INVESTIGATION OF GENOTYPE/CHEMOTYPE CORRELATIONS IN *CANNABIS SATIVA*

By

LINDSEY ALLEN

Bachelor of Science in Organismic Biology

Northeastern State University

Broken Arrow, OK

2011

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
December 2015

INVESTIGATION OF GENOTYPE/CHEMOTYPE

CORRELATIONS IN *CANNABIS SATIVA*

Thesis Approved:

Dr. Gerwald Koehler, Ph. D.

Thesis Adviser

Bruce A. Benjamin, Ph. D.

Committee Member

Ronald R. Thrasher, Ph. D.

Committee Member

Mrs. Jane Pritchard

Committee Member

ACKNOWLEDGEMENTS

First, I would like to thank those who have significantly contributed to my research. The Tulsa Police Department Evidence Custodian, Staff Sergeant Kim Lee assisted in enabling my selection of their seized *C. sativa* samples. Director of the Tulsa Police Department, Tara Valouch helped tremendously in the establishment of my GC-MS technique, as well as helping me obtain and weigh samples from the Tulsa Police Department. Dr. Mark Payton I owe gratitude for helping calculate my statistical analysis. Dr. Jun Fu provided much help with the instrumentation of the biology student laboratories. Lindsey Yoder I would also like to thank for performing the Ion Torrent PGM sequencing that I analyzed. Lastly, I would like to thank Dr. Robert Allen for proposing that I consider researching the current investigation as well as the continual insight during the research.

Second, I would like to thank the professors who have encouraged me the most to continue in my academic career and strive for success. I would like to begin by thanking Dr. Dianna Spencer, who encouraged me to acknowledge and apply my passion for biotechnology. I would also like to thank my mentor, Dr. Randy Wymore, for the guidance he has provided throughout my years in graduate school. Dr. Gerwald Koehler has been an excellent adviser with my current study, and has taught me a lot about the application of techniques in molecular biology.

Lastly, I would like to acknowledge my parents for both providing me with

Acknowledgements reflect the views of the author and are not endorsed by committee members or Oklahoma State University.

continuous opportunities for growth throughout my life as well as instilling in me the values and virtues I strive to possess and reflect through life. I could not have made it this far in my life without their unending support and the advice they have given. I will always be proud that I am their daughter, and hope to continue to provide goodness to the world as they have done.

There are many others whom I know I am neglecting to mention who have supported me throughout my graduate studies, and who I greatly appreciate. It is an honor to have been given the opportunity to study at OSU-CHS School of Forensic Sciences, and I can only hope that my knowledge will be beneficial in my future endeavors and in the lives of others.

Acknowledgements reflect the views of the author and are not endorsed by committee members or Oklahoma State University.

Name: LINDSEY N. ALLEN

Date of Degree: DECEMBER, 2015

Title of Study: INVESTIGATION OF GENOTYPE/CHEMOTYPE CORRELATIONS IN *CANNABIS SATIVA*

Major Field: FORENSIC SCIENCE

Abstract: Though *Cannabis sativa* has been used as a medical treatment for centuries, recent American political controversy surrounding the legalization of marijuana has generated interest into the mechanism of cannabinoid biosynthesis. This biosynthetic pathway involves enzymes derived from genes which produce the major cannabinoids of interest within *C. sativa*. Δ9-Tetrahydrocannabinol (THC) has been predominant in investigations due to numerous reported beneficial effects for various symptoms such as those associated with cancer treatment. Non-synonymous single nucleotide polymorphisms (SNPs) within the THCA synthase gene responsible for enzymatic production of THC change the nucleotide sequence and subsequent amino acid sequence for the enzyme. These changes could potentially alter the efficacy of the enzyme that produces THC. In the current investigation, permission was granted by Oklahoma State University to handle small samples of *C. sativa* within the School of Forensic Science student laboratories. The Tulsa Police Department provided small samples of *C. sativa* seized previously and marked "to be destroyed". In these samples, four single nucleotide polymorphisms within the THCA synthase gene were analyzed via SNaPshot® analysis and amplicon sequencing. The concentration of THC was determined using gas chromatography-mass spectrometry. Correlation statistics were performed to determine if a correlation existed between the concentration of THC and the activity of the THCA synthase gene. No correlation existed between THCA synthase SNP profiles and THC concentration. If a correlation existed, genetic analysis of the *C. sativa* THCA synthase would provide growers, providers, distributors, and users of medical marijuana with more pertinent information about the quality of *C. sativa* products available.

TABLE OF CONTENTS

Chapter Page

I. INTRODUCTION 1

Research Proposal 5

II. REVIEW OF LITERATURE 1

Legislation of *Cannabis sativa* 6
Previous Legislation 6
Recent Legislation 7
Medical Use of *Cannabis sativa* 8
Diseases 8
Complications 10
Cannabinoids of *Cannabis sativa* 11
Biosynthesis of Cannabinoids 11
Cannabidiol 14
Δ9-Tetrahydrocannabinol 15
Analysis of *Cannabis sativa* 16
Chemotype 16
Genotype 18
Summary 19

III. METHODOLOGY 20

Overview 20
Selection of *Cannabis sativa* Samples 20
Genetic Analysis 21
DNA Extraction 21
Primer Selection and Optimization 23

Chapter Page

III. METHODOLOGY ....(con't)

Genetic Analysis ....(con't)
Polymerase Chain Reaction 24
Single Nucleotide Polymorphism Analysis 24
Optimization of SNaPshot® 25
THCA Synthase Amplicon Sequencing..27
Chemical Analysis via GC/MS 28
Preparation of Chromatography..28
Cannabinoid Extraction from Plant Material..29
Chemical Analysis 30
Statistical Analysis..30
Correlation Statistics 30
Summary 31

IV. RESULTS AND INTERPRETATION 32

Overview..32
Genetic Analysis 32
Results of SNaPshot®...32
Results of THCA Synthase Amplicon Sequencing 36
Gas Chromatography-Mass Spectrometry 40
Results of Gas Chromatography-Mass Spectrometry.......................................40
Statistical Analysis...45
Results of Analysis of Variance...45
Results of Correlation Statistics...46

V. DISCUSSION AND CONCLUSION...48

Overview...48
Cannabis sativa Sampling 48
SNaPshot® Analysis...49
Amplicon Sequencing Analysis..50
GC-MS Analysis...51
Correlation Analysis 52
Comparison of Results..52
Possible Future Studies..54
Conclusion 54

REFERENCES 55

APPENDICES 61
Appendix A: SNaPshot® Data 63
Appendix B: GC/MS Data 66
Appendix C: Statistical Data 70

LIST OF TABLES

Table Page

1 *Cannabis sativa* Primer Pairs....23
2 SNaPshot® Baseline Threshold Optimization....26
3 SNaPshot® Baseline Threshold Calculations....27
4 THCA Synthase SNP Nucleotide Incorporation....34
5 Changes in Amino Acid Sequence at SNP Positions....39
6 Limit of Detection and Limit of Quantitation Calculations....40
7 GC-MS Cannabinoid Controls....42
8 Comparison of SNP Activity with Percent THC....45
9 Pearson's Correlation Coefficient Results....47

LIST OF FIGURES

Figure Page

1 Cannabinoid Biosynthetic Pathway 12
2 Inflorescence of a Mature *Cannabis sativa* Plant 13
3 Primer Alignment in C2E2 Amplicon of THCA Synthase Gene 33
4 Chromatogram of a Homozygous-Active Sample 34
5 Chromatogram of a Heterozygous Sample 35
6 Chromatogram of a Homozygous-Inactive Sample 35
7 Sequencing Results as a THCA Synthase Alignment 37
8 Sequencing Results in the C2E2 Amplicon Region 38
9 Percentage of Nucleotide Incorporation at the 16R SNP Position 39
10 GC-MS Standard Curve for CBD, THC, and Cannabinol 41
11 Calculation of Percent THC from a *C. sativa* Sample 42
12 Range of Percent THC within *C. sativa* Samples 43
13 Comparison of Percent THC and THCA Synthase SNP Activity 44
14 ANOVA Results Comparing Percent THC and SNP Activity 46

CHAPTER I

INTRODUCTION

The controversy surrounding the legalization of marijuana has existed for many years and remains a polarizing topic in American society and especially politics.(1) Prohibition of marijuana, or *Cannabis sativa L (C. sativa)*, began with the passage of state legislation in the early 1900s.(2) On a federal level, *C. sativa* is currently listed as a Schedule I controlled substance under the Comprehensive Drug Abuse Prevention and Control Act of 1970.(3) Since its initial listing in 1970, *C. sativa* has become the most widely used illicit drug within the U.S.(1) Beginning in 1996, however, states have individually started decriminalizing the use of *C. sativa* for both medical and recreational purposes. Currently, 23 states and the District of Columbia have enacted legislation that permits the medical use of *C. sativa*.(4)

Although states within the U.S. have started enabling the medical use of *C. sativa*, a variety of legal and other issues continue to present problems. While some states permit the use of *C. sativa* for medical purposes, individuals receiving, prescribing, growing, and distributing the plant are still subject to punishment for violations of federal law and can therefore be charge with federal crimes.(1) Other problems include more medically relevant dilemmas about *C. sativa*, including the variation of dosage level within the plant, variation of individuals' unique

physiological response to the plant, and variation of types and levels of other chemical components within the plant.[1],[5]

C. sativa contains more than 60 chemical compounds, called cannabinoids.[6],[7] Of these, Δ9- tetrahydrocannabinol (THC) is the main psychoactive ingredient.[1] Another cannabinoid called cannabidiol (CBD), also present in *C. sativa*, may have various beneficial effects on individuals when used as a medical treatment for uncontrollable seizures.[8] Studies have shown that these cannabinoids may be helpful in treating a variety of illnesses ranging from Crohn's disease to fibromyalgia as well as ameliorating the pain and toxic effects of chemotherapy for a variety of patients suffering from malignant disease.[1],[9],[10] However, the concentration of the cannabinoids within *C. sativa* prescribed to these individuals is often unknown.[1]

A variety of factors influence the chemical composition of *C. sativa*: the geographical origin of the plant, the part of the plant selected for use, the way in which the plant parts are stored, and the specific growth conditions used during propagation of the plant.[7] Of the numerous techniques available to analyze the chemical composition of *C. sativa*, the majority involve chromatography. Chromatography separates chemical compounds based on mass or size, solubility, or attraction to chemical characteristics of a solid matrix.[11] The primary methods used to determine the type and quantity of cannabinoids present in samples of *C. sativa* are thin layer chromatography, high pressure liquid chromatography, and gas chromatography. Mass spectrometry can be paired with gas or liquid chromatography to identify and quantify the amounts of cannabinoids in extracts.[12] Often used in forensic investigations, these techniques provide information about the chemical composition of the sample in question.[11]

Other methods to assess the characteristics of *C. sativa* involve genetic analysis. In fact, a partially completed genome sequence of the *C. sativa* plant has been published in Gen Bank, although numerous gaps remain in the sequence.[13] The discovery of cannabinoids within *C.*

sativa sparked interest in the biological mechanism(s) used by the plant to synthesize these compounds biosynthetically. A complete understanding of cannabinoid biosynthesis remains unclear and continues to be investigated.[14] However, recent evidence suggests that THC and CBD are synthesized through the activity of two enzymes, THCA synthase and CBDA synthase respectively, from a precursor known as cannabigerolic acid (CBG) that is derived from a pathway involving olivetolic acid, one of the precursors of all cannabinoids.[15] After enzymatic activity occurs, the cannabinoids Δ9-tetrahydrocannabinolic acid (THCA) and cannabidiolic acid (CBDA) spontaneously decarboxylate into their neutral forms THC and CBD, respectively. Because of high interest in THC, the main psychoactive compound, research has focused on the enzyme THCA synthase that converts CBG to Δ9-tetrahydrocannabinolic acid or THCA, which then spontaneously is decarboxylated (especially during heating) to THC. The enzyme involved in the synthesis of THCA is expressed by the THCA synthase gene that has been identified and characterized by nucleotide sequencing. The THCA synthase gene contains no introns and essentially represents the coding sequence for the enzyme.[16] Several studies have identified specific single nucleotide polymorphisms (SNPs) within the THCA synthase gene that are claimed to correlate with the production of THC. [16],[14],[17] Some studies claim to be able to distinguish "active" from "inactive" THCA synthase genes based upon the SNPs present in the gene and the chemical cannabinoid content of the mature plant. [16],[17]

Variations within the genes that code for synthesis of THCA have been analyzed. These variations, or polymorphisms, occur at specific nucleotide positions within the gene. One published study revealed that *C. sativa* can have up to 62 single-nucleotide polymorphisms in the THCA synthase gene.[16] Furthermore, the study showed that these mutations in the THCA synthase gene can either create premature stop codons, or change up to 37 amino acids within the primary sequence.[16] The changes in amino acid sequence in the THCA synthase gene can affect the activity of the THCA synthase enzyme.[17] Rotherham and Harbison, for example, analyzed

four specific single-nucleotide polymorphisms (SNPs) in the THCA synthase gene.[17] The authors claimed that these specific SNPs correlated with THCA synthase activity and THC content in the mature plant. Thus, their results demonstrated that these SNPs determined the level of activity of THCA synthase within individual *C. sativa* plants, as revealed by the THC content at maturity. Samples were found to genotype as homozygous for the active form of THCA synthase, heterozygous for both active and inactive forms, or homozygous for the inactive form of THCA synthase, and these genotypes correlated with the chemical content of THC in the plants.[17]

The study by Rotherham and Harbison created an opportunity for further investigation of the relationship between single-nucleotide polymorphisms and their effects on the activity of THCA synthase, and hence the levels of cannabinoids present in *C. sativa*. In the case of their study, forensically seized samples were obtained in New Zealand. Similarly, the research presented here assessed the genotype and chemotype relationship in *C. sativa* samples seized in the Tulsa area by the Tulsa Police Department. Thus, the genotype/chemotype relationships in plants possibly originating from a different geographical location of the world could be investigated. In addition, the work of Rotherham and Harbison focused upon SNPs existing within a 399 base-pair region of the 1635 base-pair THCA synthase gene, and was concerned with only four SNPs known to cause significant changes in the amino acid sequence of the synthase polypeptide.[17] Thus, additional SNPs elsewhere within the gene could be examined. Lastly, the Rotherham and Harbison study distinguished what constituted "drug-type" and "fiber-type" samples by analyzing the level of cannabinoids by gas chromatography-mass spectrometry (GCMS). The specific quantities of cannabinoids in each sample were correlated to the results from genetic analysis. The potential correlation between these two analyses could be useful for the standardization of *C. sativa* used in medical practice.[17]

Research Proposal

The proposed research will investigate *C. sativa* for the reported correlation between the quantities of the various cannabinoids in the chemical composition of the mature plant and the genotype of the THCA synthase gene. This correlation would enable the prediction of the final composition of cannabinoids in a mature plant based on the genetic analysis of seeds, pollen, or young plants before maturity which is the time cannabinoid biosynthesis is maximal. This information would be useful for growers to be confident of the characteristics of the plant they will ultimately harvest for medical purposes. To determine the chemotype (independent variable), I will use gas chromatography with mass spectrometry to analyze cannabinoids present in *C. sativa* plant material.[(11)] To determine the genotype of the same samples (dependent variable), I will perform genomic analysis of 4 specific single-nucleotide polymorphisms in the THCA synthase gene directly responsible for the synthesis of THCA; this enzyme is involved in producing one of the principal cannabinoids in *C. sativa*.[(17)]

The results of this study could confirm the correlation between genotype and chemotype of *C. sativa*, initially reported by Rotherham and Harbison.[(17)] This confirmation could allow genotypic analysis of seeds or other plant material in a predictive manner concerning the cannabinoid content of mature plants. This information would enable growers, medicinal providers, dispensaries, and users of medicinal marijuana to choose the quality of desired plant product based on the genetic blueprint for production of cannabinoids by *C. sativa* seeds.

The following chapter will review the available literature about *C. sativa*. This review will include a discussion of the legislation of *C. sativa*, the medical uses of *C. sativa*, the cannabinoids in *C. sativa*, as well as the techniques available to analyze *C. sativa*.

CHAPTER II

LITERATURE REVIEW

As discussed previously, controversial political debates increasingly focus on the use of medical and recreational marijuana. *C. sativa* is also the most commonly used illicit drug in the world.[(6),(15)] For thousands of years people have used *Cannabis sativa L. (C. sativa)*, or marijuana, for a number of reasons including recreational and medicinal purposes.[(8), (15), (18)] However, because of its psychotropic effects, among other factors, *C. sativa* is currently listed as a federally controlled substance within the United States.[(3)] People experience these pharmacologic effects because of the specific chemical compounds of *C. sativa* called cannabinoids.[(8)] The following review will discuss literature concerned with legislation of *C. sativa*, the previously demonstrated medicinal treatments using *C. sativa*, the primary cannabinoids found in strains of *C. sativa*, as well as the justification for further analysis of the cannabinoids within *C. sativa*.

Legislation of *Cannabis sativa*

Previous Legislation

Regulation of *C. sativa* in the United States began in the early 1900s from a state legislation perspective. Until this time, the use of *C. sativa* for medicinal purposes was legal. However, New York and Utah enacted legislation that restricted the use or sale of *C. sativa*.

Soon after, many other states also enacted legislation that restricted the sale or use of marijuana. The shift from state regulation to federal regulation occurred with the passage of two important Acts.[(2)] The Uniform Narcotic Drug Act, enacted in 1932, prohibited anyone from "manufacturing, possessing, selling, purchasing, prescribing, administering, or giving away any narcotic drug."[(19)] The Marihuana Tax Act, enacted in 1937, required any manufacturer, distributor, prescriber, vendor, or individual who gives away marijuana to register and pay taxes in order to provide their service.[(19)] These two Acts were the first to federally enforce the regulation and restriction of sale and use of *C. sativa*.

