospray ionization multi-stage tandem mass spectrometry. Rapid Commun Mass Sp 2010; 24:1305–11.

[7] Díaz-Cruz MS, López de Alda MJ, López R, Barceló D. Determination of estrogens and progestogens by mass spectrometric techniques (GC/MS, LC/MS and LC/MS/MS). Journal of Mass Spectrometry 2003; 38:917–23.

[8] Razboršek MI, Vončina DB, Doleček V, Vončina E. Determination of oleanolic, betulinic and ursolic acid in *Lamiaceae* and mass spectral fragmentation of their trimethylsilylated derivatives. Chromatographia 2008; 67:433–40.

[9] Shen D, Pan M, Wu Q, Park C, Juliani HR, Ho C et al. LC-MS method for the simultaneous quantitation of the anti-inflammatory constituents in oregano (*Origanum* Species). Journal of agricultural and food chemistry 2010; 58:7119–25.

[10] Fai YM, Tao CC. A review of presence of oleanolic acid in natural products. 2009;1–271.

[11] Djarmati Z, Jankov RM, Djordjević A, Ribar B, Lazar D, Engel P. Carnosic acid 12-methyl ether-y-lactone, a ferruginol-type diterpene from *S. officinalis.* Phytochemistry 1992; 31:1307–9.

[12] Chen H, Talaty NN, Takáts Z, Cooks RG. Desorption electrospray ionization mass spectrometry for high-throughput analysis of pharmaceutical samples in the ambient environment. Analytical Chemistry 2005; 77:6915–27.

[13] Lu Y, Foo LY. Polyphenolics of *Salvia*-a review. Phytochemistry 2002; 59:117–40.

[14] Ollanketo M, Peltoketo A, Hartonen K, Hiltunen R, Riekkola M. Extraction of sage (*Salvia officinalis L.*) by pressurized hot water and conventional methods: antioxidant activity of the extracts. European Food Research and Technology 2002; 215:158–63.

[15] Cuvelier ME, Berset C, Richard H. Antioxidant constituents in sage (*Salvia officinalis).* Journal of Agricultural and Food Chemistry 1994; 42:665–9.

[16] Lamien-Meda A, Neil M, Lohwasser U, Börner A, Fraanz C, Novak J. Investigation of antioxidant and rosmarinic acid variation in the sage collection of the genebank in Gatersleben. Journal of Agricultural and Food Chemistry 2010; 58:3813–19.

[17] Cavero S, Jaime L, Martín-Álvarez PJ, Señoráns FJ, Reglero G, Ibañez E. In vitro antioxidant analysis of supercritical fluid extracts from rosemary (*Rosmarinus officinalis L.*). European Food Research and Technology 2005; 221:478–86.

[18] Walker JB, Sytsma KJ, Treutlein J, Wink M. *Salvia* (Lamiaceae) is not monophyletic: implications for the systematics, radiation, and ecological specializations of *Salvia* and tribe Mentheae. American Journal of Botany 2004; 91:1115–25.

[19] Bodnar Willard MA. Forensic analysis of *Salvia divinorum* and related *Salvia* species using chemometric procedures. MS Thesis (2010).

[20] Tychonievich, J. *Salvia* clades. Personal Communication. 2009.

[21] Abreu ME, Müller M, Alegre L, Munné-Bosch S. Phenolic diterpene and α-tocopherol contents in leaf extracts of 60 *Salvia* species. Journal of the Science of Food and Agriculture 2008; 88:2648–53.

[22] Schwarz K, Ternes W. Antioxidative constituents of *Rosmarinus officinalis* and *Salvia officinalis*. Z Lebensm Unters Forsch 1992; 195:99–103.

[23] Adam C. Essential mathematics and statistics for forensic science. Hobken, NJ: Wiley-Blackwell 2010.

Chapter 5: Conclusions and Future Work

5.1 Gas Chromatography-Mass Spectrometry Analysis of *Salvia divinorum* and Related *Salvia* Species

Salvia divinorum and four related *Salvia* species were extracted in dichloromethane and analyzed using gas chromatography-mass spectrometry (GC-MS). All total ion chromatograms (TICs) were then subjected to data pretreatment procedures (*i.e.* background correction, smoothing, retention time alignment, and normalization) to reduce the non-chemical sources of variance. Data analysis procedures (*i.e.* Pearson product moment correlation (PPMC) coefficients, principal components analysis (PCA), and Student's *t*-tests) were performed to assess the discrimination of *S. divinorum* from the related species based on the volatile compounds extracted from each. All five species contained common compounds; however, *S. divinorum* was the only species containing salvinorins A, B, and D. PPMC coefficients calculated between replicates of each species showed moderate to strong correlation, which indicated acceptable precision in the extraction, instrumental analysis, and data pretreatment procedures. PPMC coefficients also showed moderated to strong correlation between *S. divinorum* and the other *Salvia* species. However, the internal standard added to all five species increased the correlation between the species. Therefore, this issue needs to be further investigated in the future. Despite some chemical similarities, *S. divinorum* was visually and statistically discriminated from the four related *Salvia* species on the PCA scores plot, which accounted for a total of 77% of the variance, using Student's *t*-tests at the 99% confidence level. The apparent contradiction between the PPMC and PCA results can be explained by considering how PPMC coefficients and the PCA scores plot were calculated.

