APPLICATIONS OF FORENSIC PLANT SCIENCE IN DRUG TRAFFICKING AND ENVIRONMENTAL CRIMES

_____

A Dissertation

Presented to

The Faculty of the Department of Forensic Science

Sam Houston State University

_____

In Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

_____

by

Madeline G. Roman

May, 2020

APPLICATIONS OF FORENSIC PLANT SCIENCE IN DRUG TRAFFICKING AND ENVIRONMENTAL CRIMES

by

Madeline G. Roman

_____

APPROVED:

Rachel Houston, PhD
Committee Director

Bobby LaRue, PhD
Committee Co-Director

Sheree Hughes-Stamm, PhD
Committee Member

Aaron Lynne, PhD
Committee Member

Phillip Lyons, PhD
Dean, College of Criminal Justice

DEDICATION

I would like to thank my advisor, Dr. Rachel Houston, for all her support and encouragement and the late nights she spent helping to prepare this dissertation. I can't imagine anyone better to guide me through this journey and prepare me for the future. Thank you for your patience and understanding through the times when everything that possibly could go wrong did.

This accomplishment would not have been possible without the support of my family and friends. Thank you for believing in me and listening to me rant about the challenges of scientific research for hours on end. Mom and Dad, I'm proud to join the ranks as another Dr. Roman. Thank you for encouraging me to dream big and work hard to achieve my goals. Lastly, thank you to my fellow graduate students, especially Team DNA, for all your help and for making this journey an enjoyable one.

ABSTRACT

Roman, Madeline G., *Applications of forensic plant science in drug trafficking and environmental crimes*. Doctor of Philosophy (Forensic Science), May, 2020, Sam Houston State University, Huntsville, Texas.

Forensic plant science deals with the use of plants as evidence in court. Plant genetic techniques, such as DNA barcoding or DNA fingerprinting, can be used to combat trafficking of illicit drugs by providing leads for law enforcement concerning entry points into the country and linking cases. Additionally, they can be used in court as evidence of environmental crimes, including illegal logging, which often go unpunished due to lack of forensic evidence. DNA barcoding is a technique that involves sequencing regions of the genome to identify a species or population of origin (biogeographical origin), and DNA fingerprinting involves individualizing samples based on their unique genetic profile, usually by using short tandem repeat markers (STRs).

Cannabis sativa L. is the source of both an illegal drug, marijuana, and a legal crop, hemp. Marijuana is the most commonly used illicit drug in the United States, and due to state-specific legalization of the drug, law enforcement must prevent and investigate trafficking of marijuana between states, as well as from international sources (e.g., at the border with Mexico). Current methods of identifying *C. sativa* use microscopic features of the plant or quantify delta-9-tetrahydrocannabinol (THC), the psychoactive component of marijuana. A DNA barcoding method could assist investigations by indicating the biogeographical origin and crop type of a sample and providing a means for linking cases from common growers and distributors. In the first phase of this study, seven polymorphic regions in the chloroplast genome of *C. sativa* were reported and explored as DNA barcodes for determining biogeographical origin and crop type. An MPS assay was then

developed to genotype these hotspots in a high throughput manner, which will facilitate the creation of a worldwide haplotype database, similar to the model of human mitochondrial haplotypes.

Additionally, single nucleotide polymorphisms (SNPs) in the tetrahydrocannabinolic acid (THCA) synthase gene were evaluated for their ability to distinguish between marijuana and hemp. The majority of marijuana samples and hemp flowers were classified correctly; however, other variables influence cannabinoid content in *C. sativa,* resulting in incorrect classifications for some sample types (i.e., hemp seeds and cannabigerol strains). Quantification of THC and THCA is the gold standard for distinguishing between marijuana and hemp, but several sample types (including juvenile plants, seeds, roots, and trace residues) may yield inconclusive chemical results. An alternate DNA approach should be taken with these samples, and the chloroplast DNA barcoding regions proposed in this dissertation may offer a viable future approach.

Papaver somniferum (opium poppy) is the source of opiates and opioids, a class of narcotic drugs with high abuse potential. Users who become addicted after being prescribed opiates may turn to alternatives, such as heroin or poppy seed tea, once their prescriptions end. There is currently no forensic method for genetically individualizing samples in cases of poppy seed tea overdoses. The Drug Enforcement Agency's Heroin Signature Program uses chemical analyses to determine the origin of heroin samples; however, addition of a genetic method would supplement the program and be capable of analyzing difficult sample types (such as trace residues found on drug paraphernalia). Three DNA extraction methods were evaluated for poppy seeds, and a novel quantitative real-time PCR assay was developed and validated for future genetic studies involving *P. somniferum*. STR markers

from the literature were evaluated, and a preliminary STR multiplex was used in a proof-of-concept study to show the potential of future STR panels for individualizing or determining biogeographical origin of heroin or poppy seed tea samples.