Years later, a prevalence of narcotic abusers motivated the public to call for reform. Thus in 1951, the United States Congress passed the Boggs Act, that listed marijuana as an illicit substance along with other narcotics for the first time. The Boggs Act also called for more strict consequences for abusers and sellers of narcotics and *C. sativa*. With the passage of the Narcotics Control Act of 1956, the penalties for the illegal activities increased again. Specifically for abusers and sellers of *C. sativa,* the increase would enable more control from law enforcement on restriction of use, sale, and manufacturing of marijuana. The Narcotics Control Act also increased the enforcement of narcotics regulation.[(2),1]

Recent Legislation

After nearly a century of state and federal legislation opposing the manufacture, possession, use, sale, and distribution of *C. sativa,* California became the first state to legalize its use for medical purposes in 1996.[(20)] The decriminalization of *C. sativa* has since spread

[1] When the Narcotics Drug Act of 1957 was passed, it enabled U.S. Customs and Narcotics Bureau agents to carry weapons, as well as arrest suspected violators without need of a warrant. The government could appeal judicial cases in which evidence was suppressed. It also provided witnesses with immunity in exchange for testimony. A new category of laws was enacted regarding communication of drug sale, trade, and trafficking to enable use of wiretapping. Furthermore, individuals found guilty of drug offenses were required to register with U.S. immigration authorities upon entrance or leaving the U.S. Illegal immigrants who were found guilty of drug offenses would also be deported under this new provision. See references: 2. Bonnie RJ, Whitebread II CH (1970).

throughout the United States. Currently, 23 states and the District of Columbia have legislation enabling the medical use of *C. sativa*.[4] Regardless, *C. sativa* remains listed as a schedule I federally controlled substance.[3] Therefore, although state legislature enables patients to use medicinal marijuana with authorization, federal law provides for enforcement of penalties for these individuals.[21]

Medical Use of *Cannabis sativa*

C. sativa has been used as a medication for centuries. Bostwick discusses previous uses of the plant; he states that a physician in the 1830s originally prescribed *C. sativa* for pain and vomiting, much like the medicinal uses today.[1] More modern research about the medical benefits of *C. sativa* is extensive. The best documentation about the use of the plant pertains to research involving *C. sativa* as a treatment for the side effects of chemotherapy.[8] Much of the initial research on *C. sativa* in the 1990s studied the reversal of weight loss seen in cancer patients as well as AIDS patients because of its appetite promoting effects.[1],[22]

Diseases

Scientists and medical experts have found that numerous symptoms of diseases and ailments other than cancer can be benefitted when patients are treated with *C. sativa*. In a comprehensive review, Gurley et al state that *C. sativa* alleviates symptoms such as nausea and vomiting, muscle spasms, loss of appetite, menstrual pain, pain associated with child birth, seizures, and anxiety. The effects of treatment using *C. sativa* have been studied in illnesses such as cholera and rabies, glaucoma, pain syndromes, addiction syndromes, and withdrawal syndromes.[8]

Other studies target specific illnesses such as Crohn's disease. A study by Naftali et al investigated the potential benefits of the treatment of Crohn's disease with *C. sativa*.[9] In a study

of 30 patients, every single patient stated that using *C. sativa* improved their symptoms. Specifically, results showed treatment with *C. sativa* decreased the number of bowel movements, the number of necessary prescribed treatments (including steroids), and the number of necessary surgeries. These results demonstrated a measurable significant improvement of the symptoms of Crohn's disease in 21 of the 30 patients who were treated with *C. sativa*.(9)

Along with Crohn's disease, medical use of *C. sativa* for other inflammatory bowel diseases has been investigated. Allegretti et al surveyed patients with IBD to assess the effects of the treatment on their symptoms.(23) In this study, the majority of patients described use of *C. sativa* as "very helpful" to or "completely relieving" of their abdominal pain, nausea, and loss of appetite. However, the patients noted the least amount of improvement in their diarrheal symptoms. Because most of the symptom showed improvement with the exception of diarrhea, results of the study show the physiological mechanism of *C. sativa* metabolism remains unclear. These results contribute to the fact that IBD has not been approved statewide for treatment with medicinal *C. sativa*.(23)

Another disease that is currently being researched for treatment with *C. sativa,* particularly for the effects on patients' pain, is fibromyalgia. One study performed assessments on 56 patients.(10) The results show that the patients experienced "significant relief of pain, stiffness, relaxation, somnolence and perception of well-being" after treatment with *C. sativa*. Thus, all symptoms assessed in this study showed significant improvement after administration of *C. sativa*.(10)

Epilepsy and other seizure disorders are conditions for which approval has been obtained in some states for treatment with *C. sativa*.(8) Studies by Welty et al and Szaflarski and Bebin have reviewed previous research involving the treatment of epilepsy with *C. sativa* and found that evidence supporting the claim is "scarce",(24) "anecdotal, weak, and occasionally

contradictory”.[18] However, numerous other ailments have been approved for treatment with *C. sativa*, including migraines, glaucoma, and arthritis.[4]

Complications

As much literature as is available about the beneficial effects of medical treatment using *C. sativa* for various illnesses, many articles describe the negative effects or drawbacks of *C. sativa* treatment. Pertaining to the legalization standards, one study notes that, regarding *C. sativa* in medical use:

> “[P]roducts are not produced under the guidance of good manufacturing practices (GMP) and are not subject to regulations governing labeling, purity, and reliability. In other words, there is no guarantee of consistency between products, or even differing lots produced by the same manufacturer.”[24]

Leung discussed other complications regarding treatment in another article.[5] This study showed that 1 in 10 people who have used *C. sativa* develop dependence. Leung also indicated that when people smoke *C. sativa,* levels of cannabinoid quantity vary with individual consumption based on unique physiology and technique of inhalation. Furthermore, the cannabinoid content of *C. sativa* can also vary based on the geographic origin of the plant, the part of the plant being used, the way the plant is stored, or the way in which the plant is grown. Thus, standardization of a prescription is difficult due to the titration, or unknown concentration, and dose efficacy of *C. sativa*.[5]

Many studies show overwhelming evidence of withdrawal symptoms after patients cease treatment with *C. sativa* [1],[5],[8],[18],[25],[26] The most common symptoms of withdrawal include sleep disturbance,[1],[18],[25],[26] changes in appetite,[18],[25],[26] irritability,[1],[25],[26] anxiety,[1],[26] weight loss,[1],[25] restlessness,[1],[25] cravings,[1],[18] and aggression.[18],[25] Szaflarski and Bebin report other symptoms of withdrawal from *C. sativa* treatment such as insomnia, delirium,

moodiness, somnolence, fatigue, and diarrhea.[18] Ramesh et al also note withdrawal symptoms of anger and sweat or chills.[25] Depression has also been investigated as a withdrawal effect from *C. sativa* treatment.[26]

The connection between the consumption of *C. sativa* and psychosis has also been studied.[1],[6],[8] In a comprehensive review, Gurley et al indicate that no direct evidence exists concerning psychosis presumably caused by the use of *C. sativa*.[8] However, the study notes that the use of *C. sativa* can worsen the pre-existing symptoms of patients who already have psychotic illness. In another review, Bostwick showed a correlation between use of *C. sativa* and psychotic illness; however, he states “the question of whether cannabis causes psychosis remains unresolved.”[1] The review assessed many studies involving *C. sativa* treatment for psychotic illnesses such as schizophrenia. Similar to the study by Gurley et al, Bostwick demonstrated that use of *C. sativa* exacerbates the symptoms of illness within individuals with psychotic disorders.[1],[8]

Cannabinoids of *Cannabis sativa*

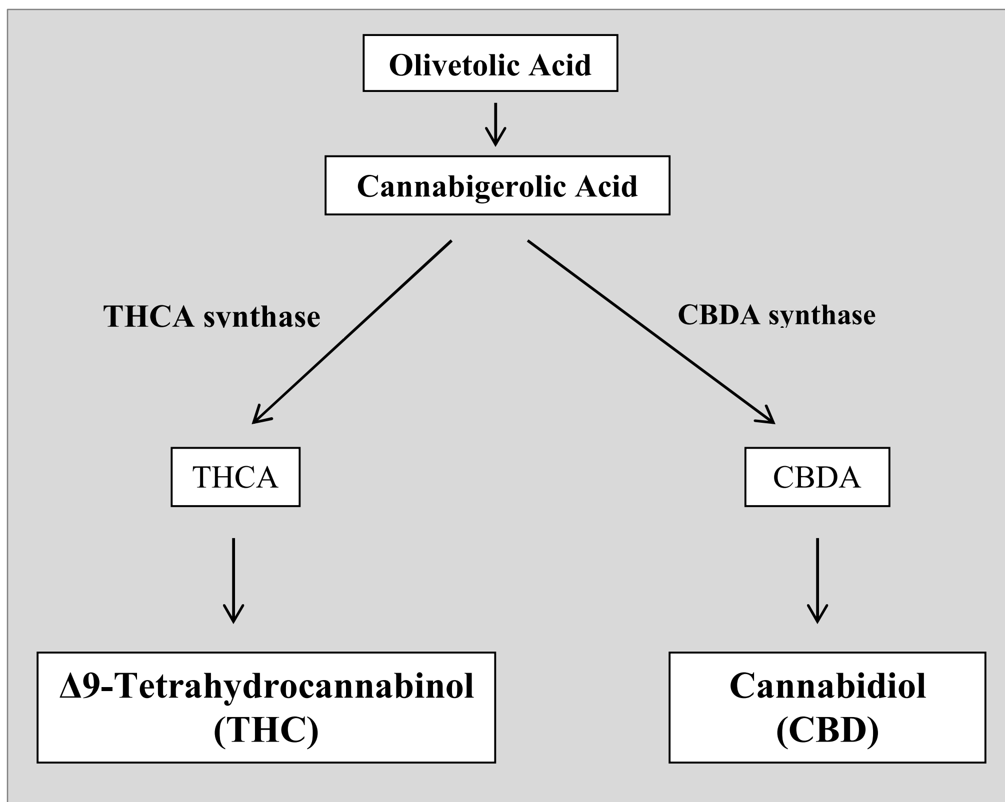
The psychotropic effects of THC present in *C. sativa* are the primary reasons for the legal restrictions of *C. sativa* use.[3] Within various strains of *C. sativa* are compounds that influence these psychotropic effects. These compounds are collectively called cannabinoids. *C. sativa* contains over 400 chemical compounds, of which more than 60 are cannabinoids.[6],[7]

Biosynthesis of Cannabinoids

As discussed previously, recent interest in the medical/recreational use of marijuana has sparked investigations of the biosynthetic pathway(s) used by the plant to produce cannabinoids. The primary cannabinoids, cannabidiol, or CBD, and Δ9-tetrahydrocannabinol, or THC, are products from the non-enzymatic decarboxylation of the actual end products of biosynthesis:

CBDA and THCA. These acids are formed from a primary compound, olivetolic acid, that is converted to cannabigerolic acid or CBGA, which then is converted to either CBDA or THCA from the action of THCA or CBDA synthase, the enzymes that functionally produce CBDA and THCA, respectively. The figure below shows the principal steps in the biosynthetic pathway.

Figure 1: Cannabinoid Biosynthetic Pathway



The biosynthetic pathway above depicts how olivetolic acid forms cannabigerolic acid, which then breaks into THCA or CBDA through the action of THCA synthase and CBDA synthase, respectively. The two cannabinoids THCA and CBDA spontaneously decarboxylate to form the corresponding neutral forms of THC and CBD.

Biosynthesis of cannabinoids is maximal in the glandular trichomes of the flowering buds, or inflorescence of the mature female *C. sativa* plant (Figure 2).[15]

Figure 2: Inflorescence of a Mature *Cannabis sativa* Plant

Shown here is an example of a mature female *Cannabis sativa* plant. The dried inflorescences analyzed in the current study are the light green, bushy material grown above the leaves. These inflorescence, or buds, are the flowering part of the plant, and are the primary site of cannabinoid biosynthesis within *C. sativa*.

One report by the United Nations describes the concentration of cannabinoids in other parts of the plant; the report demonstrates that leaves next to the inflorescence may also contain high levels of THC, but the concentrations are lower than that of the flowering buds. Also noted are the minimal levels of THC in stems and the lack of any significant THC in seeds.(12) Along with variation of cannabinoid concentration based on parts of the plant, concentration of cannabinoids within *C. sativa* plants varies based on geographical origin.(5),(7),(12) These variations within cannabinoid concentration may influence the marijuana user's selection of plant.(12)

Cannabidiol

Present in high quantities in some strains, cannabidiol (CBD) is one of the main cannabinoids of *C. sativa*.(7),(27),(28),(29) Some studies show that CBD concentrations can reach as high as 40% of the dry weight of mature plants in some strains. Although CBD does not affect individuals in a psychoactive manner, research demonstrates that the cannabinoid can be used to treat other conditions.(7) Rajesh et al discuss the beneficial effects of CBD.(7) The study indicates that CBD may serve as an antioxidant or anti-inflammatory agent and also showed immunomodulatory effects. Results from this study demonstrate that CBD may serve to treat diabetes mellitus along with treating pain, inflammation disorders, and multiple sclerosis.

Another study by Gomes et al discussed the antipsychotic properties of treatment using *C. sativa* containing high concentrations of CBD.(29) The study showed that CBD may have potential to treat the cataleptic symptoms of Parkinson's disease, but the contradictory results warrant further investigation. Along with Parkinson's disease and other striatal disorders, the antipsychotic effects of CBD were reviewed by Zuardi and coauthors (2006). Similar to the study by Gomes et al, the review by Zuardi et al reveals that CBD reduced catalepsy, or muscular rigidity, showing that CBD has potential for use as an antipsychotic treatment.(29),(30) Additionally, the study reviewed the possible use of CBD as a treatment for schizophrenia and found that CBD effectively treats psychotic symptoms in 2-4 weeks. Using CBD, patients

experience equivalent improvement to other antipsychotic treatments. However, unlike other treatments, CBD produced no side effects when treating psychotic illness.[30]

Δ9-Tetrahydrocannabinol

Contrary to the non-psychotropic effects produced by CBD, Δ9-tetrahydrocannabinol (THC) is attributed as the main psychoactive cannabinoid in *C. sativa*.[24],[25],[26] Research has shown that the two cannabinoids (CBD and THC) have adverse or opposing actions and effects[6],[27] Furthermore, CBD has been demonstrated to decrease, block, and prevent the effects of THC.[6] In the study by Gomes described above, the author states that CBD attenuates or reduces the effects of THC.[29] THC and CBD are also described as isomers,[31] which are compounds that contain the same atomic elements but differ in structure.[32] These two cannabinoids, therefore, share many similarities while producing quite different effects on individuals.

Although THC and CBD are the most abundant cannabinoids,[31] the proportions of these two compounds often vary within strains of *C. sativa*. In fact, the THC to CBD ratio is described as the definitive characteristic of *Cannabis* variation.[27],[33],[34] Often defined by percentage of THC content, the strains of *C. sativa* used for current medical purposes typically have higher percentages of THC than other cannabinoids.[33] Tambaro and Bortolato also note that strains of cannabis distributed in the illegal arena typically present higher concentrations of THC and such strains typically have a much reduced content of CBD.[34] Thus, there seems to be an inverse relationship between the levels of THC and CBD in a given strain, perhaps reflecting the competition between THCA synthase and CBDA synthase for the CBG precursor.

Potency of *C. sativa* is generally expressed in the quantity of THC present in the plant.[8] Since the psychotropic effects produced by THC are one reason the plant is federally controlled,[3] much research has been performed on the pharmacological effects of the cannabinoid.

Specifically, studies have investigated the agonistic actions of THC on the cannabinoid receptors CB1 and CB2.[7],[26] Additionally, one study by Holland et al investigated the toxicological characteristics of THC. The study stated that cases of overdose from THC have rarely been reported. Furthermore, in reference to the effects of THC within the body, the study demonstrated that "the detection window of cannabis impairment is poorly defined."[35] The primary research on cannabinoid action focuses on THC; the pharmacological actions of other cannabinoids are poorly understood and warrant further investigation.

Analysis of *Cannabis sativa*

Chemotype

The content or chemical composition of cannabinoids in *C. sativa*, known as the chemotype, can vary significantly in different strains of *C. sativa* plants. Factors such as geographic origin of the plant, as discussed previously, and cultivation methods can alter the cannabinoid content. The study of cannabinoid biosynthesis can involve both the chemical and genetic analysis of plants. Research has demonstrated that both techniques have been useful for understanding the nature of cannabinoid biosynthesis within *C. sativa*.

Chemotyping typically focuses on the separation of the cannabinoids using chromatography. These techniques allow the compounds to be separated based on the chemical structure. The most notably used chromatographic techniques for cannabinoids are thin-layer chromatography,[11],[12],[36],[37],[38] gas chromatography,[11],[12],[36],[37],[38] and high performance liquid chromatography.[11],[12],[36],[37] Gas chromatography and liquid chromatography are often paired with mass spectrometry in order to further resolve and quantify the compounds.[34]

Thin-layer chromatography uses a stationary phase along with solvents to separate the cannabinoids of *C. sativa* extracts.[12] The types of solid and liquid phases used vary based on the

polarity of compounds of interest.[36] Some studies have used silica plates as the stationary phase when chemically analyzing extracts of *C. sativa*.[12],[36],[37],[38] An extended report by the United Nations Office on Drugs and Crime (UNODC) explains suitable techniques that have been validated for analysis of *C. sativa*. The report specifies the parameters for preparing thin-layer chromatography; the UNODC states that cannabinoids such as THC and CDB are easily soluble in many organic solvents. Thus, a variety of methods are available for separation of cannabinoids via thin layer chromatography.[12] The principal limitation of thin layer chromatography, however, is that the technique lacks sensitivity and may require multiple systems to separate complex mixtures. Another issue of thin-layer chromatography results from the ambiguous separation and subsequent identification of cannabinoids from lack of adequate performance dependent on choice of method.[36]

In comparison to thin-layer chromatography, gas chromatography is more frequently used for the separation of cannabinoids in *C. sativa*. Using this technique, *C. sativa* extract is heated until vaporized (which converts THCA and CBDA into their neutral THC and CBD counterparts), and carried by a gas through a thin column containing the stationary phase in order to separate the compounds.[11] THC exists in plant material as a mixture of THC and THCA (tetrahydrocannabinolic acid). THCA spontaneously decarboxylates into THC in the presence of heat;[33] thus, the heat associated with gas chromatography automatically converts any THCA into THC and allows for the complete determination of THC content.[12] A study by Hazekamp et al, that investigated various chromatographic techniques for the chemical analysis of *C. sativa* showed that many cannabinoids are easily fragmented when analyzed with gas chromatography coupled with mass spectrometry (GC/MS).[36] Therefore GC-MS is a valuable technique for the analysis of cannabinoid content in *C. sativa*.