5.2 Liquid Chromatography-Mass Spectrometry Analysis of *Salvia divinorum* and Related *Salvia* Species

S. divinorum and four related *Salvia* species were also extracted in acetonitrile and analyzed using liquid chromatography-mass spectrometry (LC-MS) in positive ionization mode. Once again, TICs were subjected to data pretreatment procedures (*i.e.* retention time alignment and normalization) and the same data analysis procedures were performed. Common compounds were shared among some species; however, *S. divinorum* was the only species containing salvinorins A, B, C, D, and F. PPMC coefficients were calculated between replicates of each species and showed moderate to strong correlation, which indicated acceptable precision in the extraction, instrumental analysis, and data pretreatment procedures. PPMC coefficients calculated between *S. divinorum* and the four other *Salvia* species showed weak to moderate correlation. In addition, *S. divinorum* was visually and statistically discriminated from the four related *Salvia* species on the PCA scores plot, which accounted for a total of 82% of the variance, using Student's *t*-tests at the 99% confidence level.

5.3 Comparison of Gas Chromatography-Mass Spectrometry and Liquid Chromatography-Mass Spectrometry for the Analysis of *Salvia divinorum* and Related *Salvia* Species

GC-MS is limited to the analysis of volatile compounds, while LC-MS can also be used to analyze non-volatile compounds. In addition, LC-MS^n was previously shown to be more sensitive than GC-MS; therefore, more non-volatile compounds are commonly detected when using LC-MS and/or LC-MS^n [1]. For example, salvinorins A, B, C, D, and F were identified in *S. divinorum* extracts analyzed using LC-MS, while only salvinorins A, B, and D were identified in *S. divinorum* extracts analyzed using GC-MS. As previously stated, PPMC coefficients calculated between replicates of each species analyzed using GC-MS and LC-MS both showed

moderate to strong correlation. This indicated acceptable precision in extraction and analysis for both techniques. Mean PPMC coefficients calculated between *S. divinorum* and the other *Salvia* species based on the volatile compounds were greater than 0.8, while mean PPMC coefficients calculated based on more polar non-volatiles were below 0.8. However, the internal standard increased the correlation between the species analyzed using GC-MS. Therefore, further investigation of the use of the internal standard is needed in the future. Using PCA and Student's *t*-tests, *S. divinorum* was visually and statistically discrimination from the four related *Salvia* species on the scores plot based on both the GC-MS and LC-MS data. Overall, the chemical profiles generated using the LC-MS data were more discriminatory than those generated using the GC-MS data based on TICs, PPMC coefficients, PCA, and Student's *t*-tests.

5.4 Utility of Gas Chromatography-Mass Spectrometry and Liquid Chromatography-Mass Spectrometry for the Analysis of *Salvia divinorum*

GC-MS, which is widely available in forensic laboratories, and LC-MS can be used to identify *S. divinorum* based on the presence of salvinorin A [2]. Identification of *S. divinorum* in this manner is possible since salvinorin A is only known to be found in *S. divinorum* thus far [3].

Identification of *S. divinorum* based on a full chemical fingerprint using LC-MS is necessary in states were only *S. divinorum* is regulated and not salvinorin A, if salvinorin A is later identified in other plants. However, the chemical composition of any plant material can vary over time and based on the location where the plant is cultivated [4, 5]. Thus, the analyst must be aware of this variability when identifying *S. divinorum* using a chemical fingerprint. Therefore, both techniques are useful in analyzing *S. divinorum*.

5.5 Future Work

Further work can always improve previous research. In this research, *S. divinorum* and four related *Salvia* species were extracted and analyzed to determine if *S. divinorum* could be discriminated from four other *Salvia* species based on chemical fingerprints generated by GC-MS and LC-MS. The extraction and analysis procedures in this research can be further investigated. Firstly, different extraction solvents can be investigated. In this research, dichloromethane and acetonitrile were used. It is polarity and temperature of the solvent that determines which compounds are extracted from the plant material. Thus, other solvents or mixture of solvents may extract more compounds from *S. divinorum* than observed in this research. For example, the use of acetone, toluene, hexane, or mixtures of these solvents can be investigated. Identifying more compounds in *S. divinorum* provides more information that can be used to discriminate *S. divinorum* from other plant material.

Since GC-MS is limited to the analysis of volatile compounds, extracts containing non-volatile compounds can be derivatized to modify compound volatility. Therefore, future work may include the investigation of an extraction procedure with a dertivatization step so that more compounds in *S. divinorum* can be detected using GC-MS. For example, the use of bistrimethylsilyltrifluoroacetamide (BSTFA), and N-methyl-trimethylsilyltrifluoroacetamide (MSTFA) as derivatizating reagents can be investigated. Thus, a more complete chemical fingerprint of *S. divinorum* may be generated.

In this research, progesterone was used as the internal standard. However, as progesterone is not stable in solution, the internal standard solution had to be prepared daily and added to each extract. This caused variability in the abundance of the internal standard in the resulting TICs, which affected subsequent data analysis procedures. Therefore, if an internal

standard is needed in future research, using a compound that is more stable over time is beneficial. For example, the use 5-α-androstane, and cholesterol as an internal standard can be investigated. Furthermore, the addition of the internal standard to the extracts analyzed by GC-MS caused *Salvia* species to be more strongly correlated to one another based on PPMC coefficients. Therefore, future work may include the investigation of the use of an internal standard.