Eucalyptus is a genus of gum trees (eucalypts) planted around the world for use in the production of paper pulp, hardwood, essential oils, and other industrial products. Illegal logging of eucalypts and other trees costs the world economy billions of dollars annually, and cases of wood theft are often dismissed due to a lack of forensic evidence. Over 1,200 STRs have been discovered in eucalypts, but there has been a lack of forensic research testing these markers for evidence of illegal logging. This project evaluated nine STR markers for *Eucalyptus* and applied them to a case of illegal logging to demonstrate the utility of STR analysis for providing evidence in court.

KEY WORDS: Forensic plant science, *Cannabis sativa*, *Papaver somniferum*, *Eucalyptus*, Illegal logging, DNA barcoding, DNA fingerprinting, Short tandem repeats, Massively parallel sequencing, Chloroplast DNA.

ACKNOWLEDGEMENTS

The author would like to thank Blake Young for his work on the opium poppy project and Jessica Cheng for her assistance in the laboratory and for providing the chemical data for the THCA synthase project. The author would also like to thank Alejandra Figueroa for her kind donation of marijuana DNA extracts and Dr. Madeleine Gates for providing the poppy seed samples. Thank you also to Dr. Brandt Cassidy for his helpful suggestions on the poppy project and Dr. Abishek Muralidhar for his assistance with the PDQeX *phyto*GEM Plant Kit.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

ABBREVIATIONS

AFLP	Amplified fragment length polymorphism
bp	Base pairs
CBC	Cannabichromene
CBCA	Cannabichromene acid
CBD	Cannabidiol
CBDA	Cannabidiolic acid
CBG	Cannabigerol
CBGA	Cannabigerolic acid
CE	Capillary electrophoresis
CO1	Cytochrome c oxidase subunit 1
cpDNA	Chloroplast DNA
CSA	Controlled Substances Act
CTAB	Cetyltrimethyl ammonium bromide
ddNTP	Dideoxynucleotide triphosphates
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphates
EST-SSR	Expressed sequence tag simple sequence repeat
F_{ST}	Fixation index
GC-MS	Gas chromatography-mass spectrometry
HDMP	Heroin Domestic Monitoring Program
HPS	Heroin Signature Program
INDEL	Insertion/deletion polymorphism
ISFG	International Society of Forensic Genetics
ISSR	Inter simple sequence repeat
MPS	Massively parallel sequencing
mRNA	Messenger ribonucleic acid
mtDNA	Mitochondrial DNA
nDNA	Nuclear DNA
NGS	Next generation sequencing

NIDA	National Institute on Drug Abuse
OSAC	Organization of Scientific Area of Committees
PCA	Principal component analysis
PCR	Polymerase chain reaction
qPCR	Quantitative PCR
RAPD	Random amplified polymorphic DNA
rDNA	Ribosomal DNA
RFLP	Restriction fragment length polymorphism
SNP	Single nucleotide polymorphism
SSR	Simple sequence repeat
STR	Short tandem repeat
SWGDAM	Scientific Working Group on DNA Analysis Methods
THC	Tetrahydrocannabinol
THCA	Tetrahydrocannabinolic acid
VNTR	Variable number tandem repeat

GLOSSARY

Allele	Alternate versions of a gene found at the same locus.
Capillary electrophoresis	An electrokinetic separation technique by which DNA fragments are separated by size and detected by a laser that measures fluorescence.
DNA barcoding	A technique that assesses the sequence of a specific location of the genome and compares it to a database to identify species or populations.
DNA fingerprinting	A phrase coined by Jeffreys in 1985 to describe a method by which polymorphic loci are assessed, resulting in a DNA profile unique to a single individual (also called individualization).
Massively parallel sequencing	A method of determining the nucleic acid sequence of DNA targets that is capable of analyzing hundreds of targets in hundreds of samples simultaneously. Also referred to as next generation sequencing.
Polymerase chain reaction	A method used to make copies of a specific DNA sequence. PCR cycles consist of a series of temperature changes to separate DNA strands (denaturation), allow primers to anneal to the DNA (annealing), and allow the polymerase to incorporate nucleotides (extension).

CHAPTER I

Introduction

Forensic plant science

Overview

Forensic plant science, or more traditionally "forensic botany," is a branch of forensic science that deals with plant life and its use as evidence in court [1]. Forensic plant science consists of multiple subdisciplines, including plant anatomy and taxonomy, plant ecology, palynology, limnology, and molecular plant biology [1-3]. Plant anatomy and taxonomy involves the use of anatomical and morphological features to classify unknown plant material as belonging to a particular species or family. This can be used to link suspects or victims to specific locations. Plant ecology deals with the relationship of plants and their environment, including other organisms. Additionally, plant succession patterns can be used to elucidate the postmortem interval; for example, weeds underneath a cadaver will discolor and die after a minimum period of time due to lack of sunlight [4]. Palynology is the study of pollen, and identification of a particular species' pollen grains on a suspect or victim may link them to a crime scene. Limnology, the study of freshwater ecology, typically involves the identification of diatoms, microscopic algae present in the lungs and other organs of drowning victims. Lastly, molecular plant biology is the use of DNA techniques, including DNA fingerprinting and barcoding, to identify species, populations, or individuals. Molecular plant biology is the focus of this dissertation.