Genotype

Although chromatography allows for the separation and identification of cannabinoids through chemical analysis, molecular analysis of the cannabinoid synthase genes can also predict cannabinoid content according to literature.(17) Some of the preliminary genetic studies on *C. sativa* utilized the molecular analysis for identification or tracing purposes. For example, Linacre and Thorpe used the universal chloroplast transfer RNA gene sequences for the identification of *C. sativa* plants from different geographic origins.(39)

Knowledge of the biosynthetic pathway for cannabinoids, THC in particular, has caused considerable interest in the molecular biology of the THCA synthase gene.(16),(33),(40),(41),(42) In a study by Marks et al (2009), RNA extracted from the glandular trichomes of *C. sativa* was used to construct a cDNA library that was then used to quantify the THCA synthase content in different tissues of the *C. sativa* plant.(33) Results of the study showed that maximal expression of THCA synthase occurs in the trichomes. Another study by Sirikantaramas et al cloned the THCA synthase gene to predict the characteristics of the THCA synthase protein and found that the gene encodes a polypeptide consisting of over 500 amino acid residues.(42) Kojoma et al extended the results from Sirikantaramas et al and discovered polymorphisms within the THCA synthase gene, some of which would result in the production of non-functional THCA synthase enzyme due to the introduction of either a premature stop codon or alterations in the amino acid sequence.(16),(42) Over 60 nucleotide substitutions were identified in the THCA synthase gene by Kojoma et al that resulted in 37 amino acid substitutions within the synthase polypeptide chain.(16)

Rotherham and Harbison also used the reported sequence of the THCA synthase gene to develop a molecular assay designed to reliably analyze four specific single-nucleotide polymorphisms (SNPs) of the THCA synthase gene of *C. sativa* that changed the amino acid

sequence of the protein in a way that was predicted to inactivate the enzyme. The results from this study showed that these four polymorphisms differentiate between the active and inactive form of THCA synthase. The primers designed by Rotherham and Harbison for the molecular analysis of the four specific SNPs allowed for minisequencing of the specific SNP sites using a process generally known as SNaPshot® analysis. The conclusion from the study of Rotherham and Harbison was that their assay of nucleotides present at the four SNP positions could reliably distinguish between active (drug-type) and inactive (fiber-type) THCA synthase genes in a plant and therefore predict the THC content of a mature *C. sativa* plant.[17]

Summary

Research of *C. sativa* has provided the medical community with a foundation for the controversial debates about treatment using medical marijuana.[24] Though medical use of *C. sativa* has been prevalent throughout history,[8,16,18] the plant remains listed as a federally-controlled substance within the United States.[3] Legislation within certain states, however, has enabled individuals to treat symptoms of illnesses such as AIDS or glaucoma with *C. sativa*.[4] The psychotropic effects produced by THC are among the factors that limit the use of *C. sativa* under federal regulation.[3] Research into the biosynthetic pathway of THC and other cannabinoids such as CBD continues to provide insight into the specific effects of *C. sativa*.[6,15,33,31,40,41] Many of the details about the chemical composition and biosynthesis remain unclear.[14] Therefore, these gaps in research warrant further investigation of the chemotype and genotype of *C. sativa*.[17] The methods proposed in this thesis will analyze some of the issues surrounding the chemotype and genotype of *C. sativa.*

CHAPTER III

METHODOLOGY

Overview

Analytical techniques for assessment of cannabinoid content of *Cannabis sativa* include both genomic analysis and gas chromatography coupled with mass spectrometry (GC-MS). Genomic analysis includes DNA extraction, agarose gel electrophoresis, single nucleotide sequencing (i.e., SNaPshot® analysis),[17] capillary electrophoresis, and genomic sequencing. Genomic analysis using the method of Rotherham and Harbison allows for the detection of active and inactive THCA synthase genes of *C. sativa* according to the conclusions of their study. [17] Recall that this gene is directly responsible for the synthesis of the enzyme that converts CBG to THCA in mature *C. sativa* plants. Gas chromatography coupled with mass spectrometry was chosen to detect and quantify cannabinoid content.[12]

Selection of *Cannabis sativa* Samples

In order to analyze *C. sativa* samples in the School of Forensic Science laboratory at Oklahoma State University-Center for Health Sciences, permission was sought from the OSU-CHS administration to work with small amounts of marijuana within the forensic student

laboratories. After this permission was granted, a request was sent to the Tulsa Police Department (TPD) to provide small amounts of seized *C. sativa* evidence from former casework that was marked "to be destroyed." The TPD Evidence Custodian and the TPD Crime Laboratory Director obtained permission to support our project from the command staff within TPD.

After permission was granted to work with former evidence from the TPD, samples were selected from the property room. Because storage of *C. sativa* may alter the cannabinoid content, samples were selected from case numbers created within the last two years. Separate case numbers were also selected in an attempt to avoid repeat analysis of the same plant material in the different sample selection events (three in total over the course of the present study). In addition, care was taken to select seized samples composed of individual inflorescences, or buds, to help ensure the likelihood that plant material was unique to an individual plant and also to maximize the THC chemical content of the material subjected to GC/MS.

Because *C. sativa* is a Schedule I federally-controlled substance,[3] precautions were taken for access and storage of samples. During analysis, samples were accounted for at all times. When samples were not needed in the laboratory for analysis, they were stored in a safe in an evidence room with controlled access. Only the Chair of the School of Forensic Science and the Director of Quality have access to the safe. A chain of custody log was also kept for the samples to show proof when the samples were handled, by whom, and the quantity of sample that was removed.

Genetic Analysis

DNA Extraction

Extraction of DNA was performed using an organic extraction protocol employing digestion of plant material in TNE buffer (10 mM Tris-Cl, pH 8.0, 0.15M NaCl and 0.1 mM EDTA; Thermo-Fischer, Waltham, MA) containing proteinase K (40 µg/mL; Promega Corp,

Madison, WI) and sodium dodecyl sulfate (0.5% final concentration; Affymetrix, Santa Clara, CA) with incubation for 2 hours at 65°C. Digested samples were then subjected to organic solvent extraction with a mixture (9:0.96:0.04 v/v/v) of phenol (MP Biomedicals, Santa Ana, CA) and chloroform:isoamyl alcohol (Sigma Aldrich, St. Louis, MO). Before initial extraction, an extraction buffer was prepared that contained 25 µL SDS, 20 µL Proteinase K, and 955 µL TNE buffer. Samples of 10-15 mg of plant material were weighed out and placed in labeled 0.6 mL microfuge tubes. A spatula of glass beads was added to the tubes as well, in order to assist in breakage of plant material during mixing on the vortex mixer. Then 250 µL of extraction buffer was added to each tube and incubated at 65 °C for 2 hours.

After incubation, for each sample, a hole was pierced into the bottom of the 0.6 mL microfuge tube before it was placed into a labeled 1.8 mL microfuge tube. The liquid was centrifuged away from the residual plant material during a 1 minute centrifugation at 6000 x g. A mixture of phenol:chloroform:isoamyl alcohol (9:0.96:0.04 v/v/v) solution was prepared and an equal volume of organic solvent was added to each extract in the 1.8 mL microfuge tube. The samples were vortexed and placed in a centrifuge at 10,000 x g for 3 minutes. The upper aqueous phase from each sample was recovered with a micropipette and placed into a labeled clean 1.8 mL microfuge tube. An equal amount (~250 µL) of chloroform:isoamyl alcohol (24:1) was added to each tube and the samples were vortexed again and centrifuged at 10,000 x g for 3 minutes. Again, the upper aqueous phase for each sample was collected using a micropipette and placed into a clean, labeled 1.8 mL microcentrifuge tube. DNA was recovered from extracts using Zymo Clean and Concentrator-25 technology (Zymo Research, Irvine, CA) per manufacturer instructions, with the exception of a modification to the elution step. After washing, samples were eluted two times in succession using 15 µL aliquots of TE^{-4} (10mM Tris, 0.1mM EDTA) at 65°C. The extracts were then ready for processing using polymerase chain reaction.

Primer Selection and Optimization

Two different rounds of PCR required a variety of primer pairs for the SNaPshot® reaction. The first pair, called C2E2, directs the amplification of a 400 base-pair segment of THCA synthase gene that harbors the four single nucleotide polymorphisms (SNPs) reported by Rotherham and Harbison.[17] These SNPs were determined by Rotherham and Harbison to correlate with an active/inactive status of the THCA synthase gene. The C2E2 primers were reconstituted in TE^{-4} to 100 µM stock solution, and aliquots of 10X primer pairs were made by performing a 1:10 dilution with TE^{-4}. The numbered pairs of primers were selected based on the SNaPshot® parameters used by Rotherham and Harbison, with the exception that three thymine nucleotides were added to the 5' end of the 8F primer as shown below to obtain better separation of SNaPshot® products during capillary electrophoresis. The concentrations of the SNaPshot® primers were as described in the Rotherham and Harbison study.[17] The THCa and THCb primers were used to sequence the entire THCA synthase gene via Ion Torrent PGM.

Table 1: *Cannabis sativa* Primer Pairs

Primer name	Primer Sequence (5' – 3')	Length (bp)	Tm (1M Na^+)
C2 F	CAAACTKGTTGYTGTCCCATC	21	81°C
E2 R	CGTCTTCTTCCCAGCTGATC	21	82°C
8F(*)	<u>TTT</u>GAGTTGGGTATTAAAAAAACTGATTGCAAAGAATT	38	92°C
9F	CAACCATCTTCTACAGTGGTGTTGTAAATT	30	86°C
16R	TCRACTAGACTATCCACTCCACCA	24	82°C
17R	TACTGTAGTCTTATTCTTCCCATGATTATCTGTAATATTC	40	87°C
THCa	TGAAGAAAAAAAATGAATTGCTCAGCATTTTTC	33	69°C
THCb	TCTATTTAAAGATAATTAATGATGATGCGGTGG	33	66°C

(*) The underlined nucleotides were added to the sequence published by Rotherham and Harbison.[17]

Polymerase Chain Reaction

Polymerase Chain Reaction (PCR) was performed using GoTaq Hot Start® DNA polymerase (Promega Corp, Madison, WI) per manufacturer instructions. The thermal cycling parameters used were those described by Rotherham and Harbison.[17]

To perform minisequencing, a 400 base-pair (bp) sequence of the THCA synthase gene was first amplified by PCR. The C2E2 primer (Table 1) directed the amplification of the 400 bp amplicon using approximately 1 ng of genomic DNA template. Following amplification, the 400 bp amplicon was visualized via agarose gel electrophoresis both to confirm that product had been amplified and also to get a rough estimate of product concentration. A 1.5% gel was made by adding 1.125 g agarose (Invitrogen, Waltham, MA) to 75 mL 1X TAE (Tris-Acetate-EDTA, pH 8.0) buffer (Invitrogen; Waltham, MA) and heating the mixture to solution in a microwave for approximately 90 seconds. Following partial cooling of the gel solution, 15 µL of 10 mg/mL ethidium bromide (MP Biomedicals, Solon, OH) was added to the molten agarose which was then poured into a gel mold and cooled to room temperature. Then, 5 µL aliquots of PCR products were added to 10 µL of 1X TAE containing tracking dyes (bromophenol blue, Sigma-Aldrich, St. Louis, MO; xylene cyanol, BioRad, Berkeley, CA; and Ficoll, Sigma-Aldrich, St. Louis, MO). Samples were loaded onto the gel along with two 100 bp size ladders (Invitrogen, Waltham, MA). The gels were electrophoresed at 70 volts for approximately two hours and then placed in a UV illuminator to confirm the presence of the 400 bp template for minisequencing.

Single Nucleotide Polymorphism Analysis

In order to analyze the four single nucleotide polymorphisms, the SNaPshot® minisequencing kit was used (Thermo Fisher, Waltham, MA). SNaPshot® technology enables the nucleotide base at a predetermined position within the DNA molecule to be identified in a minisequencing approach. For this research, primers for the SNaPshot® reaction revealed the

nucleotide at each of the four positions that Rotherham and Harbison suggested allowing active and inactive THCA synthase genes to be distinguished.[17]

SNaPshot® reactions were performed following the manufacturer's protocol. Briefly, the 400 bp minisequencing template was subjected to removal of PCR primers and deoxynucleotide triphosphates (dNTPs) following a 1 hour incubation with a mixture of exonuclease and alkaline phosphatase (EXO/SAP, Affymetrix, Cleveland, OH). The dilution of SNaPshot® products was also modified from a 1:10 dilution described by Rotherham and Harbison to between 1:50 and 1:100 dilution in sterile deionized water which improved the overall quality of SNaPshot® results. All other SNaPshot® reaction steps were performed according to methods described by Rotherham and Harbison.[17] Similar to PCR, SNaPshot® minisequencing uses DNA polymerase and primers to amplify template DNA. However, SNaPshot® extends primers by a single nucleotide through the use of dideoxynucleotide triphosphates (ddNTPs). Minisequencing primers were those described by Rotherham and Harbison with the exception of the addition of three thymine nucleotides to the 5' end of the 8F primer, as discussed previously.[17] This allowed for better separation between the minisequencing products containing the 8F and 9F nucleotide polymorphisms. Visualization of SNaPshot® products was performed on ABI 3130 capillary electrophoresis instrument using Gene Scan™ 120LIZ™ (Applied Biosystems, Waltham, MA) size standard and GeneMapper® (Applied Biosystems, Waltham, MA) analysis software (version 3.2), which allowed for visualization of electropherograms of capillary electrophoresis results.

Optimization of SNaPshot®

Methodology of SNaPshot® reactions were initially followed according to Rotherham and Harbison.[17] The SNaPshot® settings were established in GeneMapper® (Applied Biosystems, Waltham, MA) analysis software (version 3.2). After these settings were adjusted for optimal detection of the specific SNPs, the initial analysis of samples showed an abundant amount of off-

scale results. This effect happens when too much PCR product is loaded into the capillary 96-well plate. In order to counteract the occurrence of off-scale results, the dilution step within the SNaPshot® reaction described by Rotherham and Harbison was increased as discussed previously.[17]

Further optimization included determination of minimum relative fluorescence unit (RFU) in the allele peak heights for the SNPs. Two samples were selected that were previously determined to be homozygous active and homozygous inactive, respectively. These samples were processed with SNaPshot® five times to determine the minimum RFU setting appropriate for detection of SNPs. By visualizing the areas of fluorescence baseline between peaks in the electropherogram, the level of baseline "noise" could be determined by setting the detection of RFU for each fluorescent dye at 1 RFU. Any values below 1 RFU were determined to be 0. The baseline noise showed values above 0 RFU where peaks for SNP positions were absent (Table 2).

Table 2: SNaPshot® Baseline Threshold Optimization

Sample	8A	8I	9A	9I	16A	16I	17A	17I
6043-1	62		84		49		63	
6043-2	39		52		39		40	
6043-3	46		56		42		47	
6043-4	71		89		62		73	
6043-5	35		44		29		73	
6922-1		14		31		36		43
6922-2		0		13		28		20
6922-3		0		18		20		24
6922-4		0		17		21		0
6922-5		0		17		16		15

The inactive sample 6043 and the active sample 6922 were selected for their THCA synthase activity determined from SNaPshot® results. These samples were processed with SNaPshot® five times to determine the baseline threshold for the predetermined single nucleotide polymorphism positions within the THCA synthase gene. Shown are the relative fluorescent units (RFU) for the absent active peaks of 6043 and the absent inactive peaks of 6922.

These values were averaged for each absent peak across all SNPs and the standard deviation from these values was calculated. Three times the standard deviation was added to the average to determine the minimum RFU detection threshold. The values of each calculation are as follows.

Table 3: SNaPshot® Baseline Threshold Calculations

	Average RFU	Standard Deviation	3 X Standard Deviation	Baseline
Active SNP	55	17	51	55 + 51=106
Inactive SNP	16	12	36	16+36=53

Since 53 was the calculated minimum baseline RFU, the threshold RFU was set at 50. Results for the SNaPshot® reactions showed that samples either depicted homozygous-active, heterozygous, or homozygous-inactive SNPs for each position. For example, if a homozygous-active sample lacked an inactive peak at the 8F SNP position, all other SNP positions also lacked the inactive peak. The RFU of each peak was typed into an Excel spreadsheet for statistical analysis. Within the spreadsheet, ratios of active/inactive peak height RFUs were recorded for each SNP position. These ratios were averaged within each SNP position for all samples, as well as within all SNPs for each respective sample (Appendix A).

THCA Synthase Amplicon Sequencing

Once the SNaPshot® reactions were performed, a select number of samples were chosen for partial gene sequencing analysis based on THCA synthase activity level demonstrated via SNaPshot® results. Using the primers which direct the 400 bp amplicon discussed above, the seven selected samples were sequenced by staff in the Oklahoma State University School of

Forensic Sciences to further analyze the C2E2 amplicon region within the THCA synthase gene for further sequence information. The parameters for amplicon sequencing using the Ion Torrent Personal Genome Machine® (PGM™) (Life Technologies, Carlsbad, CA) were those described in the manufacturer's instructions.[43] Data was analyzed using the Integrated Genomics Viewer (Broad Institute, Cambridge, MA) program.[44]

Chemical Analysis via GC/MS

Preparation of Chromatography

Chemical analysis of cannabinoids was performed using gas chromatography coupled with mass spectrometry (GC/MS) on an Agilent 6890 GC (Agilent, Santa Clara, CA) coupled with an Agilent 5973 mass spectrometer (Agilent, Santa Clara, CA). For optimum separation of cannabinoids using gas chromatography-mass spectrometry, the Rxi®-32Sil MS fused silica column was selected (Restek, Belfonte, PA). Suggested parameters for gas chromatography that were provided by Restek were further optimized in house and included injection volume and split ratio, injection liner, injection temperature, oven temperature, carrier gas, flow rate, and detector parameters. Methapyrilene (MePy; Sigma Chemical Co., St. Louis, MO) was chosen as the internal standard and was mixed with samples at 50 µg/mL concentration in methanol (VWR International, Radnor, PA).