To further improve this research, different statistical procedures can be investigated. This can include the investigation of soft supervised statistical procedures, such as soft independent modeling by class analogy, which objectively identify an unknown extract based on how closely it is associated to known extracts. The benefit of this type of statistical procedure is if the unknown extract is not associated closely to any of the known extracts, an identification is not forced. Therefore, this statistical procedure may be useful in a forensic laboratory.

In conclusion, both techniques are useful in analyzing *S. divinorum* and discriminating *S. divinorum* from four other *Salvia* species. However, a more complete chemical fingerprint is generated using LC-MS. Therefore, if an unknown plant material is submitted to a forensic laboratory and can not be identified by GC-MS, an alternative procedure for identifying the plant material based on several compounds using LC-MS should be considered.

References

References

[1] 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; 20:131–6.

[2] Watson JT, Sparkman OD. Introduction to mass spectrometry: instrumentation, applications and strategies for data interpretation. 4th ed. West Sussex, England: John Wiley & Sons Ltd, 2007.

[3] Giroud C, Felber F, Augsburger M, Horisberger B, Rivier L, Mangin P. *Salvia divinorum:* an hallucinogenic mint which might become a new recreational drug in Switzerland. Forensic Sci Int 2000; 112:143–50.

[4] Siebert, DJ. Localization of salvinorin A and related compounds in glandular trichomes of the psychoactive sage, *Salvia divinorum*. Annals of Botany 2004; 93:763–71.

[5] Carrer RP, Vanderlinde R, Dutra S, Marcon A, Echeverrigaray S. Esstential oil varation among Brazilian accessions of *Salvia guaranitica* L.. Flavour Fragr. J. 2007; 22:430–4.

APPENDICES

APPENDIX A

PROVISIONAL IDENTIFICATION OF COMPOUNDS IN THE FIVE SALVIA SPECIES ANALYZED USING GAS CHROMATOGRAPHY-MASS SPECTROMETRY

Table A.1: Provisional Identification of Compounds in *S. divinorum* Analyzed Using Gas Chromatography-Mass Spectrometry

Compound Number	Approximate Retention Time (minutes)	Molecular Mass (*m/z*)	Major Fragment Ions (*m/z*) (Order of decreasing abundance)	Provisional Identification of Compound and/or Compound Class*
1	23.24	380	71, 57, 97, 85, 83, 84, 113, 194, 127, 154	Heptacosane
2	24.03	410	69, 81, 95, 77, 94, 137, 387, 295, 80, 154	Squalene
3	24.47	314	124, 314, 272, 229, 91, 79, 105, 107, 93, 133	Progesterone
4	24.61	408	71, 57, 85, 99, 83, 177, 55, 118, 125, 141	Nonacosane
5	24.82	**	69, 204, 96, 81, 53, 107, 91, 164, 109, 97	Acetate
6	25.94	436	71, 57, 85, 56, 96, 99, 97, 443, 295, 127	Hentriacontane
7	26.17	390	94, 291, 237, 166, 304, 356, 476, 305, 107, 77	Salvinorin B
8	26.89	432	94, 273, 432, 121, 95, 107, 166, 220, 91, 404	Salvinorin A
9	27.17	478	71, 57, 85, 55, 97, 99, 113, 56, 70, 69	Tetratriacontane
10	27.29	432	94, 179, 372, 105, 97, 206, 400, 77, 237, 383	Salvinorin D
11	27.44	414	397, 91, 268, 414, 145, 273, 95, 60, 81, 159	Sitosterol
12	28.35	492	57, 71, 85, 206, 55, 99, 97, 463, 134, 88	Pentatriacontane

*Provisional identification based on library searches and literature sources
**Unknown molecular mass

Table A.2: Provisional Identification of Compounds in *S. guaranitica* Analyzed Using Gas Chromatography-Mass Spectrometry

Compound Number	Approximate Retention Time (minutes)	Molecular Mass (*m/z*)	Major Fragment Ions (*m/z*) (Order of decreasing abundance)	Provisional Identification of Compound and/or Compound Class*
2	24.03	410	69, 81, 105, 95, 342, 93, 68, 149, 77, 68	Squalene
3	24.43	314	124, 314, 272, 229, 91, 79, 105, 107, 133, 77	Progesterone
6	25.94	436	57, 71, 85, 99, 55, 113, 69, 83, 97, 127	Hentriacontane
9	27.18	478	57, 71, 85, 99, 55, 83, 69, 111, 72, 141	Tetratriacontane
11	27.46	414	414, 396, 329, 303, 399, 381, 159, 95, 107, 213	Sitosterol
13	9.91	134	105, 77, 134, 76, 51, 106, 63, 50, 104, 89	Phthalide
14	11.75	204	161, 91, 105, 204, 79, 120, 77, 93, 81, 162	β-Cubebene
15	25.52	340	143, 340, 202, 128, 129, 95, 142, 157, 203, 141	Ester
16	25.65	338	160, 141, 115, 145, 129, 128, 91, 95, 94, 159	Alkene
17	26.18	430	430, 165, 431, 164, 205, 166, 432, 55, 121, 57	Vitamin E
18	26.45	338	141, 94, 95, 81, 357, 145, 338, 128, 142, 222	Dione
19	26.66	340	340, 95, 94, 133, 82, 91, 128, 189, 77, 105	Dione
20	26.83	340	135, 94, 121, 81, 91, 340, 95, 77, 129, 105	Nitrolepidine
21	27.30	399	164, 94, 81, 95, 67, 329, 77, 121, 161, 491	Benzenesulfonamide
22	27.80	426	218, 203, 219, 105, 95, 94, 121, 189, 81, 204	β-Amyrin
23	28.09	426	218, 145, 136, 122, 161, 55, 203, 109, 189, 426	α-Amyrin

*Provisional identification based on library searches and literature sources

Table A.3: Provisional Identification of Compounds in *S. nemorosa* Analyzed Using Gas Chromatography-Mass Spectrometry

Compound Number	Approximate Retention Time (minutes)	Molecular Mass (*m/z*)	Major Fragment Ions (*m/z*) (Order of decreasing abundance)	Provisional Identification of Compound and/or Compound Class*
3	24.45	314	124, 314, 272, 229, 91, 79, 105, 93, 77, 107	Progesterone
4	24.63	408	57, 71, 85, 55, 97, 83, 69, 56, 41, 86	Nonacosane
6	25.95	436	57, 71, 85, 55, 69, 83, 99, 97, 56, 70	Hentriacontane
9	27.19	478	57, 7, 85, 55, 97, 99, 83, 69, 56, 113	Tetratriacontane
11	27.47	414	414, 107, 93, 97, 396, 303, 255, 55, 329, 145	Sitosterol
14	11.76	204	161, 105, 204, 79, 73, 120, 69, 93, 80, 91	β-Cubebene
17	26.19	430	430, 165, 431, 164, 166, 205, 57, 121, 69, 203	Vitamin E
24	10.98	204	93, 69, 79, 9, 95, 77, 55, 161, 175, 94	β-Caryophyllene
25	15.68	123	95, 55, 123, 68, 82, 403, 221, 488, 98, 57	Alkene
26	20.57	213	72, 69, 59, 96, 97, 126, 82, 165, 56, 114	Amide
27	26.25	328	313, 328, 153, 285, 299, 69, 314, 181, 268, 329	Phenyl
28	26.58	478	57, 71, 85, 97, 99, 55, 192, 56, 83, 69	11-n-Decyltetracosane
29	28.37	506	57, 71, 96, 97, 85, 206, 69, 179, 339, 82	Hextriacontane

*Provisional identification based on library searches and literature sources

Table A.4: Provisional Identification of Compounds in *S. officinalis* Analyzed Using Gas Chromatography-Mass Spectrometry

Compound Number	Approximate Retention Time (minutes)	Molecular Mass (*m/z*)	Major Fragment Ions (*m/z*) (Order of decreasing abundance)	Provisional Identification of Compound and/or Compound Class*
1	23.24	380	57, 71, 85, 99, 55, 113, 84, 141, 69, 83	Heptacosane
2	24.07	410	69, 81, 67, 68, 287, 122, 255, 123, 70, 94	Squalene
3	24.47	314	124, 314, 272, 229, 91, 79, 105, 107, 133, 93	Progesterone
4	24.65	408	71, 57, 85, 97, 99, 113, 124, 72, 182, 169	Nonacosane
6	25.97	436	57, 71, 85, 97, 99, 113, 69, 55, 83, 56	Hentriacontane
9	27.21	478	57, 71, 85, 99, 97, 69, 55, 83, 113, 127	Tetratriacontane
11	27.48	414	414, 213, 396, 399, 303, 159, 97, 145, 57, 95	Sitosterol
17	26.21	430	165, 430, 164, 431, 166, 205, 432, 57, 121, 206	Vitamin E
23	28.11	426	97, 218, 161, 87, 189, 69, 203, 95, 81, 217	α-Amyrin
25	15.70	123	95, 68, 123, 55, 97, 82, 69, 57, 124, 83	Alkene
29	28.40	506	57, 71, 85, 55, 69, 99, 83, 113, 70, 97	Hexatriacontane
30	3.95	136	93, 91, 92, 77, 79, 121, 94, 136, 80, 105	α-Pinene
31	4.19	136	93, 121, 79, 107, 91, 67, 77, 94, 136, 95	Camphene
32	4.56	136	93, 69, 91, 77, 79, 80, 92, 67, 94, 121	β-Pinene
33	4.60	136	93, 69, 208, 92, 94, 77, 67, 53, 70, 79	β-Myrcene
34	5.32	154	81, 154, 108, 139, 111, 84, 71, 93, 69, 83	Eucalyptol
35	6.43	152	110, 81, 109, 68, 95, 67, 69, 82, 79, 55	Thujone
36	6.60	152	110, 81, 95, 109, 67, 69, 68, 70, 55, 79	Thujone

*Provisional identification based on library searches and literature sources

Table A.4 (cont'd)