Identification of ions for each of the three cannabinoids [THC, CBD, and Cannabinol (CBN)] was accomplished by running known drug standards (Cerilliant, Round Rock, TX) and choosing ions from their respective mass spectra. Cannabinoid concentrations were determined using a standard curve constructed by analyzing signal strengths produced by known amounts of THC, CBD, and CBN on the GC/MS. Standard curves were produced using an eight point curve ranging in concentration from 200 µg/mL to 1.063 µg/mL of each cannabinoid. For each standard at each concentration, a ratio of the signal produced by the drug versus the signal

produced by the internal standard was determined; this normalized each injection of sample into the GC/MS based upon a constant methapyrilene standard. Standard curves routinely exhibited R^2 values of >0.99.

Cannabinoid concentrations in unknown samples were estimated by plotting the ratio of THC, CBD, or CBN to the internal standard on the standard curve. Included with each GC/MS run were three controls that were produced by diluting reference cannabinoid standards to concentrations of 200µg/mL, 50 µg/mL, and 6.25 µg/mL representing high, medium, and low concentration controls. The lot numbers of the cannabinoids used for the controls were different from those of the standard curves, thus ensuring that the instrument was performing correctly between sampling. A limit of detection was calculated by averaging the total background noise for each negative control sample and multiplying the value by three. The limit of quantitation was also calculated by multiplying the same background noise average by 10. These limits set the minimum value for detection and quantitation of drugs in the *C. sativa* samples.

Cannabinoid Extraction from Plant Material

Using an analytical scale and sterile forceps, 15 mg of plant material from seized drug samples were weighed and placed into labeled 0.6 µL microfuge tubes. These weights were recorded on the evidence log described previously. To extract the cannabinoids from the plant material, 250 µL of 50 µg/mL MePy/MeOH was added to each sample. The tubes were placed in a heat block at 65 °C for 30 minutes. After incubation, a hole was pierced in the bottom of each 0.6 mL tube that was then placed into a labeled 1.8 mL microfuge tube. The samples were centrifuged at 10,000 x g for 3 minutes to collect the liquid extract while the plant material remained in the 0.6 mL microfuge tube. Following centrifugation, 4 µL of each extract was placed into a labeled GC-MS vial with 196 µL of 50 µg/mL MePy/MeOH. Using this approach it was possible to ultimately calculate the percent cannabinoid based upon the weight of the plant material extracted.

Chemical Analysis

To ensure that the instrument is prepared to run samples, the mass spectrometer was autotuned prior to running samples. Methanol was added to the appropriate vial for cleaning between each sample. A blank of methanol was also run prior to samples to ensure the column was free and clear of any contaminants. A negative control consisting of 50 µg/ml MePy/MeOH was run as well to ensure that methapyrilene was selected as internal standard and to measure the amount of background noise for each cannabinoid. Three unique ions were selected for each cannabinoid based on abundance in the mass spectrogram.

Batches were input based on sample name, vial number, and date of run. Because the instrument is set up with an auto injector, each set of injections allowed for eight vials. One standard curve was run prior to any samples for quantification of cannabinoids. A control was run as vial 1 for each set, and samples were run in vials 2-8.

Statistical Analysis

Correlation Studies

Correlation statistics were computed by Dr. Mark Payton of the Oklahoma State University Statistics Department to assess the possible correlation between the THCA SNP activity and cannabinoid composition of *C. sativa* samples. First, an analysis of variance, or ANOVA was performed to compare the percentage of THC from GC-MS data with characteristics of three SNPs from SNaPshot® data. In order to perform the ANOVA, numerical values were used in place of homozygous-active, heterozygous, and homozygous-inactive nominal descriptions. Thus, homozygous-active samples were numbered 0, heterozygous samples were numbered 1, and homozygous-inactive samples were numbered 2. Then, the ratio of active/inactive peak height RFU was computed from SNaPshot® results (discussed previously). To determine if a correlation was present between these values and percent THC, the potential

correlation between each possible combination of RFU value ratio at a given SNP to percent THC obtained from a sample was calculated using Pearson's coefficient. A correlation would determine that percent THC could be predetermined based on genetic analysis via SNaPshot® within *C. sativa* samples.

Summary

Using SNaPshot® analysis and gas chromatography-mass spectrometry, the presence of the polymorphisms and THC content within *C. sativa* samples can be determined. This determination could potentially show that a correlation exists between cannabinoid content and genetic variations within *C. sativa*. This research could potentially provide a resource to growers, providers, dispensaries, and users of medicinal marijuana to know the quality of medicinal marijuana plant product based on the analysis of seeds, pollen, and immature plant material. The knowledge gained from this study could save these individuals time spent growing plants as well as stress from the lack of knowledge about their current product. It could also optimize treatment of illnesses using medicinal marijuana, giving the individuals affected by those illnesses more peace of mind.

CHAPTER IV

RESULTS AND INTERPRETATION

Overview

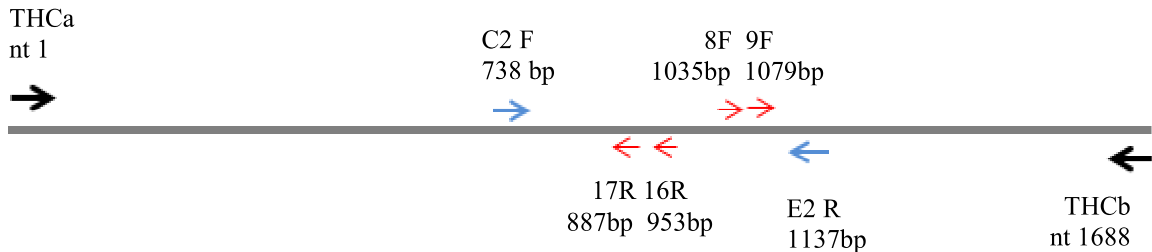
Investigative methods of genetic and chemical analysis of *Cannabis sativa* samples discussed previously include SNaPshot® analysis, amplicon sequencing, and gas chromatography-mass spectrometry. The SNaPshot® reactions were performed on separate days prior to GC-MS for each set of samples. The results and interpretation of results follow.

Genetic Analysis

Results of SNaPshot®

In order to examine the results from the SNaPshot® reactions, a brief review of the functions of the process is necessary. Primers 8F and 9F designed by Rotherham and Harbison are forward primers that are complimentary to the non-coding sequence of the THCA synthase gene.(17) Contrarily, primers 16R and 17R are reverse primers that are complimentary to the coding strand of the THCA synthase gene. A graphic displayed in Figure 3 provides a visual interpretation of this mechanism.(17)

Figure 3: Primer Alignment in C2E2 Amplicon of THCA Synthase Gene.



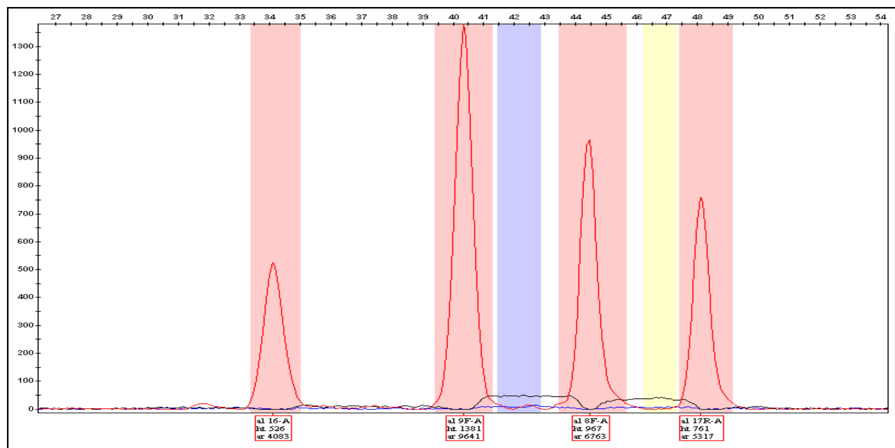
The visual representation adapted from Rotherham and Harbison shows the THCA Synthase gene spanning across 1688bp of the *C. sativa* genome.[17] The C2E2 amplicon is located 738 bp downstream from the THCA synthase start codon. The locations of the C2 and E2 primers discussed in the methodology section are shown here. The SNP positions within the gene are also demonstrated from their individual primer locations. The forward (F) SNP primers 8F and 9F are complimentary to the non-coding strand, while the reverse (R) SNP primers 16R and 17R are complimentary to the coding strand.

The SNaPshot® mastermix contains dideoxynucleotides (ddNTPs) that allow the sequence to extend one extra nucleotide after the end of the primer. Within active (wild-type) samples, a dideoxythymine is extended from the forward primers. Since the forward primers are complementary to the non-coding strand, a thymine is incorporated within the nucleotide sequence. Similarly, within active (wild-type) samples, a dideoxythymine is extended from the reverse primers. However, because the reverse primers are complementary to the coding strand, an adenine is incorporate within the nucleotide sequence. Because SNPs are changes within the DNA sequence, the inactive samples that contain SNPs at the predetermined positions incorporate nucleotides other than the active (Table 4).

Table 4: THCA Synthase SNP Nucleotide Incorporation				
	8F	9F	16R	17R
Active	Thymine	Thymine	Adenine	Adenine
Inactive	Guanine	Adenine	Thymine	Guanine

The results of the SNaPshot® reactions were visualized in GeneMapper (version 3.2) genetic analysis software. Examples of homozygous-active, heterozygous, and homozygous-inactive chromatograms are shown in Figures 4, 5, and 6 respectively.

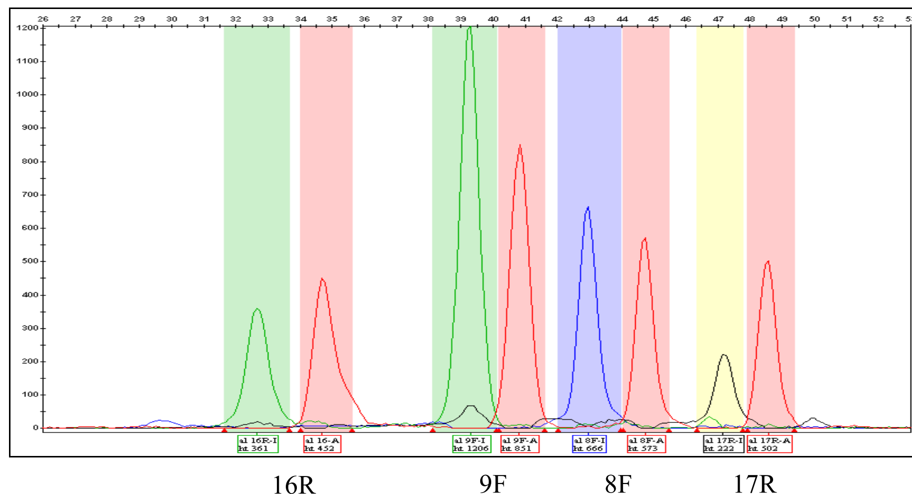
Figure 4: Chromatogram of a Homozygous-Active Sample



16R 9F 8F 17R

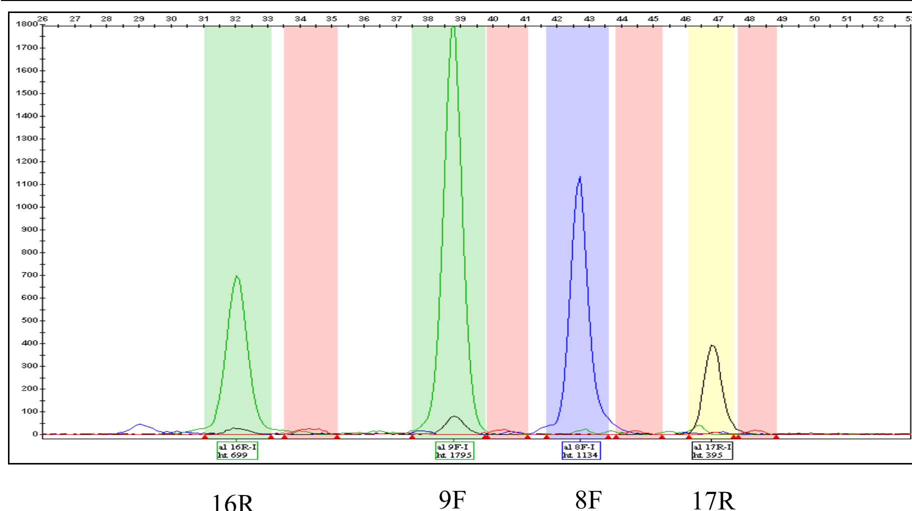
The red peaks depict the extension of a thymine from the 3' end of the primers discussed previously within the SNaPshot® reaction, demonstrating that the sample types active for these four SNP positions within THCA synthase.

Figure 5: Chromatogram of a Heterozygous Sample



The multicolored peaks (other than red in Figure 4) depict the extension of an adenine for green, a guanine for blue, and a cytosine for yellow (visualized as black) in place of other nucleotides within the template strand in the SNaPshot® reaction. The results demonstrate that the sample types heterozygous for these four SNPs within THCA synthase.

Figure 6: Chromatogram of a Homozygous-Inactive Sample



The multicolored peaks depict the extension of an adenine for green, a guanine for blue, and a cytosine for yellow (visualized as black) in place of other nucleotides from the template in the SNaPshot® reaction. Because the chromatogram is lacking the active thymine (red) peaks, the results demonstrate that the sample types homozygous-inactive for these four SNPs within THCA synthase.

Once the chromatograms were analyzed considering the optimized analytical threshold of 50 RFU, the relative fluorescent units (RFU) for each SNP position of the samples were recorded. These units were then assessed by forming a ratio of active to inactive peak height RFU. The samples with a ratio of 0 were determined to be homozygous-inactive (I), the samples with a ratio between 0 and 1 were heterozygous (A/I), and the samples with a ratio of 1 were determined to be homozygous-active (A). With 75 *C. sativa* samples analyzed, only three samples typed as homozygous-active. Furthermore, out of 75 samples analyzed, only four samples typed as homozygous-inactive. Therefore, 68 samples were determined to be heterozygous for active and inactive THCA synthase single nucleotide polymorphisms. As mentioned in the methodology, samples that typed homozygous-active depicted only active peaks at all four SNP positions; samples that typed heterozygous depicted both active and inactive peaks at all four SNP positions, and samples that typed homozygous-inactive depicted only inactive peaks at all four SNP positions. Therefore, there were no mosaic representations within the results in which a sample would contain only an active peak at one SNP position and only an inactive peak at another SNP position. The lack of mosaic results reflects that the four SNPs within the THCA synthase gene act as a uniform haplotype.

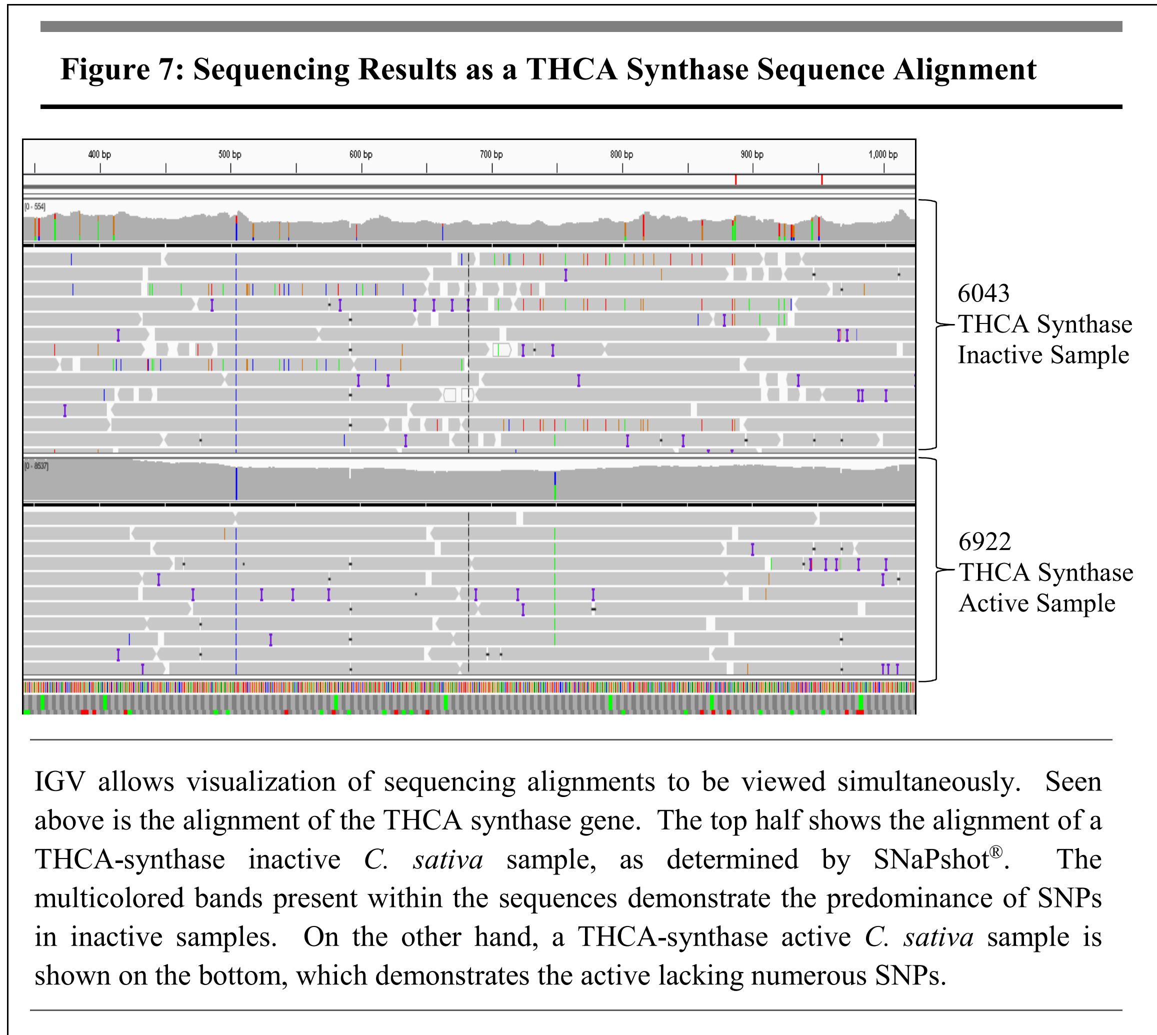
Results of THCA Synthase Amplicon Sequencing

In order to analyze the SNPs of the 400 bp amplicon, the THCA synthase gene was selected as the reference gene from GenBank (AB057805.1). The samples processed with genomic sequencing using the Ion Torrent PGM platform were selected based on zygosity. Seven samples were sequenced and processed through IGV software, including active samples #5725 and #6922; heterozygous samples #0565, #0777, #1688, and #5725; and homozygous samples #6043, and #7436 (Appendix A).