Compound Number	Approximate Retention Time (minutes)	Molecular Mass (*m/z*)	Major Fragment Ions (*m/z*) (Order of decreasing abundance)	Provisional Identification of Compound and/or Compound Class*
37	7.10	152	95, 81, 108, 152, 109, 69, 83, 55, 67, 110	Camphor
38	7.45	154	95, 110, 139, 93, 55, 67, 79, 111, 84, 69	Borneol
39	9.06	196	95, 93, 121, 136, 92, 108, 80, 91, 110, 67	Bornyl Acetate and/or Isoborynyl Acetate
40	9.93	168	153, 108, 107, 93, 109, 111, 69, 55, 99, 79	Bicyclic Compound
41	11.00	164	93, 69, 91, 79, 120, 107, 105, 106, 67, 95	Cyclopropane Compound
42	11.47	204	93, 80, 121, 92, 91, 107, 94, 79, 77, 109	α-Caryophyllene
43	13.20	222	109, 161, 93, 105, 69, 107, 91, 122, 204, 81	Epiglobulol
44	13.39	220	109, 138, 67, 93, 123, 95, 68, 55, 81, 82	Diene
45	18.05	290	137, 95, 81, 257, 93, 69, 123, 107, 109, 79	Epimanool
46	21.09	330	286, 230, 215, 243, 218, 271, 204, 217, 287, 203	Isocarnosol
47	21.37	286	271, 286, 272, 128, 287, 215, 229, 141, 228, 135	Phenyl
48	21.75	286	286, 215, 271, 204, 287, 202, 187, 141, 216, 272	Dione
49	26.44	422	57, 71, 85, 56, 97, 55, 99, 69, 70, 83	Squalane
50	26.98	422	57, 71, 85, 99, 97, 55, 69, 83, 113, 56	Triacontane
51	27.66	450	57, 71, 85, 56, 97, 55, 83, 99, 69, 113	Dotriacontane
52	27.81	478	57, 71, 85, 83, 69, 99, 55, 113, 70, 97	9-n-Octylhexacosane
53	29.72	456	203, 218, 97, 415, 282, 133, 95, 149, 135, 119	Oleanolic Acid

*Provisional identification based on library searches and literature sources

Table A.5: Provisional Identification of Compounds in *S. splendens* Analyzed Using Gas Chromatography-Mass Spectrometry

Compound Number	Approximate Retention Time (minutes)	Molecular Mass (*m/z*)	Major Fragment Ions (*m/z*) (Order of decreasing abundance)	Provisional Identification of Compound and/or Compound Class*
2	24.06	410	69, 81, 95, 136, 121, 68, 137, 123, 93, 67	Squalene
3	24.46	314	124, 229, 272, 314, 91, 79, 105, 93, 107, 95	Progesterone
4	24.63	408	57, 71, 85, 99, 69, 55, 113, 56, 83, 111	Nonacosane
6	25.96	436	57, 71, 85, 97, 99, 69, 55, 83, 113, 56	Hentriacontane
9	27.20	478	57, 71, 85, 55, 99, 83, 69, 97, 113, 70	Tetratriacontane
11	27.48	414	414, 81, 107, 105, 145, 95, 57, 55, 303, 93	Sitosterol
17	26.20	430	165, 430, 164, 431, 166, 205, 57, 55, 69, 97	Vitamin E
23	28.12	426	218, 189, 123, 107, 203, 136, 219, 122, 79, 109	α-Amyrin
25	15.70	123	95, 68, 82, 123, 57, 55, 97, 71, 69, 83	Alkene
29	28.38	506	57, 71, 85, 97, 55, 99, 69, 83, 56, 113	Hextriacontane
53	29.71	456	207, 203, 73, 133, 59, 179, 206, 125, 208, 189	Oleanolic Acid
54	26.42	342	94, 91, 95, 117, 342, 105, 121, 81, 77, 145	8-Episalviarin
55	27.03	406	189, 91, 110, 340, 97, 129, 95, 81, 143, 105	Cardenolide

*Provisional identification based on library searches and literature sources

APPENDIX B

PROVISIONAL IDENTIFICATION OF COMPOUNDS IN THE FIVE SALVIA SPECIES ANALYZED USING LIQUID CHROMATOGRAPHY-MASS SPECTROMETRY

Table B.1: Provisional Identification of Compounds in *S. divinorum* Analyzed Using Liquid Chromatography-Mass Spectrometry

Compound Number	Approximate Retention Time (minutes)	Molecular Mass (*m/z*)	Molecular Mass + Hydrogen (*m/z*)	Adduct Ions (*m/z*)	Major Fragment Ions (*m/z*) (Order of decreasing abundance)	Major Ions of Enhanced Product Ion 1 (*m/z*)	Major Ions of Enhanced Product Ion 2 (*m/z*)	Provisional Identification of Compound and/or Compound Class*
1	9.53	390	391	413, 423, 803	373, 391, 374, 355, 259, 413, 327, 392, 341, 414	(of 391) 145, 105, 119, 107, 131, 91, 173, 117, 143, 121	(of 366) 142, 193, 423, 199, 91, 105, 119, 145, 213, 195	Salvinorin B
2**	10.13	406	407	424, 835	424, 429, 341, 373, 407, 329, 413, 389, 445, 214	(of 424) 157, 293, 309, 355, 183, 209, 179, 195, 185, 265	(of 407) 191, 149, 105, 107, 133, 147, 173, 223, 90, 104	Terpenoid
3**	10.32	378	379	396, 401	311, 214, 371, 401, 496, 343, 312, 361, 279, 158	(of 524) 157, 309, 355, 185, 156, 291, 337, 183, 235	(of 373) 191, 129, 209, 323, 401, 91, 107, 115, 121, 127	Terpenoid
4	10.73	432	433	455	373, 341, 391, 214, 374. 323, 455, 279, 413, 355	(of 373) 121, 341, 145, 91, 105, 231, 107, 213, 223, 129	(of 341) 105, 145, 91, 109, 117, 173, 103, 121, 147, 157	Salvinorin D
5	11.20	374	375	397, 771	325, 357, 326, 397, 214, 358, 331, 279, 87, 158	(of 325) 325, 105, 145, 179, 297, 91, 223, 117, 143, 119	(of 331) 201, 279, 105, 128, 149, 93, 105, 155, 173, 397	Salvinorin F

*Provisional identification based on literature sources

**Possibly two or more compounds coeluting at this retention time

Table B.1 (cont'd)

Compound Number	Approximate Retention Time (minutes)	Molecular Mass (*m/z*)	Molecular Mass + Hydrogen (*m/z*)	Adduct Ions (*m/z*)	Major Fragment Ions (*m/z*) (Order of decreasing abundance)	Major Ions of Enhanced Product Ion 1 (*m/z*)	Major Ions of Enhanced Product Ion 2 (*m/z*)	Provisional Identification of Compound and/or Compound Class*
6	12.09	432	433	455, 887	373, 374, 455, 341, 313, 355, 314, 375, 295, 456	(of 331) 295, 135, 121, 107, 91, 105, 131, 187, 277, 145	(of 408) 91, 107, 121, 145, 187, 199, 135, 131, 211, 221	Salvinorin A
7	12.69	474	475	497, 971	415, 497, 475, 305, 261, 416, 355, 279, 214, 323	(of 279) 155, 156, 157, 142, 105, 131, 145, 129, 121, 141	(of 499) 497, 219, 95, 145, 147, 189, 194, 219, 498, 91	Salvinorin C
8	18.59	264	265	282, 304	282, 304, 283, 265, 247, 563, 305, 284, 266, 248	(of 313) 93, 97, 107, 95, 135, 121, 111, 100, 91, 282	(of 279) 282, 201, 247, 107, 265, 97, 177, 283, 304, 149	Sesquiterpenoid
9**	20.83	426	427	444, 449	409, 410, 449, 259, 327, 271, 285, 203, 450, 257	(of 429) 409, 107, 95, 109, 119, 105, 121, 133, 159, 91	(of 447) 335, 149, 105, 613, 91, 119, 95, 595, 93, 107	Unidentified Compound

*Provisional identification based on literature sources

**Possibly two or more compounds coeluting at this retention time

Table B.2: Provisional Identification of Compounds in *S. guaranitica* Analyzed Using Liquid Chromatography-Mass Spectrometry

Compound Number	Approximate Retention Time (minutes)	Molecular Mass (*m/z*)	Molecular Mass + Hydrogen (*m/z*)	Adduct Ions (*m/z*)	Major Fragment Ions (*m/z*) (Order of decreasing abundance)	Major Ions of Enhanced Product Ion 1 (*m/z*)	Major Ions of Enhanced Product Ion 2 (*m/z*)	Provisional Identification of Compound and/or Compound Class*
8	18.77	264	265	282, 304	282, 304, 283, 265, 247, 563, 305, 284, 348, 266	(of 282) 93, 91, 95, 107, 97, 121, 114, 135, 86, 95	(of 331) 270, 107, 298, 316, 331, 91, 93, 109, 301, 92	Sesquiterpenoid
10	9.56	340	341	363, 703	341, 363, 295, 342, 323, 364, 296, 324, 249, 343	(of 357) 95, 161, 165, 179, 141, 129, 115, 155, 128, 178	(of 293) 165, 128, 171, 215, 178, 117, 128, 131, 153, 95	Diterpenoid
11	10.16	398	399	421, 819	339, 399, 321, 421, 340, 293, 381, 322, 400, 357	(of 343) 178, 165, 179, 191, 128, 203, 207, 202, 141, 205	(of 279) 91, 129, 178, 128, 131, 179, 219, 115, 165, 141	Terpenoid
12**	10.36	340, 398	341, 399	421, 703	341, 295, 342, 323, 363, 421, 296, 324, 381, 343	(of 323) 237, 165, 141, 295, 169, 155, 179, 154, 129, 167	(of 295) 165, 155, 153, 179, 167, 169, 154, 178, 202, 203	Terpenoid
13	12.36	398	399	421, 819	339, 399, 421, 321, 340, 245, 293, 357, 400, 422	(of 345) 128, 155, 141, 129, 142, 178, 153, 179, 165, 115	(of 339) 179, 142, 178, 189, 165, 155, 178, 192, 115, 141	Terpenoid

*Provisional identification based on literature sources

**Possibly two or more compounds coeluting at this retention time

Table B.2 (cont'd)