When visualizing the entire THCA synthase gene, a remarkable predominance of single nucleotide polymorphisms was seen in THCA synthase-inactive and heterozygous samples in

comparison with THCA synthase-active samples. Figure 6 below displays the sequencing alignments of two samples that typed as inactive and active for the THCA synthase gene in SNaPshot®.

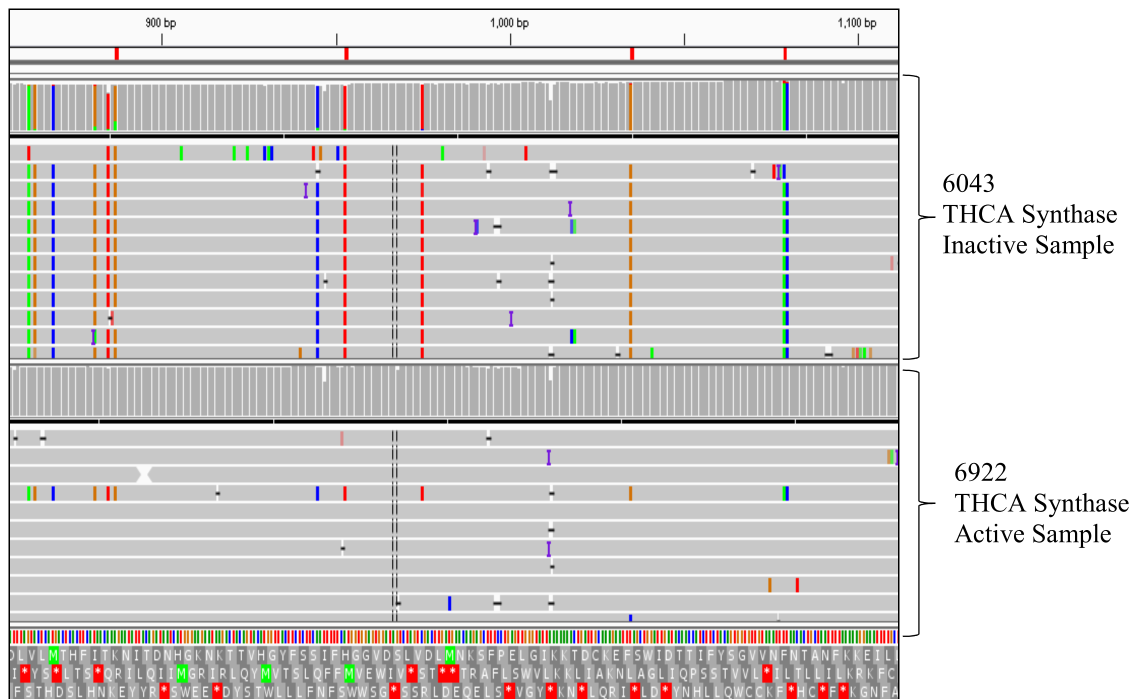
Figure 7: Sequencing Results as a THCA Synthase Sequence Alignment



IGV allows visualization of sequencing alignments to be viewed simultaneously. Seen above is the alignment of the THCA synthase gene. The top half shows the alignment of a THCA-synthase inactive *C. sativa* sample, as determined by SNaPshot®. The multicolored bands present within the sequences demonstrate the predominance of SNPs in inactive samples. On the other hand, a THCA-synthase active *C. sativa* sample is shown on the bottom, which demonstrates the active lacking numerous SNPs.

All of the SNPs that were reported from Rotherham and Harbison were present in each homozygous-inactive samples processed on Ion Torrent PGM.[17] Figure 7 shows examples of active and inactive samples within the C2E2 amplicon region of the THCA synthase gene.

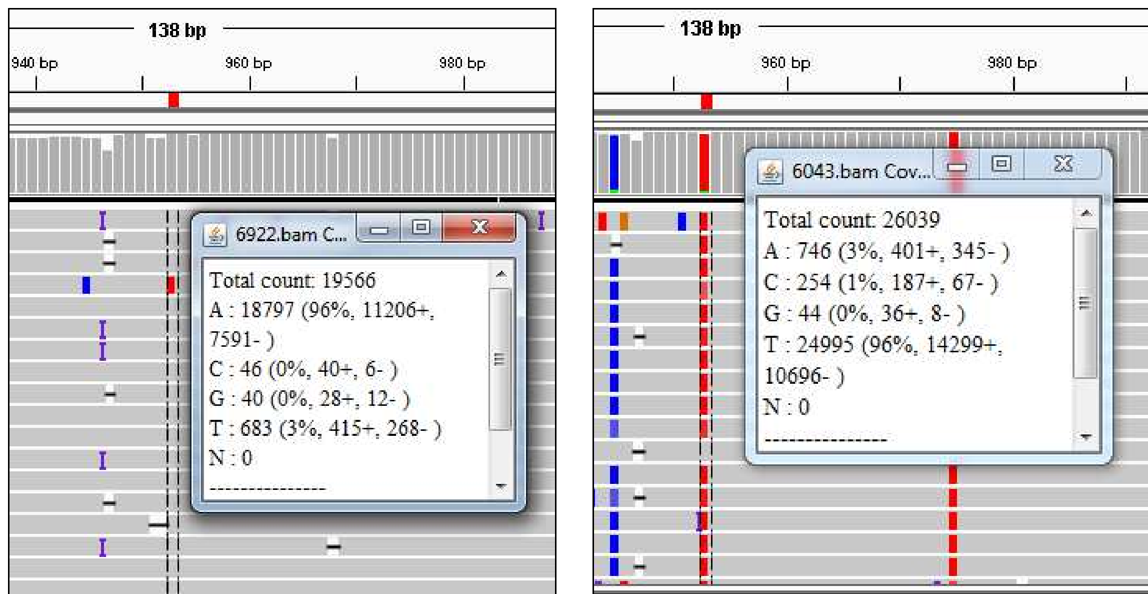
Figure 8: Sequence Results in the C2E2 Amplicon Region



As seen with the THCA synthase genomic sequencing results, the inactive samples retain the single nucleotide polymorphisms within the genome. All four SNPs analyzed were present in each inactive sample, along with a multitude of other SNPs. The active samples, however, demonstrate wild-type sequence alignments, shown by the lack of SNPs at the four positions within the sequence. The four vertical red bars at the top of the figure indicate the SNP positions within the C2E2 amplicon as reported by Rotherham and Harbison.[17]

The results from the amplicon sequencing allowed for the quantitation of SNPs within a sample. Based on the number of replicate alignments, the percentage of each nucleotide present in the replicate strands sequenced on the Ion Torrent PGM can be determined for each position in the sequence. Figure 8 displays the alignment of an active sample and an inactive sample with the percentages shown for the 9F SNP position.

Figure 9: Percentage of Nucleotide Incorporation at the 16R SNP Position.



6922-THCA Synthase Active | 6043-THCA Synthase Inactive

The percentage of nucleotides present in the replicates for the 16R position is shown for the active (6922) and inactive (6043) samples above. The presence of adenine for the active is displayed at 96%, whereas the inactive sample counts adenine at 3%. However, the inactive sample presents thymine at 96%, showing the difference in quantities of SNPs and zygosity.

Use of IGV also allows for the translation of the reading frame from the DNA sequence to the amino acid sequence. As such, the active and inactive amino acid sequences are displayed (Table 5). Changes in the amino acid sequence shown below are caused by the SNPs.

Table 5: Changes in Amino Acid Sequence at SNP Positions

	17R	16R	8F	9F
Wild-Type (active)	Lysine (AAG)	Histidine (CAT)	Phenylalanine (TTT)	Phenylalanine (TTT)
SNP (inactive)	Arginine (AGG)	Leucine (CTT)	Leucine (TTG)	Tyrosine (TAC)

The nucleotide sequence in Table 5 at the 9F SNP position reflects the established thymine to adenine polymorphism, as well as the thymine to cytosine polymorphism. The thymine to cytosine transition is present in all inactive samples sequenced on Ion Torrent PGM.

Gas Chromatography-Mass Spectrometry

Results of Gas Chromatography-Mass Spectrometry

In order to evaluate the results of GC/MS, the values of limit of detection (LOD), limit of quantitation, (LOQ), and the standard curves were first analyzed. In order to calculate the LOD and LOQ, the measurement of total concentration was added from each ion for each cannabinoid from the three negative controls (aliquots of 50µg/mL methapyrilene in methanol). These values were then averaged for each cannabinoid. The LOD and LOQ calculations are shown in Table 6 below. These limits were set to establish the minimum values to determine if the drugs within the samples were quantifiable.

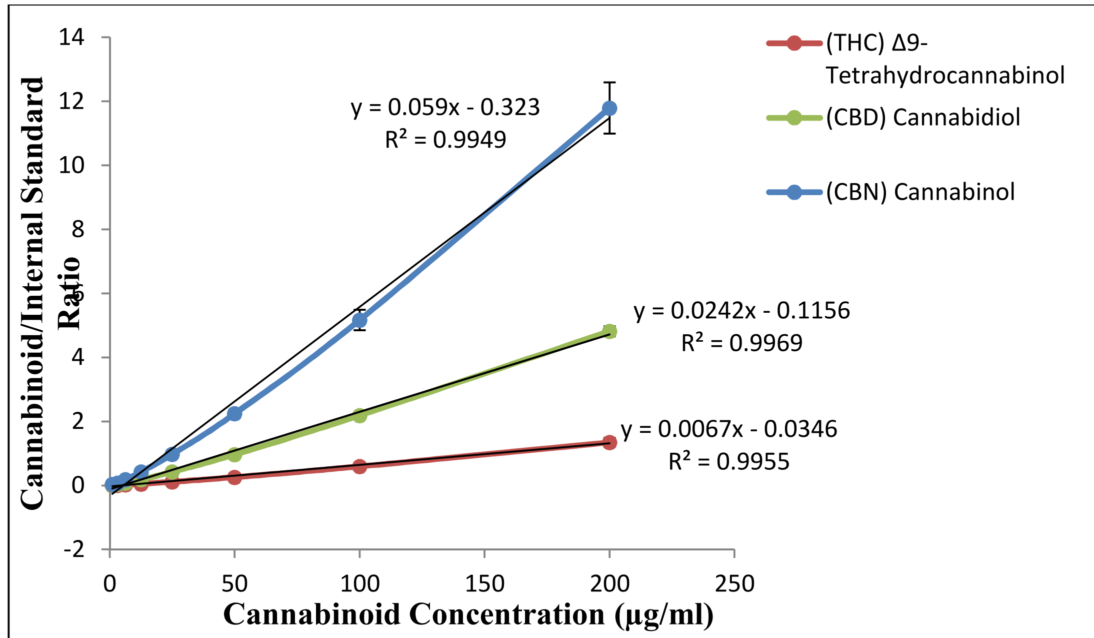
Table 6: Limit of Detection and Quantitation Calculations (in absorbance units)

	CBD	THC	CBN
Average Background Noise Detection	637.5	450	800
Limit of Detection (3 times average)	1912.5	1350	2400
Limit of Quantitation (10 times average)	6375	4500	8000

The standard curves discussed in the methodology section enabled the quantitation of cannabinoids within the samples. Three curves were made, and the values of each cannabinoid

for each concentration were averaged to compare with *C. sativa* samples. An average standard curve is displayed in Figure 10.

Figure 10: GC-MS Standard Curve for CBD, THC, and Cannabinol



Above are the plotted standard curves for each cannabinoid, expressed by color depicted in the legend. The known concentration was compared to the ratio of cannabinoid/methapyrilene concentration taken from the mass spectrometer output. This created the data in the form of a standard curve, with linearity R^2 values depicted above 0.99. The function formula is used to compute the concentration of each cannabinoid in unknown samples. The error bars in each curve represent the potential error, calculated by one standard deviation for each data point.

In order to ensure that the GC-MS was operating correctly between runs, three concentration controls were used for each cannabinoid and included with each batch of samples analyzed. Table 7 shows the concentration of cannabinoids calculated from plotting the ratio of cannabinoid/internal standard in the standard curve formula of each cannabinoid.

Table 7: GC-MS Cannabinoid Controls

	Day 1			Day 2			Day 3		
Known	CBD	THC	CBN	CBD	THC	CBN	CBD	THC	CBN
200 µg/mL	185.05	239.70	192.31	195.98	239.70	192.31	205.40	205.4	207.67
50 µg/mL	43.96	45.03	44.12	45.85	45.03	44.12	43.08	46.73	43.68
6.125 µg/mL	7.98	8.67	8.90	7.56	8.67	8.90	8.19	8.73	8.54

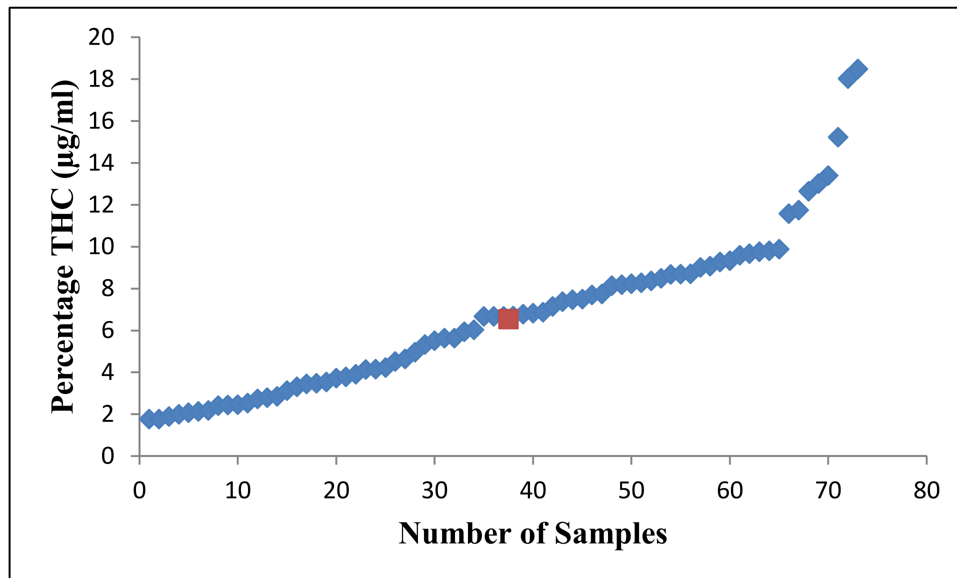
Using the linear curve equation from the standard curves on the ratio of cannabinoid to internal standard from the MS output was input for X. This value was then multiplied 50 for the dilution of extracts in MePy/MeOH, multiplied by 0.25 for the initial 250 µl extraction volume, and divided by 15000 for the initial 15 mg weight of sample. The final value was multiplied by 100 to produce a percentage. An example of a sample calculation is shown n Figure 11 for the percentage of THC in a sample.

Figure 11: Calculation of Percent THC from a *C. sativa* Sample

MS output THC/MePy ratio	Standard Curve(THC) (from Figure 10)	Dilution Factor, Extraction, Weight	Percent THC (µg/mL)
	y=0.0067x-0.0346	$X \frac{50 \times 0.25mL}{15000\mu g}$	X 100
1.012	$=\frac{(1.012+0.346)}{0.0067}$ =156.2089552	=0.13074129	=13.01%

Because *C. sativa* is often characterized in terms of the THC/CBD ratio, the ratios for THC and CBD were computed across all samples (Appendix B). Since the genetic analysis pertained to the investigation of THCA synthase, the primary focus of chemical analysis involved only THC. The percentages of THC within a sample ranged from 0.9% to 18% THC with an average of 6.54% (Figure 12).

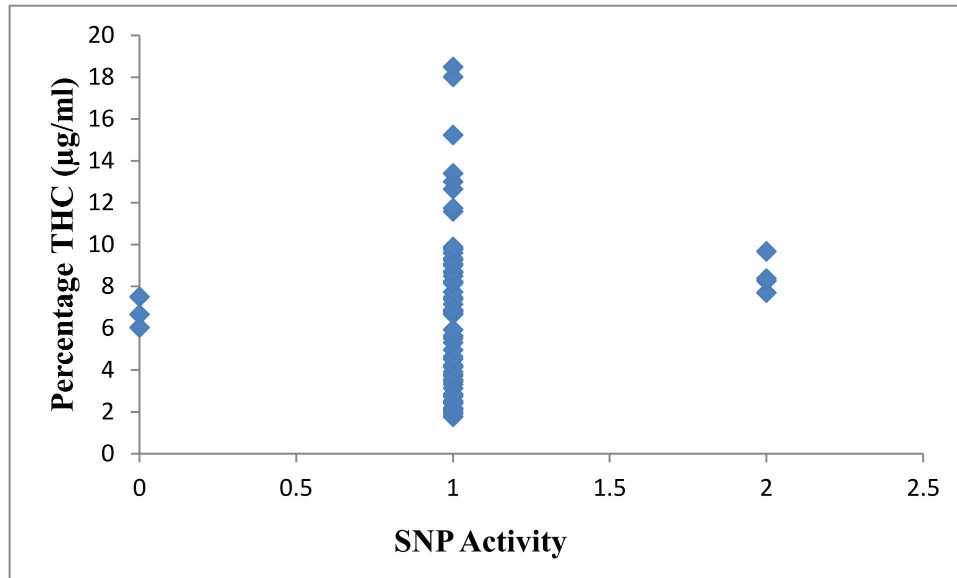
Figure 12: Range of Percent THC within *C. sativa* Samples



Displayed is a representation of the ranges of THC percentage within *C. sativa* samples, sorted from the lowest to highest concentration. The percentage of THC within 75 illicit *C. sativa* samples from Tulsa Police Department seizures ranged from 0.9% to 18%, shown in blue. The average percentage THC was 6.54%, shown in red.