Compound Number	Approximate Retention Time (minutes)	Molecular Mass (*m/z*)	Molecular Mass + Hydrogen (*m/z*)	Adduct Ions (*m/z*)	Major Fragment Ions (*m/z*) (Order of decreasing abundance)	Major Ions of Enhanced Product Ion 1 (*m/z*)	Major Ions of Enhanced Product Ion 2 (*m/z*)	Provisional Identification of Compound and/or Compound Class*
14**	18.22	456	457	479, 935	479, 439, 191, 480, 440, 411, 203, 393, 501, 331	(of 302) 479, 95, 91, 93, 105, 107, 109, 121, 134, 435	(of 280) 203, 119, 105, 133, 95, 107, 93, 121, 147, 109	Oleanolic and/or Ursolic Acid
15	21.46	430	431	453	431, 430, 429, 432, 445, 331, 301, 397, 433, 446	(of 331) 165, 137, 109, 119, 166, 121, 431, 122, 107, 205	(of 495) 165, 137, 139, 494, 107, 91, 109, 93, 105, 149	Vitamin E
16	21.74	***	***	***	429, 522, 430, 301, 331, 523, 239, 185, 397, 141	(of 429) 637, 165, 107, 105, 91, 93, 105, 145, 187, 121	(of 409) 522, 697, 298, 270, 523, 524, 301, 93, 525, 95	Unidentified Compound

*Provisional identification based on literature sources

**Possibly two or more compounds coeluting at this retention time

***Unknown molecular mass, molecular mass + hydrogen, and adduct ions

Table B.3: Provisional Identification of Compounds in *S. nemorosa* Analyzed Using Liquid Chromatography-Mass Spectrometry

Compound Number	Approximate Retention Time (minutes)	Molecular Mass (*m/z*)	Molecular Mass + Hydrogen (*m/z*)	Adduct Ions (*m/z*)	Major Fragment Ions (*m/z*) (Order of decreasing abundance)	Major Ions of Enhanced Product Ion 1 (*m/z*)	Major Ions of Enhanced Product Ion 2 (*m/z*)	Provisional Identification of Compound and/or Compound Class*
8	18.63	264	265	282, 304	282, 283, 304, 563, 247, 265, 564, 284, 305, 248	(of 282) 93, 91, 107, 95, 135, 97, 121, 111, 100, 114	(of 191) 91, 93, 97, 135, 95, 109, 121, 304, 83, 94	Sesquiterpenoid
14**	18.19	456	457	479, 935	479, 439, 480, 191, 440, 411, 457, 481, 203, 936	(of 245) 119, 203, 479, 105, 133, 107, 95, 93, 191, 147	(of 263) 119, 107, 133, 203, 105, 95, 121, 109, 191, 93	Oleanolic and/or Ursolic Acid
15	21.33	430	431	453	431, 430, 432, 429, 445, 895, 411, 329, 896, 413	(of 329) 165, 137, 119, 166, 431, 268, 109, 167, 136	(of 730) 599, 237, 597, 595, 895, 268, 317, 237	Vitamin E
17	20.91	***	***	***	429, 430, 447, 469, 397, 329, 960, 954, 448, 470	(of 469) 165, 149, 177, 137, 163, 429, 191, 205, 119, 135	(of 869) 573, 469, 149, 595, 163, 317, 107, 165, 261, 119	Unidentified Compound

*Provisional identification based on literature sources

**Possibly two or more compounds coeluting at this retention time

***Unknown molecular mass, molecular mass + hydrogen, and adduct ions

Table B.4: Provisional Identification of Compounds in *S. officinalis* Analyzed Using Liquid Chromatography-Mass Spectrometry

Compound Number	Approximate Retention Time (minutes)	Molecular Mass (*m/z*)	Molecular Mass + Hydrogen (*m/z*)	Adduct Ions (*m/z*)	Major Fragment Ions (*m/z*) (Order of decreasing abundance)	Major Ions of Enhanced Product Ion 1 (*m/z*)	Major Ions of Enhanced Product Ion 2 (*m/z*)	Provisional Identification of Compound and/or Compound Class*
8	18.62	264	265	282, 304	282, 283, 304, 265, 247, 563, 284, 305, 266, 248	(of 282) 93, 91, 121, 95, 107, 97, 111, 135, 240, 109	(of 304) 91, 107, 93, 95, 304, 313, 133, 92, 105, 109	Sesquiterpenoid
14**	18.12	456	457	479, 935	439, 479, 191, 440, 285, 411, 480, 457, 301, 203	(of 439) 203, 119, 107, 105, 95, 133, 147, 121, 189, 109	(of 285) 256, 91, 93, 107, 103, 285, 95, 121, 105, 131	Oleanolic and/or Ursolic Acid
15	21.30	430	431	453	431, 430, 432, 429, 285, 301, 205, 433, 331, 445	(of 589) 165, 599, 595, 137, 895, 317, 109, 119, 261, 107	(of 569) 285, 283, 527, 191, 215, 243, 117, 569, 115, 131	Vitamin E
18	10.24	300	301	323	301, 302, 323, 303, 214, 297, 324, 271, 158, 279	(of 271) 286, 301, 168, 140, 287, 121, 258, 285, 112, 271	(of 301) 286, 153, 271, 105, 119, 131, 140, 153, 168, 301	3-O-Methylkaempferol
19**	15.02	330, 332	331, 333	354, 355, 683, 685	287, 205, 331, 288, 245, 285, 191, 355, 333, 353	(of 331) 115, 117, 215, 103, 197, 145, 285, 267, 225, 131	(of 301) 103, 217, 149, 191, 115, 131, 128, 129, 205, 145	Carnosol and Carnosic Acid