In comparing the results from the gas chromatography with results from SNaPshot®, the comparison demonstrated that the majority of samples were heterozygous for SNaPshot®, as described previously. Figure 13 below shows the visual representation of comparisons between SNP activity and percentage THC.

Figure 13: Comparison of Percent THC and THCA Synthase SNP Activity



Within SNaPshot, the samples that typed as THCA Synthase active were labeled 0, the samples that typed as heterozygous were labeled 1, and the samples that typed as THCA synthase inactive were labeled 2. In comparing, the highest percentage of THC in the active samples was 7.5% and the lowest percentage of THC within the inactive samples was 7.7% (Appendix A).

Since the THC activity in theory should correlate with percentage of THC, it is of importance to note that the samples that typed as homozygous-inactive for THCA synthase had slightly higher percentages of THC than the samples that typed homozygous-active for THCA synthase, shown in Table 8.

Table 8: Comparison of SNP Activity with Percent THC

Sample	SNP Activity	%THC (µg/mL)
9961	0	6.04
6922	0	6.67
5725	0	7.51
9944	2	7.71
6043	2	8.28
7436	2	8.38
8811	2	9.68

Above shows a summary of comparisons of only samples that typed as active or inactive for THCA Synthase from SNaPshot results with %THC for each sample. Samples that were active were numbered 0, whereas samples that were inactive were numbered 2. Notice that a higher percentage of THC is present in THCA synthase Inactive samples.

.

Statistical Analysis

Results of Analysis of Variance

The ANOVA comparison of percent THC to SNP activity showed that there is no statistically significant difference in the percent THC of active (0), heterozygous (1), or inactive (2) samples. ANOVA results are depicted in Figure 14.

Figure 14: ANOVA Results Comparing Percent THC and SNP Activity

Type 3 Tests of Fixed Effects				
Effect	Numerator DF	Denominator DF	F Value	Pr > F
SNP	2	72	0.57	0.5687

Least Squares Means			
Effect	SNP	Average % THC	Standard Error
SNP	0 (Active)	6.7405	2.2033
SNP	1 (Heterozygous)	6.4250	0.4628
SNP	2 (Inactive)	8.5118	1.9081

The ANOVA results above compared the percentage THC across all SNP positions. Though it may seem like the difference between 8.5118 from SNP 2 and 6.4250 and 6.7405 from the other SNPs is large, the P value of 0.5687 demonstrates that there is no statistically significant difference between the average percentages of THC of the SNPs.

Results of Correlation Statistics

In order to asses which variables contributed to potential variation or correlation, the RFU ratios of each individual SNP position was calculated, as discussed in the methodology section. Using simple statistics, the mean, standard deviation, sum, minimum, and maximum values between 16 variables/combinations of variables and percentage THC were calculated. These values were used to calculate the Pearson's correlation coefficient for each variable. Not one variable or combination of variables revealed a statistically significant P value (Table 9).

Table 9: Pearson's Correlation Coefficient Results

Pearson's Correlation Coefficients		
N=75		
Variable	H0: Rno=0	Prob > |r|
PCT THC	1.00000	
PCT 16	0.04368	0.7098
PCT 9	0.04673	0.6906
PCT 8	0.03796	0.7464
PCT 17	0.00241	0.9836
PCT ALL	0.03352	0.7753
PCT 16, 9, 8	0.04291	0.7147
PCT 16, 9, 17	0.03188	0.7860
PCT 16, 8, 17	0.02869	0.8070
PCT 9, 8, 17	0.03017	0.7972
PCT 16, 9	0.04544	0.6987
PCT 16, 8	0.04084	0.7280
PCT 16, 17	0.02344	0.8418
PCT 9, 8	0.04237	0.7181
PCT 9, 17	0.02596	0.8251
PCT 8, 17	0.02119	0.8568

The results from Pearson's correlation coefficients above demonstrate that there is no statistically significant correlation between percentage THC and any individual/combination of SNP position ratio (percentage active/inactive peak height RFU, as discussed in Methodology).

Pearson's correlation coefficient demonstrates either negative correlation between two variables (0) or a positive correlation between two variables (1). The results of the Pearson's coefficient calculations demonstrate that there was no statistically significant correlation between percentage THC between and within all SNPs within a sample, between any SNP position within a sample, or between combinations of SNPs within a sample, or between any sample.

CHAPTER V

DISCUSSION AND CONCLUSIONS

Overview:

The principle goal of the current investigation was to determine if there was correlation between genetic variations within THCA synthase, or genotype, and of THC, or chemotype, in drug-type *Cannabis sativa* plants, also known as marijuana. The rationale for the study was derived from an article that concluded that THC composition within the plant could be determined from genetic analysis.[17] By expanding the methodology, the initial findings of other research was improved upon. A more in-depth discussion of each particular analytical method follows.

Cannabis sativa Sampling

Although permission was granted to obtain and work with *C. sativa* samples that were selected from TPD seizures, the original source of cannabis was subject to bias. Because illicit *C. sativa*, also known as marijuana, typically refers to samples with higher THC percentage, the selection of samples taken from the illicit market biased those samples to have higher percentages of THC.[27] Prior knowledge about source of growth, geographical origin, or absolute age of the samples was also unavailable and may have subjected results to bias. Furthermore, unlike the

study by Rotherham and Harbison, samples within the current study were not categorized based on GC-MS results prior to testing via SNaPshot® analysis.[17]

Another sampling factor that may have contributed to inconsistent results was the lack of negative control. Experimental design typically includes a negative control which is substance that behaves similar to the sample being tested, but lacks proper characteristics to produce results. In this experiment, a form of *C. sativa* known to lack high percentages of THC, called hemp, would serve as a promising negative control. This non-drug type of *C. sativa* was also used as a negative control in the article by Rotherham and Harbison.[17] However, in order to obtain a hemp sample from a legal source such as law enforcement, a DEA licensure is required. This license remains pending within OSU Forensic Sciences laboratories, and as such hemp samples were unavailable for chemical testing. DNA from hemp was available however, as it is possible to purchase hemp seeds for a food snack.[45] It should be noted that the genomic DNA from hemp seeds represents a population of molecules from different meiotic events as the material extracted was seeds. Since each seed is an individual plant, each would represent a different combination of genetic markers.

SNaPshot® Analysis

The method of the SNaPshot® reaction was altered in order to optimize results for the study. The occurrence of off-scale peaks in the early genetic analyses necessitated further dilution of initial C2E2 PCR products. Organic phenol:chloroform:isoamyl alcohol extraction was used in lieu of methods by Rotherham and Harbison because of required laboratory protocols within the student forensic DNA laboratory. The DNA was quantified using a Nanodrop spectrophotometer to ensure adequate amount of DNA was present for initial PCR. However, the concentration of DNA was not equivalent across all samples, which may have contributed to the need for various levels of dilution.

From the SNaPshot® results, the samples that typed as heterozygous varied in characteristics of the SNaPshot® profile in terms of the peak heights (Appendix A). Certain samples presented with higher active peaks than inactive, and vice versa. These results are demonstrated in the ratios of active/inactive peak height used for statistical analysis (Appendix A). Although samples were characterized only as active, heterozygous, or inactive, variations within level of activity could result in division of samples into subcategories heterozygous-active or heterozygous-inactive based upon the RFU ratios in the different heterozygous genotypes. The occurrence of subsets within heterozygote samples could be a result of possible polyploidy.

Polyploidy is the term used when a typically diploid organism contains, for example, one (triploid) or two (tetraploid) extra sets of chromosomes. Aneuploidy refers to organisms that contain an abnormal number of chromosomes; as such, these organisms may lack a chromosome or have an extra chromosome.[(46)] The occurrence of polyploidy within *C. sativa* has been attributed to the use of clonal propagation, which is the reason "why it is not possible to apply Hardy–Weinberg biostatistics."[(47)] Another potential cause of polyploidy and aneuploidy within *C. sativa* plants could be the use of biochemical modification. One example of modification includes soaking seeds or young plant roots in colchicine, that prevents cytokinesis in meiosis and results in the increase of chromosomes within the cell.[(46)] The occurrence of polyploidy and aneuploidy will obviously affect typical diploid zygosity, therefore preventing the use of traditional Hardy-Weinberg understanding of allelic balance and also affecting possible correlations between genotype and chemotype in marijuana plants.

Amplicon Sequencing Analysis

Along with SNaPshot® heterozygote ratios, the variation within samples that type as heterozygous for THCA synthase SNPs was also demonstrated in the results from the Ion Torrent PGM genomic sequencing. The percentages of nucleotides within heterozygote SNPs varied

across multiple SNP positions. One of the sequenced samples revealed a composition of 32% adenine and 68% guanine for the 8F SNP position. In some heterozygotes, the ratios were closer to 75%:25% for the different SNPs. If the sample typed as a true heterozygote per Hardy-Weinberg equilibrium, the percentages would appear closer to 50%:50%.

The polyploidy issue within the tested *C. sativa* samples could not be addressed in the current research since the flowering buds of the plant had already been dried. If the plants had been living or in better condition, karyotyping methods could have been employed to determine the number of chromosomes for the plant.[48] The determination of number of chromosomes within samples could have resolved much of the uncertainty regarding zygosity within the *C. sativa* plants investigated.

GC-MS Analysis

The matrix effects of THC, CBD, and CBN were originally tested using oregano as a negative control. However, use of oregano would not correlate with a negative *C. sativa* sample, since only *C. sativa* plants may have other cannabinoids that effect the composition of the THC, CBD, or CBN. Because non-drug *C. sativa*, or hemp, was unobtainable for chemical testing, the possibility of testing true matrix effects was not possible within Oklahoma State University School of Forensic Sciences laboratories. The Restek column that was used separates six total cannabinoids. Without using a true negative control to test for matrix effects, it was not possible to determine if other cannabinoids such as cannabichromene affected the results of THC, CBD, or CBN.

When observing the values of low controls, the 6.125 µg/mL measurements fell outside the range of +/-20% reproducibility. After the first day that control results were analyzed, an additional low control at 12.5 µg/mL was added. Results for this control were within acceptable limits.

It is also important to highlight factors that influence THC within *C. sativa*. The report by the UNODC states that THC undergoes degradation over time, and can degrade in certain environmental factors including light, air, and humidity.(11) Therefore, the true concentration of THC within the samples in the current study could have been affected by these factors because the storage conditions and age of samples prior to TPD seizure were unknown.

Correlation Analysis

The cannabinoid biosynthesis within *C. sativa* remains unclear. There is no conclusive evidence stating that the THCA synthase gene is the only means of synthesizing THC within *C. sativa*. The gaps within the *C. sativa* genome that remain to be investigated could potentially lead to more knowledge about cannabinoid biosynthesis. Thus, other undetermined factors may influence cannabinoid biosynthesis and accordingly production of THC.

The addition of SNPs within the THCA synthase gene other than those investigated here would increase the likelihood of determining activity and potential correlation of production of THC. Although the four SNPs investigated here did not correlate with THC production, other SNPs that were not investigated could also lead to deactivation of the THCA synthase gene via changes in the amino acid sequence. Until all of the SNPs within the THCA synthase gene are investigated simultaneously, the correlation between SNPs within the THCA synthase gene and THC concentration cannot be determined with 100% certainty.

Comparison of Results

Although the initial interest in the current study was a derivative of the Rotherham and Harbison study, the evidentiary samples selected by Rotherham and Harbison were from New Zealand, which may differ dramatically from the samples of the United States.(17) The study by Rotherham and Harbison reveals numerous errors in their results.

First, the authors state that only 51 of the 75 drug-type *C. sativa* samples were quantified via GC-MS. However, their results show that each sample had a concentration of THC that fell within certain ranges; these results are confounding with their previous statement. Although the authors state the presence of THC was confirmed via thin-layer chromatography for samples that were not quantified via GC-MS, this screening confirmation may not adequately quantify a concentration of THC, and should not be reported as a concentration.

Secondly, the authors describe their results from SNaPshot® in terms of SNP position with according nucleotide change and change in amino acid sequence. Rotherham and Harbison failed to include the change in amino acid sequence for the 17R SNP position. They also erroneously reported the amino acid change for both the 8F and 9F SNP positions. Their results state that 8F, at 1035 bp, changes a thymine to a guanine, which corresponds accurately with the results in the current study. However, they state that this change results in a change in the amino acid sequence from a Phenylalanine to a Leucine. The codon for Leucine lacks a guanine altogether. For the 9F position, the authors state that the nucleotide sequence changes a thymine in the active form to a adenine in the inactive form, which also corresponds accurately with the results in the current study. Yet the authors state that amino acid change for the 9F SNP position is from lysine to arginine, which is also incorrect. Lysine codons lack a thymine.

Lastly, the results of Rotherham and Harbison are confounding because the authors did not include a description of determination of peak variation within SNP positions, or a description of how they determined a baseline threshold. They did state that there was an artifact in the 16 SNP position within every SNaPshot® reaction.

In comparing results, the majority of the samples by Rotherham and Harbison were heterozygous, which is similar to the current study. It is difficult to compare results with the study by Rotherham and Harbison since the authors included the 15 hemp or non-drug type *C.*

*sativ*a samples within their results, whereas the current study only analyzed samples from forensic seizures. It would be interesting to compare results if they had separated their results into the samples that were previously determined to be drug-type.

Possible Future Studies

A potential consideration for future studies would be to take samples directly from mature plants to test for ploidy and to continue with SNaPshot® analysis once diploid zygosity was determined. In order to confidently determine these factors, it would be best to grow plants of known genotype under the same conditions and harvest the mature female inflorescence at the same time. Incorporation of all of the previously determined SNPs within the THCA synthase gene would also create a more definitive determination of amino acid sequence and subsequent level of THCA synthase activity. It would be preferable to sequence all of the samples that were processed with SNaPshot®. Funding in the sequencing area of the School of Forensics Science at OSU was limited, thus further sequencing of samples was not possible. Other options for genetic analysis would include the investigation of the CBD synthase gene within *C. sativa* samples.

Conclusion

The purpose of the present study was to investigate the potential correlation between percentage of Δ9-tetrahydrocannabinol and zygosity of single nucleotide polymorphisms (SNPs) within the THCA synthase gene. The use of the SNaPshot® minisequencing kit and gas chromatography-mass spectrometry allowed analysis of these components. Correlation statistics were performed between the percentages of THC from GC-MS data and the zygosity of SNP positions from SNaPshot® data. The statistical results showed that there was no correlation between percentage THC and zygosity of single nucleotide polymorphisms within the THCA synthase gene.

REFERENCES

1. Bostwick JM. Blurred boundaries: the therapeutics and politics of medical marijuana. *Mayo Clin Proc.* 2012;87(2):172-186. doi:10.1016/j.mayocp.2011.10.003.

2. Bonnie RJ, Whitebread II CH. The forbidden fruit and the tree of knowledge: an inquiry into the legal history of american marijuana prohibition. *Virginia Law Review.* 1970;56(6):971-1203. doi:10.2307/1071903.

3. Comprehensive Drug Abuse and Prevention Control Act, 21 USC §812 (1970)

4. Medical Marijuana. ProCon Web site. http://medicalmarijuana.procon.org/view.resource.php?resourceID=000881. Updated October 30, 2014. Accessed November 25, 2014

5. Leung L. Cannabis and its derivatives: review of medical use. *J Am Board Fam Med.* 2011;24(4):425-462. doi:10.3122/jabfm.2011.04.100280.

6. Atakan Z. Cannabis, a complex plant: different compounds and different effects on individuals. *Ther Adv Psychopharmacol.* 2012;2(6):241-254. doi:10.1177/2045125312457586.

7. Rajesh M, Mukhopadhyay P, Bátkai, S, et al. Cannabidiol attenuates high glucose-induced endothelial cell inflammatory response and barrier disruption. *Am J Physiol Heart Circ Physiol.* 2007;293(1):H610-619. doi:10.1152/ajpheart.00236.2007.

8. Gurley RJ, Aranow R, Katz M. Medicinal marijuana: a comprehensive review. *J Psychoactive Drugs.* 1998;30(2):137-147. doi:10.1080/02791072.1998.10399683.

9. Naftali T, Lev LB, Yablekovitz D, Half E, Konikoff F. Treatment of Crohn's disease with cannabis: an observational study. *Isr Med Assoc J.* 2011;13(8):455-458.

10. Fiz J, Durán M, Capellà D, Caronell J, Farré M. Cannabis use in patients with fibromyalgia: effects on symptoms relief and health-related quality of life. *PLoS One.* 2011;6(4):e18440. doi:10.1371/journal.pone.0018440.

11. Fraud N, Gottschalk U. Using chromatography to separate complex mixtures. *Chem Eng Prog.* 2010;106(12):27-31.

12. United Nations Office on Drugs and Crime. Recommended methods for the identification an analysis of cannabis and cannabis products. http://www.unodc.org/documents/scientific/ST-NAR-40-Ebook.pdf. Updated 2009. Accessed November 3, 2014.

13. Nucleic Acids Res. 2013 Jan;41(Database issue):D36-42. doi: 10.1093/nar/gks1195. Epub 2012 Nov 27.

14. Bakel H van, Stout JM, Cote AG, et al. The draft genome and transcriptome of *Cannabis sativa*. *Genome Biol.* 2011;12(10):R102. doi:10.1186/gb-2011-12-10-r102.

15. Gagne SJ, Stout JM, Liu E, Boubakir Z, Clark SM, Page JE. Identification of olivetolic acid cyclase from *Cannabis sativa* reveals a unique catalytic route to plant polyketides. *Proc Natl acad Sci U S A.* 2012;109(31):12811-12816. doi:10.1073/pnas.1200330109.

16. Kojoma M, Seki H, Yoshida S, Muranaka T. DNA polymorphisms in the tetrahydrocannabinolic acid (THCA) synthase gene in "drug-type" and "fiber-type" *Cannabis sativa L*. *Forensic Sci Int.* 2006;159(2-3):132-140. doi:10.1016/j.forsciint.2005.07.005.

17. Rotherham D, Harbison S. Differentiation of drug and non-drug cannabis using a single nucleotide polymorphism (SNP) assay. *Forensic Sci Int.* 2011;207(1-3):193-197. doi:10.1016/j.forsciint.2010.10.006.

18. Szaflarski JP, Bebin EM. Cannabis, cannabidiol, and epilepsy—from receptors to clinical response. *Epilepsy Behav.* 2014. doi:10.1016/j.yebeh.2014.08.135.

19. Miller DE. Narcotic Drug and Marihuana Controls. In: Washington, D.C.; 1966:1-13. Available at: http://files.eric.ed.gov/fulltext/ED026662.pdf. Accessed November 25, 2014.

20. Newbern AE. Good cop, bad cop: federal prosecution of state-legalized medical marijuana use after United States v. Lopez. *California Law Review.* 2000;88(5):1575-1634.

21. Marijuana resource center: state laws related to marijuana. Office of National Drug Control Policy. http://www.whitehouse.gov/ondcp/state-laws-related-to-marijuana. Accessed November 25, 2014

22. Grant I, Atkinson JH, Gouaux B, Wilsey B. Medical marijuana: clearing away the smoke. *Open Neurol J.* 2012;6:18-25. doi:10.2174/874205X01206010018.

23. Allegretti JR, Courtwright A, Lucci M, Korzenik JR, Levine J. Marijuana use patterns among patients with inflammatory bowel disease. *Inflamm Bowel Dis.* 2013;19(13):2809-2814. doi:10.1097/01.MIB.0000435851.94391.37.

24. Welty TE, Luebke A, Gidal BE. Cannabidiol: Promise and Pitfalls. *Epilepsy Curr.* 2014;14(5):250-252. doi:10.5698/1535-7597-14.5.250.

25. Ramesh D, Schlosburg JE, Wiebelhaus JM, Lichtman AH. Marijuana dependence: not just smoke and mirrors. *ILAR J.* 2011;25:295-308. doi:10.1093/ilar.52.3.295.

26. Weinstein A, Gorelick DA. Pharmacological treatment of cannabis dependence. *Curr Pharm Des.* 2011;17(14):1351-1358. doi:10.2174/138161211796150846#sthash.TK13ZBHM.dpuf.

27. Morgan CJ, Freeman TP, Schafer G, Curran HV. Cannabidiol attenuates the appetitive effects of D9-Tetrahydrocannabinol in humans smoking their chosen cannabis. *Neuropsychopharmacology.* 2010;35(9):1879-1885. doi:10.1038/npp.2010.58.

28. Shier AR de M, Ribeiro NP de O, Silva AC de O e, et al. Cannabidiol, a *Cannabis sativa* constituent, as an anxiolytic drug. *Braz J Med Biol Res.* 2012;34:S104-S117.

29. Gomes FV, Bel EAD, Guimarães FS. Cannabidiol attenuates catalepsy induced by distinct pharmacological mechanisms via 5-HT_{1A} receptor activation in mice. *Prog Neuropsychoparmacol Biol Psychiatry.* 2013;46(1):43-47. doi:10.1016/j.pnpbp.2013.06.005.

30. Zuardi A, Crippa J, Hallak J, Moreira F, Guimaraes F. Cannabidiol, a *Cannabis sativa* constituent, as an antipsychotic drug. *Braz J Med Biol Res.* 2006;39(4):421-429.

31. Taura F, Sirikantaramas S, Shoyama Y, Yoshikai K, Shoyama Y, Morimoto S. Cannabidiolic-acid synthase, the chemotype-determining enzyme in the fiber-type Cannabis sativa. *FEBS letters* 2007;581(16):2929-2934. doi:10.1016/j.febslet.2007.05.043.

32. Isomer. *Merriam-Webster Online Dictionary* 2015. Available at: http://www.merriam-webster.com/dictionary/isomer?show=0&t=1417913815. Accessed December 5, 2014.

33. Marks MD, Wenger JP, Omburo SN, et al. Identification of candidate genes affecting Δ -tetrahydrocannabinol biosynthesis in Cannabis sativa. *J Exp Bot.* 2009;60(13):3715-3726. doi:10.1093/jxb/erp210.

34. Tambaro S, Bortolato M. Cannabinoid-related agents in the treatment of anxiety disorders: current knowledge and future perspectives. *Recent Pat CNS Drug Discov*. 2012;7(1):25-40.

35. Holland MG, Schwope DM, Stoppacher R, Gillen SB, Huestis MA. Postmortem redistribution of Δ9-tetrahydrocannabinol (THC), 11-hydroxy-THC (11-OH-THC), and 11-nor-9-carboxy-THC (THCCOOH). *Forensic Sci Int.* October 10,;212(1-3):247-251. doi:10.1016/j.forsciint.2011.06.028.

36. Hazekamp A, Peltenburg A, Verpoorte R. Chromatographic and spectroscopic data of cannabinoids from Cannabis sativa L. *J Liq Chromatogr Relat Technol.* 2005;28(15):2361-2382. doi:10.1080/10826070500187558.

37. Galand N, Ernouf D, Montigny F, Dollet J, Pothier J. Separation and identification of cannabis components by different planar chromatography techniques (TLC, AMD, OPLC). *J Chromatogr Sci.* 2004;42(3):130-134. doi:10.1093/chromsci/42.3.130.

38. Matsunaga T, Nagatomo H, Yamamoto I, Yoshimura H. Identification and determination of cannabinoids in commercially available cannabis seeds. *Eisei kagaku.* 1990;36(6):545-547. doi:http://dx.doi.org/10.1248/jhs1956.36.545.

39. Linacre A, Thorpe J. Detection and identification of cannabis by DNA. *Forensic Sci Int.* 1998;91(1):71-76. doi:http://dx.doi.org/10.1016/S0379-0738(97)00173-4.

40. Docimo T, Consonni R, Coraggio I, Mattana M. Early phenylpropanoid biosynthetic steps in Cannabis sativa: link between genes and metabolites. *Int J Mol Sci.* 2013;14(7):13626-13644. doi:10.3390/ijms140713626.

41. Cascini F, Passerotti S, Boschi I. Analysis of THCA synthase gene expression in cannabisL a preliminary study by real-time quantitative PCR. *Forensic Sci Int.* 2013;231(1):208-212. doi:http://dx.doi.org/10.1016/j.forsciint.2013.05.019.

42. Sirikantaramas S, Morimoto S, Shoyama Y, et al. The gene controlling marijuana psychoactivity: molecular cloning and the heterologous expression of Δ1-tetrahydrocannabinolic acid synthase from Cannabis sativa L. *J Biol Chem.* 2004;279(38):39767-39774. doi:10.1074/jbc.M403693200.

43. Ion PGM™ System for Next-Generation Sequencing. Life Technologies. Thermo Fisher Scientific Inc. http://www.lifetechnologies.com/us/en/home/life-science/sequencing/next-generation-sequencing/ion-torrent-next-generation-sequencing-workflow/ion-torrent-next-generation-sequencing-run-sequence/ion-pgm-system-for-next-generation-sequencing.html?icid=fr-pgm-main. Accessed August 10, 2015.

44.Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative Genomics Viewer. *Nat Biotechnol*. 2011;29:24-26. doi:10.1038/nbt.1754.

45. Manitoba Harvest Hemp Hearts Raw Shelled Hemp Seeds, natural flavor, 1 Pound. Amazon.com, Inc. http://www.amazon.com/Manitoba-Harvest-Hearts-Shelled-natural/dp/B00856TSCC/ref=sr_1_1?rps=1&ie=UTF8&qid=1442257618&sr=8-1&keywords=hemp+seeds. Accessed September, 14, 2015.

46: Clark RC. *Marijuana Botany An Advanced Study: The Propagation and Breeding of Distinctive Cannabis*. Berkeley, CA: Ronin Publishing; 1981. https://www.thevespiary.org/library/Files_Uploaded_by_Users/llamabox/Clarke,%20Robert%20C%20-%20Marijuana%20Botany%20An%20Advanced%20Study.pdf. Accessed September 1, 2015.

47. Bagheri M, Mansouri H. Effect of induced polyploidy on some biochemical parameters in Cannabis sativa L. *Appl Biochem Biotechnol*. 2015;175(5):2366-2375. doi:10.1007/s12010-014-1435-8.

48. Köhnemann S, Nedele J, Schwotzer D, Morzfeld J, Pfeiffer H. The Validation of a 15 STR multiplex PCR for *Cannabis* species. *Int J Legal Med*. 2012;126(4):601-606. doi:10.1007/s00414-012-0706-6.

APPENDICES

Summary

The following appendices contain raw data from analyses. Appendix A contains data from SNaPshot® analysis, wherein the zygosity column refers to homozygous-active samples numbered as 0, heterozygous samples numbered as 1, and homozygous-inactive samples numbered as 2. Appendix B contains data from GC-MS Analyses, including Standard Curves, Contols, and sampling data. Appendix C contains data from Statistical Analyses.

APPENDIX A
SNaPshot® Results

Sample	16I	16A	Ratio A/I	9I	9A	Ratio A/I	8I	8A	Ratio A/I	17I	17A	Ratio A/I	Average SNP ratio	Zygosity*
254	923	443	0.3243045	3097	776	0.2003615	1626	520	0.2423113	551	475	0.462963	0.307485064	1
255	900	333	0.270073	2958	582	0.1644068	1794	428	0.1926193	612	443	0.4199052	0.261751062	1
275	328	233	0.4153298	1177	432	0.2684897	719	317	0.3059846	215	302	0.5841393	0.393485834	1
365	213	888	0.8065395	462	1823	0.7978118	247	1152	0.8234453	133	1051	0.8876689	0.828866391	1
367	237	524	0.6885677	616	932	0.6020672	406	750	0.6487889	160	583	0.7846568	0.681020145	1
452	629	250	0.2844141	2116	440	0.172144	1308	327	0.2	431	333	0.4358639	0.273105489	1
466	587	194	0.2483995	1975	356	0.1527242	1191	262	0.1803166	396	278	0.4124629	0.248475784	1
565	429	386	0.4736196	1106	667	0.3761985	742	545	0.4234654	275	435	0.6126761	0.471489911	1
597	98	661	0.8708827	206	1304	0.8635762	137	1025	0.8820998	50	749	0.9374218	0.888495126	1
661	361	452	0.5559656	1206	851	0.4137093	666	573	0.4624697	222	502	0.6933702	0.531378686	1
777	511	75	0.1279863	1726	129	0.0695418	1050	96	0.0837696	332	105	0.2402746	0.13039309	1
866	209	137	0.3959538	565	200	0.2614379	361	164	0.312381	152	164	0.5189873	0.37218999	1
956	495	314	0.3881335	1609	583	0.2659672	978	422	0.3014286	359	421	0.5397436	0.373818203	1
1072	509	195	0.2769886	1650	336	0.1691843	994	251	0.2016064	363	264	0.4210526	0.267207996	1
1187	452	330	0.4219949	1524	684	0.3097826	939	501	0.3479167	317	443	0.5828947	0.415647224	1
1220	1117	618	0.356196	3676	1142	0.2370278	1943	728	0.2725571	631	655	0.5093313	0.343778033	1
1250	439	404	0.4792408	1499	801	0.3482609	956	588	0.380829	298	528	0.6392252	0.461888968	1
1365	613	661	0.5188383	2015	1244	0.3817122	1074	801	0.4272	380	716	0.6532847	0.495258789	1
1497	1186	866	0.4220273	4253	1718	0.287724	2303	1127	0.3285714	706	955	0.5749548	0.403319391	1
1688	289	299	0.5085034	859	578	0.4022269	535	422	0.4409613	223	406	0.645469	0.49929015	1
1697	876	555	0.3878407	3189	1017	0.2417974	1699	677	0.2849327	533	600	0.5295675	0.361034571	1
1778	330	230	0.4107143	1026	431	0.2958133	643	325	0.3357438	234	295	0.557656	0.399981839	1
1868	142	224	0.6120219	479	406	0.4587571	283	298	0.5129088	109	285	0.7233503	0.576759488	1
1991	217	102	0.3197492	604	151	0.2	397	126	0.2409178	131	102	0.4377682	0.29960881	1

Sample	16I	16A	Ratio A/I	9I	9A	Ratio A/I	8I	8A	Ratio A/I	17I	17A	Ratio A/I	Average SNP ratio	Zygosity*
2227	1169	681	0.3681081	4048	1247	0.2355052	2125	796	0.2725094	722	730	0.5027548	0.344719384	1
2251	642	213	0.2491228	2241	360	0.1384083	1367	265	0.1623775	411	272	0.398243	0.237037902	1
2255	1023	260	0.20265	3375	426	0.1120758	2123	315	0.1292043	670	335	0.3333333	0.194315852	1
2517	840	327	0.2802057	2655	581	0.1795426	1642	431	0.2079112	590	435	0.4243902	0.273012446	1
2761	594	250	0.2962085	1979	452	0.1859317	1234	342	0.2170051	416	334	0.4453333	0.286119664	1
3056	416	317	0.4324693	1293	585	0.3115016	817	431	0.3453526	308	421	0.5775034	0.416706724	1
3579	323	116	0.2642369	1054	216	0.1700787	664	158	0.1922141	214	156	0.4216216	0.262037844	1
4344	510	301	0.3711467	1789	625	0.2589064	1134	461	0.2890282	355	404	0.5322793	0.36284016	1
4606	628	119	0.1593039	2208	202	0.0838174	1390	149	0.0968161	410	151	0.2691622	0.152274909	1
5075	150	251	0.6259352	452	528	0.5387755	291	386	0.5701625	109	336	0.7550562	0.622482333	1
5546	587	202	0.2560203	2017	378	0.1578288	1282	284	0.1813538	384	266	0.4092308	0.251108406	1
5725		283	1		654	1		483	1		383	1	1	0
5774	218	151	0.4092141	507	254	0.3337714	323	197	0.3788462	122	185	0.6026059	0.431109366	1
5888	576	187	0.2450852	2036	373	0.154836	1259	268	0.1755075	416	256	0.3809524	0.239095283	1
5931	221	150	0.4043127	703	292	0.2934673	428	213	0.3322933	156	195	0.5555556	0.396407213	1
6043	528		0	1995		0	1254		0	360		0	0	2
6264	460	268	0.3681319	1666	577	0.2572448	1028	421	0.2905452	321	366	0.5327511	0.362168231	1
6352	714	189	0.2093023	2447	329	0.1185159	1480	243	0.1410331	501	244	0.3275168	0.199092009	1
6361	270	212	0.439834	971	461	0.3219274	588	331	0.3601741	200	297	0.5975855	0.429880254	1
6922		526	1		1381	1		967	1		761	1	1	0
7326	299	255	0.4602888	960	496	0.3406593	654	382	0.3687259	210	312	0.5977011	0.441843792	1
7436	699		0	1795		0	1134		0	395		0	0	2
7916	377	318	0.457554	1371	700	0.338001	847	512	0.3767476	278	431	0.6078984	0.445050245	1
8058	445	227	0.3377976	1505	449	0.2297851	925	323	0.2588141	320	309	0.491256	0.329413185	1
8089	145	391	0.7294776	333	792	0.704	217	556	0.7192755	108	533	0.8315133	0.746066606	1
8144	234	160	0.4060914	781	346	0.3070098	563	248	0.3057953	180	215	0.5443038	0.390800061	1

Sample	16I	16A	Ratio A/I	9I	9A	Ratio A/I	8I	8A	Ratio A/I	17I	17A	Ratio A/I	Average SNP ratio	Zygosity*
8147	166	159	0.4892308	542	297	0.3539928	376	229	0.3785124	128	196	0.6049383	0.456668572	1
8159	332	228	0.4071429	1064	468	0.305483	687	334	0.3271303	233	310	0.5709024	0.402664636	1
8305	221	198	0.4725537	773	451	0.3684641	533	320	0.3751465	177	271	0.6049107	0.455268752	1
8434	352	186	0.3457249	1139	365	0.2426862	715	266	0.2711519	256	261	0.5048356	0.341099638	1
8577	50	176	0.7787611	157	355	0.6933594	97	263	0.7305556	50	238	0.8263889	0.75726622	1
8717	432	194	0.3099042	1173	299	0.203125	757	251	0.2490079	245	204	0.454343	0.304095019	1
8741	1280	519	0.2884936	4408	963	0.1792962	2372	624	0.2082777	743	587	0.4413534	0.279355229	1
8742	450	131	0.2254733	1465	261	0.1512167	1082	189	0.1487018	322	185	0.3648915	0.222570834	1
8769	1451	610	0.2959728	3702	852	0.1870883	2359	682	0.2242683	838	638	0.4322493	0.28489469	1
8811	367		0	825		0	545		0	189		0	0	2
8821	1002	394	0.282235	2576	542	0.1738294	1629	446	0.2149398	568	404	0.4156379	0.271660488	1
8829	744	273	0.2684366	2657	474	0.1513893	1442	319	0.1811471	427	307	0.4182561	0.254807279	1
8859	533	268	0.3345818	1665	444	0.2105263	1105	360	0.2457338	305	286	0.4839255	0.318691857	1
8916	253	180	0.4157044	783	367	0.3191304	524	278	0.3466334	184	231	0.5566265	0.409523686	1
9032	543	376	0.4091404	1818	641	0.2606751	1194	528	0.3066202	306	406	0.5702247	0.386665092	1
9134	643	428	0.3996265	1934	736	0.2756554	1263	598	0.3213326	389	462	0.5428907	0.38487632	1
9207	462	333	0.4188679	1434	591	0.2918519	937	485	0.3410689	282	363	0.5627907	0.403644848	1
9238	554	229	0.2924649	1524	341	0.1828418	1026	284	0.2167939	313	242	0.436036	0.282034158	1
9292	403	149	0.2699275	1367	297	0.1784856	961	220	0.1862828	303	197	0.394	0.257173981	1
9384	130	342	0.7245763	343	770	0.6918239	211	556	0.7249022	102	505	0.8319605	0.743315712	1
9563	272	274	0.5018315	821	539	0.3963235	520	407	0.4390507	192	369	0.657754	0.498739936	1
9660	1399	491	0.2597884	4807	863	0.1522046	2526	584	0.1877814	849	557	0.3961593	0.248983403	1
9933	112	459	0.8038529	227	876	0.7941976	178	707	0.7988701	50	514	0.9113475	0.827067027	1
9961		596	1		1040	1		675	1		649	1	1	0
9994	311		0	846		0	569		0	189		0	0	2