*Provisional identification based on literature sources

**Possibly two or more compounds coeluting at this retention time

Table B.4 (cont'd)

Compound Number	Approximate Retention Time (minutes)	Molecular Mass (*m/z*)	Molecular Mass + Hydrogen (*m/z*)	Adduct Ions (*m/z*)	Major Fragment Ions (*m/z*) (Order of decreasing abundance)	Major Ions of Enhanced Product Ion 1 (*m/z*)	Major Ions of Enhanced Product Ion 2 (*m/z*)	Provisional Identification of Compound and/or Compound Class*
20**	15.96	346	347	369, 715	301, 219, 205, 302, 220, 369, 206, 347, 231, 303	(of 219) 231, 148, 205, 103, 115, 131, 128, 129, 219, 162	(of 205) 103, 102, 148, 130, 128, 115, 369, 323, 91, 147	Methyl Carnosate, Rosmanol and/or Epirosmanol
21	20.00	***	***	***	285, 551, 331, 287, 301, 273, 569, 205, 191, 552	(of 285) 119, 105, 133, 145, 93, 91, 107, 121, 143, 173	(of 331) 197, 215, 285, 267, 169, 225, 243, 165, 157, 191	Unidentified Compound

*Provisional identification based on literature sources

**Possibly two or more compounds coeluting at this retention time

***Unknown molecular mass, molecular mass + hydrogen, and adduct ions

Table B.5: Provisional Identification of Compounds in *S. splendens* Analyzed Using Liquid Chromatography-Mass Spectrometry

Compound Number	Approximate Retention Time (minutes)	Molecular Mass (*m/z*)	Molecular Mass + Hydrogen (*m/z*)	Adduct Ions (*m/z*)	Major Fragment Ions (*m/z*) (Order of decreasing abundance)	Major Ions of Enhanced Product Ion 1 (*m/z*)	MS^2 Major Ions of EPI 2 (*m/z*)	Provisional Identification of Compound and/or Compound Class*
8	18.61	264	265	282, 304	282, 304, 283, 247, 265, 563, 305, 284, 248, 266	(of 479) 93, 97, 95, 91, 107, 121, 111, 109, 100, 114	(of 443) 331, 107, 313, 301, 91, 93, 271, 331, 94, 95	Sesquiterpenoid
9	21.03	426	427	444, 449	409, 410, 449, 427, 444, 285, 259, 450, 327, 411	(of 409) 107, 119, 105, 409, 95, 109, 121, 91, 93, 133	(of 285) 105, 107, 613, 91, 119, 109, 93, 151, 159, 189	Unidentified Compound
14**	18.14	456	457	479, 935	439, 479, 191, 440, 480, 411, 457, 203, 393, 458	(of 302) 119, 107, 133, 105, 203, 95, 121, 91, 109, 191	(of 245) 119, 107, 479, 105, 95, 121, 93, 133, 147, 163	Oleanolic/ Ursolic Acid
15	21.32	430	431	453	431, 430, 432, 429, 425, 445, 433, 447, 407, 426	(of 430) 165, 430, 164, 137, 121, 205, 136, 119, 109, 91	(of 445) 119, 147, 175, 107, 123, 137, 139, 163, 445, 95	Vitamin E
22	9.85	400	401	423, 823	341, 339, 401, 423, 383, 342, 399, 357, 340, 421	(of 267) 341, 129, 117, 179, 105, 115, 157, 165, 128, 142	(of 303) 179, 109, 157, 341, 105, 129, 143, 178, 91, 95	Diterpenoid

*Provisional identification based on literature sources

**Possibly two or more compounds coeluting at this retention time

Table B.5 (cont'd)

Compound Number	Approximate Retention Time (minutes)	Molecular Mass (*m/z*)	Molecular Mass + Hydrogen (*m/z*)	Adduct Ions (*m/z*)	Major Fragment Ions (*m/z*) (Order of decreasing abundance)	Major Ions of Enhanced Product Ion 1 (*m/z*)	MS^2 Major Ions of EPI 2 (*m/z*)	Provisional Identification of Compound and/or Compound Class*
23	10.04	342	343	365, 707	343, 365, 344, 325, 297, 251, 231, 279, 707, 326	(of 325) 117, 129, 157, 115, 131, 128, 142, 91, 143, 141	(of 365) 192, 157, 165, 166, 167, 179, 199, 203, 207, 217	Diterpenoid
24	10.21	362	363	***	305, 363, 287, 323, 275, 364, 306, 269, 288, 231	(of 325) 363, 91, 105, 107, 95, 119, 117, 129, 121, 93	(of 365) 105, 91, 119, 93, 95, 131, 133, 147, 117, 145	Diterpenoid
25**	12.10	406, 548	407, 549	572, 587	311, 566, 293, 549, 312, 329, 371, 571, 345, 265	(of 345) 284, 312, 345, 330, 91, 115, 131, 117, 128, 105	(of 311) 217, 171, 107, 131, 145, 311, 199, 189, 265, 229	Phenolic Methyl Ether and Unidentified Compound

*Provisional identification based on literature sources

**Possibly two or more compounds coeluting at this retention time

***Unknown adduct ions