APPENDIX B
GC-MS Data

Concentration (µg/ml)	THC-1/ MEPY	THC-2/ MEPY	THC-3/ MEPY	AVG THC (3)	STD THC	Concentration (µg/ml)	CBD-1/ MEPY	CBD-2/ MEPY	CBD-3/ MEPY	AVG CBD (3)	STD CBD
1.06	0.004	0.004	0.005	0.004333333	0.00057735	1.06	0.014	0.015	0.013	0.014	0.001
3.125	0.01	0.009	0.01	0.009666667	0.00057735	3.125	0.033	0.035	0.031	0.033	0.002
6.25	0.021	0.022	0.023	0.022	0.001	6.25	0.078	0.079	0.074	0.077	0.002645751
12.5	0.049	0.049	0.055	0.051	0.003464102	12.5	0.182	0.186	0.176	0.181333333	0.005033223
25	0.111	0.109	0.124	0.114666667	0.008144528	25	0.434	0.43	0.411	0.425	0.012288206
50	0.256	0.241	0.283	0.26	0.021283797	50	1	0.965	0.932	0.965666667	0.034004902
100	0.578	0.547	0.661	0.595333333	0.058943476	100	2.257	2.16	2.149	2.188666667	0.059433436
200	1.326	1.215	1.496	1.345666667	0.14152856	200	4.957	4.65	4.848	4.818333333	0.155635257

Concentration (µg/ml)	CBN-1/ MEPY	CBN-2/ MEPY	CBN-3/ MEPY	AVG CBN (3)	STD CBN
1.06	0.037	0.036	0.032	0.035	0.002646
3.125	0.082	0.08	0.076	0.079333	0.003055
6.25	0.191	0.177	0.177	0.181667	0.008083
12.5	0.446	0.412	0.411	0.423	0.019925
25	1.04	0.948	0.935	0.974333	0.057239
50	2.395	2.131	2.203	2.243	0.13647
100	5.511	4.879	5.125	5.171667	0.318574
200	12.637	11.043	11.695	11.79167	0.801385

Controls-Day 1	CBD	THC	CBN
High (200)	185.0481928	239.6969697	192.3133
Med (50)	43.96385542	45.03030303	44.11709
Low (6.125)	7.979919679	8.666666667	8.89557
Controls-Day 2	CBD	THC	CBN
High (200)	195.982906	239.6969697	192.3133
Med (50)	45.85470085	45.03030303	44.11709
Low (6.125)	7.564102564	8.666666667	8.89557
Controls-Day 3	CBD	THC	CBN
High (200)	205.3950617	205.4	207.6706
Med (50)	43.08230453	46.73333333	43.67747
Low (6.25)	8.193415638	8.733333333	8.540956

Sample	CBD/ MePy	y=.0242x-.1156	diltuion, etc	%CBD	THC/ MePy	y=.0067x-.0346	diltuion, etc.	%THC	CBN/ MePy	y=.059x-.323	diltion, etc	%CBN
254	0	4.77686	0.003981	0.398072	1.012	156.209	0.130174	13.01741	0.043	6.20339	0.005169	0.516949
255	0.023	5.727273	0.004773	0.477273	0.91	140.9851	0.117488	11.74876	0.888	20.52542	0.017105	1.710452
275	0.007	5.066116	0.004222	0.422176	0.16	29.04478	0.024204	2.420398	0.354	11.47458	0.009562	0.956215
365	0	4.77686	0.003981	0.398072	0.559	88.59701	0.073831	7.383085	0.166	8.288136	0.006907	0.690678
367	0	4.77686	0.003981	0.398072	0.108	21.28358	0.017736	1.773632	0.186	8.627119	0.007189	0.718927
452	0.009	5.14876	0.004291	0.429063	1.191	182.9254	0.152438	15.24378	0.476	13.54237	0.011285	1.128531
466	0.374	20.2314	0.01686	1.68595	0.118	22.77612	0.01898	1.89801	0.463	13.32203	0.011102	1.110169
565	0	4.77686	0.003981	0.398072	0.28	46.95522	0.039129	3.912935	0.093	7.050847	0.005876	0.587571
597	0	4.77686	0.003981	0.398072	0.299	49.79104	0.041493	4.149254	0.024	5.881356	0.004901	0.490113
661	0	4.77686	0.003981	0.398072	0.19	33.52239	0.027935	2.793532	0.026	5.915254	0.004929	0.492938
777	0.405	21.5124	0.017927	1.7927	0.138	25.76119	0.021468	2.146766	0.151	8.033898	0.006695	0.669492
866	0	4.77686	0.003981	0.398072	0.649	102.0299	0.085025	8.502488	0.045	6.237288	0.005198	0.519774
956	0.002	4.859504	0.00405	0.404959	0.162	29.34328	0.024453	2.445274	0.104	7.237288	0.006031	0.603107
1072	0.006	5.024793	0.004187	0.418733	0.567	89.79104	0.074826	7.482587	0.203	8.915254	0.007429	0.742938
1187	0.004	4.942149	0.004118	0.411846	0.108	21.28358	0.017736	1.773632	0.453	13.15254	0.01096	1.096045
1220	0	4.77686	0.003981	0.398072	0.419	67.70149	0.056418	5.641791	0.028	5.949153	0.004958	0.495763
1250	0.019	5.561983	0.004635	0.463499	0.246	41.8806	0.0349	3.49005	0.456	13.20339	0.011003	1.100282
1365	0	4.77686	0.003981	0.398072	0.716	112.0299	0.093358	9.335821	0.062	6.525424	0.005438	0.543785
1467	0	4.77686	0.003981	0.398072	0.365	59.64179	0.049701	4.970149	0.05	6.322034	0.005268	0.526836
1688	0.004	4.942149	0.004118	0.411846	0.665	104.4179	0.087015	8.701493	0.382	11.94915	0.009958	0.995763
1697	0.08	8.082645	0.006736	0.673554	0.186	32.92537	0.027438	2.743781	0.027	5.932203	0.004944	0.49435
1778	0	4.77686	0.003981	0.398072	0.443	71.28358	0.059403	5.940299	0.194	8.762712	0.007302	0.730226
1868	0.022	5.68595	0.004738	0.473829	0.218	37.70149	0.031418	3.141791	0.744	18.08475	0.015071	1.507062
1991	0	4.77686	0.003981	0.398072	0.504	80.38806	0.06699	6.699005	0.044	6.220339	0.005184	0.518362

Sample	CBD/ MePy	y=.0242x-.1156	diltuion, etc	%CBD	THC/ MePy	y=.0067x-.0346	diltuion, etc.	%THC	CBN/ MePy	y=.059x-.323	diltion, etc	%CBN
2227	0	4.77686	0.003981	0.398072	0.394	63.97015	0.053308	5.330846	0.009	5.627119	0.004689	0.468927
2251	0.133	10.27273	0.008561	0.856061	0.133	25.01493	0.020846	2.084577	0.838	19.67797	0.016398	1.639831
2255	0.003	4.900826	0.004084	0.408402	0.511	81.43284	0.067861	6.78607	0.056	6.423729	0.005353	0.535311
2517	0.017	5.479339	0.004566	0.456612	0.067	15.16418	0.012637	1.263682	0.288	10.35593	0.00863	0.862994
2761	0.006	5.024793	0.004187	0.418733	0.737	115.1642	0.09597	9.597015	0.537	14.57627	0.012147	1.214689
3056	1.485	66.1405	0.055117	5.511708	0.243	41.43284	0.034527	3.452736	0.146	7.949153	0.006624	0.662429
3579	0.008	5.107438	0.004256	0.42562	0.3	49.9403	0.041617	4.161692	0.38	11.91525	0.009929	0.992938
4344	0.017	5.479339	0.004566	0.456612	0.663	104.1194	0.086766	8.676617	0.164	8.254237	0.006879	0.687853
4606	0.004	4.942149	0.004118	0.411846	0.27	45.46269	0.037886	3.788557	0.5	13.94915	0.011624	1.162429
5075	0.005	4.983471	0.004153	0.415289	0.983	151.8806	0.126567	12.65672	0.158	8.152542	0.006794	0.679379
5546	0.015	5.396694	0.004497	0.449725	0.126	23.97015	0.019975	1.997512	0.993	22.30508	0.018588	1.858757
5725	0	4.77686	0.003981	0.398072	0.569	90.08955	0.075075	7.507463	0.113	7.389831	0.006158	0.615819
5774	0.031	6.057851	0.005048	0.504821	0.037	10.68657	0.008905	0.890547	1.048	23.23729	0.019364	1.936441
5888	0.005	4.983471	0.004153	0.415289	0.666	104.5672	0.087139	8.71393	0.125	7.59322	0.006328	0.632768
5931	0.002	4.859504	0.00405	0.404959	0.196	34.41791	0.028682	2.868159	0.471	13.45763	0.011215	1.121469
6043	0.009	5.14876	0.004291	0.429063	0.631	99.34328	0.082786	8.278607	0.105	7.254237	0.006045	0.60452
6264	0.003	4.900826	0.004084	0.408402	0.519	82.62687	0.068856	6.885572	0.074	6.728814	0.005607	0.560734
6352	0	4.77686	0.003981	0.398072	0.339	55.76119	0.046468	4.646766	0.231	9.389831	0.007825	0.782486
6361	0	4.77686	0.003981	0.398072	0.69	108.1493	0.090124	9.012438	0.147	7.966102	0.006638	0.663842
6922	0	4.77686	0.003981	0.398072	0.502	80.08955	0.066741	6.674129	0.082	6.864407	0.00572	0.572034
7326	0.007	5.066116	0.004222	0.422176	0.624	98.29851	0.081915	8.191542	0.893	20.61017	0.017175	1.717514
7436	0.003	4.900826	0.004084	0.408402	0.639	100.5373	0.083781	8.378109	0.132	7.711864	0.006427	0.642655
7916	0.006	5.024793	0.004187	0.418733	0.502	80.08955	0.066741	6.674129	0.084	6.898305	0.005749	0.574859
8058	0	4.77686	0.003981	0.398072	0.515	82.02985	0.068358	6.835821	0.053	6.372881	0.005311	0.531073

Sample	CBD/ MePy	y=.0242x-.1156	diltuion, etc	%CBD	THC/ MePy	y=.0067x-.0346	diltuion, etc.	%THC	CBN/ MePy	y=.059x-.323	diltion, etc	%CBN
8144	0.004	4.942149	0.004118	0.411846	0.409	66.20896	0.055174	5.517413	0.332	11.10169	0.009251	0.925141
8159	0	4.77686	0.003981	0.398072	0.751	117.2537	0.097711	9.771144	0.534	14.52542	0.012105	1.210452
8305	0	4.77686	0.003981	0.398072	0.711	111.2836	0.092736	9.273632	0.614	15.88136	0.013234	1.323446
8434	0.012	5.272727	0.004394	0.439394	0.503	80.23881	0.066866	6.686567	0.633	16.20339	0.013503	1.350282
8577	0	4.77686	0.003981	0.398072	1.452	221.8806	0.1849	18.49005	0.165	8.271186	0.006893	0.689266
8717	0	4.77686	0.003981	0.398072	0.163	29.49254	0.024577	2.457711	0.06	6.491525	0.00541	0.54096
8741	0	4.77686	0.003981	0.398072	0.754	117.7015	0.098085	9.808458	0.036	6.084746	0.005071	0.507062
8742	0	4.77686	0.003981	0.398072	0.628	98.89552	0.082413	8.241294	0.446	13.0339	0.010862	1.086158
8769	0	4.77686	0.003981	0.398072	0.621	97.85075	0.081542	8.154229	0.033	6.033898	0.005028	0.502825
8811	0	4.77686	0.003981	0.398072	0.744	116.209	0.096841	9.68408	0.052	6.355932	0.005297	0.529661
8821	0.483	24.73554	0.020613	2.061295	0.169	30.38806	0.025323	2.532338	0.008	5.610169	0.004675	0.467514
8829	0	4.77686	0.003981	0.398072	0.761	118.7463	0.098955	9.895522	0.041	6.169492	0.005141	0.514124
8859	0	4.77686	0.003981	0.398072	0.266	44.86567	0.037388	3.738806	0.019	5.79661	0.004831	0.483051
8916	0	4.77686	0.003981	0.398072	0.232	39.79104	0.033159	3.31592	0.61	15.81356	0.013178	1.317797
9032	0	4.77686	0.003981	0.398072	0.588	92.92537	0.077438	7.743781	0.019	5.79661	0.004831	0.483051
9134	0	4.77686	0.003981	0.398072	0.696	109.0448	0.090871	9.087065	0.041	6.169492	0.005141	0.514124
9207	0	4.77686	0.003981	0.398072	1.043	160.8358	0.13403	13.40299	0.051	6.338983	0.005282	0.528249
9238	0	4.77686	0.003981	0.398072	0.419	67.70149	0.056418	5.641791	0.038	6.118644	0.005099	0.509887
9292	0	4.77686	0.003981	0.398072	0.541	85.91045	0.071592	7.159204	0.809	19.18644	0.015989	1.59887
9384	0.005	4.983471	0.004153	0.415289	0.251	42.62687	0.035522	3.552239	0.513	14.16949	0.011808	1.180791
9563	0	4.77686	0.003981	0.398072	0.306	50.83582	0.042363	4.236318	0.47	13.44068	0.011201	1.120056
9660	0	4.77686	0.003981	0.398072	0.329	54.26866	0.045224	4.522388	0.022	5.847458	0.004873	0.487288
9933	0	4.77686	0.003981	0.398072	0.897	139.0448	0.115871	11.58706	0.014	5.711864	0.00476	0.475989
9961	0	4.77686	0.003981	0.398072	0.451	72.47761	0.060398	6.039801	0.043	6.20339	0.005169	0.516949

APPENDIX C
Statistical Data

Dependent Variable: PCT THC		
Class Level Information		
Class	Levels	Values
SNP	3	0, 1, 2

Type 3 Test of Fixed Effects				
Effect	Numerator DF	Denominator DF	F Value	Pr>F
SNP	2	72	0.57	0.5687

Least Squares Means			
Effect	SNP	Estimate	Standard Error
SNP	0	6.7405	2.2033
SNP	1	6.4250	0.4628
SNP	2	8.5118	1.9081

Simple Statistics						
Variable	N	Mean	Std Sev.	Sum	Minimum	Maximum
PCT THC	75	6.54892	3.79387	491.16915	0.89055	18.49005
PCT 16	75	0.40708	0.21612	30.53107	0	1.00000
PCT 9	75	0.31301	0.22928	23.47596	0	1.00000
PCT 8	75	0.34172	0.22847	25.62883	0	1.00000
PCT 17	75	0.53234	0.21083	39.92566	0	1.00000
PCT All	75	0.39854	0.21878	29.89038	0	1.00000
PCT 16_9_8	75	0.35394	0.22394	26.54529	0	1.00000
PCT 16_9_17	75	0.41748	0.21606	31.31090	0	1.00000
PCT 16_8_17	75	0.42705	0.21636	32.02852	0	1.00000
PCT 9_8_17	75	0.39569	0.21979	29.67681	0	1.00000
PCT 6_9	75	0.36005	0.22179	27.00352	0	1.00000
PCT 16_8	75	0.37440	0.22178	28.07995	0	1.00000
PCT 16_17	75	0.46971	0.21223	35.22837	0	1.00000
PCT 9_8	75	0.32737	0.22875	24.55239	0	1.00000
PCT 9_17	75	0.42268	0.21617	31.70081	0	1.00000
PCT 8_17	75	0.43703	0.21665	32.77724	0	1.00000

Pearson Correlation Coefficients, N = 75		
Prob > |r| under H0: Rho=0		
Variable	PCT THC	R^2
PctTHC	1.00000	
Pct16	0.04368	0.7098
Pct9	0.04673	0.6906
Pct8	0.03796	0.7464
Pct17	0.00241	0.9836
PctAll	0.03352	0.7753
Pct16_9_8	0.04291	0.7147
Pct16_9_17	0.03188	0.7860
Pct16_8_17	0.02869	0.8070
Pct9_8_17	0.03017	0.7972
Pct16_9	0.04544	0.6987
Pct16_8	0.04084	0.7280
Pct16_17	0.02344	0.8418
Pct9_8	0.04237	0.7181
Pct9_17	0.02596	0.8251
Pct8_17	0.02119	0.8568

VITA

Lindsey Nicole Allen

Candidate for the Degree of

Master of Science

Thesis: INVESTIGATION OF GENOTYPE/CHEMOTYPE CORRELATIONS IN *CANNABIS SATIVA*

Major Field: Forensic Sciences

Biographical:

Education:

Completed the requirements for the Master of Arts in your major at Oklahoma State University, Center for Health Sciences, Oklahoma in September ,2015.

Completed the requirements for the Bachelor of Science in Organismic Biology with a minor in chemistry at Northeastern State University, Broken Arrow, Oklahoma in December, 2011.

Experience: Worked as a Research Assistant at Oklahoma State University-Center for Health Sciences. Worked as a Drug Screen Analyst at Advanced Pain Specialists of Tulsa.

Professional Memberships: Member of the American Academy of Forensic Science.