History

Plant material was first used as forensic evidence in the Lindbergh kidnapping case in 1932, which is considered the birth of forensic plant science [5, 6]. The 20-month-old

son of the American aviator Charles Lindbergh and Anne Morrow Lindbergh was abducted on the evening of March 1, 1932 from his second-floor nursery in Hopewell, New Jersey. The kidnapper left behind a ransom note demanding $50,000, traces of mud on the nursery floor, footprints in the mud under the nursery window, and a ladder, which had broken during the ascent or descent. A total of twelve ransom notes were delivered in the following months, and after payment on April 2, a thirteenth note contained instructions to retrieve the child from a boat near Martha's Vineyard, Massachusetts. A search based on the instructions revealed neither the boat nor the baby.

On May 12, the child's body was found partially buried near a highway less than five miles from the Lindbergh home. The body was badly decomposed, the skull was crushed, and parts of the body were missing. The coroner found that death had occurred two months previously and the cause of death was a blow to the head.

Investigators originally identified the suspect, Bruno Richard Hauptmann, by tracking the gold certificates used to pay the ransom. Among the evidence against him at trial was the ladder left at the crime scene. The police had observed that the ladder was crudely built and brought in a wood expert, Arthur Koehler from the Forest Service, United States Department of Agriculture. He identified the types of wood used, examined toolmarks and nail hole patterns, and concluded that some of the wood was likely to have been used previously in indoor construction. After Hauptmann was arrested, investigators matched the toolmarks on the ladder to tools owned by Hauptmann and matched the wood to flooring in his attic [5, 6]. Other evidence, such as handwriting analysis, were also important in the trial. The jury found Hauptmann guilty of murder in the first degree and sentenced him to death. The ruling was later upheld by the Supreme Court, and Hauptmann

was executed by electrocution on April 3, 1936. This case brought forensic plant science to the attention of the American public, and it has since been used to provide evidence in other types of cases. For example, identification of pollen signatures on a suspect's clothing compared to those of an alleyway where a crime was alleged to have taken place allowed investigators to determine that a suspect had been at the crime scene [7]. Individualization of plants by their DNA fingerprints was first used in court in 1992 in Arizona, when seed pods in the back of a suspect's truck were genetically matched to a particular tree where the body of a woman was discovered [8]. However, plant evidence is not routinely used in forensic investigations, and more research is needed to fully explore the applications of forensic plant science.

Applications

Forensic plant science has a range of applications. As evidenced by the Lindbergh case, it has the potential to link a suspect to a crime scene [5]. It can also provide evidence of drowning in freshwater by identification of diatoms in the lungs and organs of drowning victims [9], and identification of pollen grains or other plant matter may link a suspect or victim to a specific location or type of location (e.g., climate zone) [7]. Identification of plant matter in the stomach of a deceased person may corroborate witness statements or alibis based on the victim's last meal [10]. Forensic plant science is also used in the taxonomic identification of poisonous plants, such as poison hemlock (*Conium maculatum* L.), deadly nightshade (*Atropa belladonna*), strychnine (*Strychnos ignatii*), and ricin (*ricinius communis*), to name a few [1]. Similarly, it can be used in the identification of illegal drugs of plant origin (e.g., heroin, marijuana, and peyote) or to identify environmentally important or protected species [1, 2].

Drug Trafficking

Many drugs of plant origin are addictive and produce euphoria and hallucinations. They are also easy and inexpensive to cultivate, extract, and purify, making them staples in the drug trade. These drugs may be lethal in high doses, and consequently, identification of plant material is important in death investigations due to drug overdose.

Cannabis sativa L. is the source of the drug marijuana. The principal psychoactive component is delta-9-tetrahydrocannabinol (THC), naturally present in the plant [11]. In the United States, some states allow the medicinal use of marijuana for treatment of a number of conditions. It has been shown to reduce anxiety, relieve inflammation and pain, control nausea and vomiting from cancer chemotherapy, and stimulate appetite in people with cancer [12]. However, it is also the most used recreational drug in the United States [13] due to the euphoric and hallucinogenic properties of THC.

One of the most important classes of illegal drugs comes from *Papaver somniferum*, the opium poppy. The latex of the plant contains compounds known as opiates, including morphine, codeine, and thebaine [14]. While morphine and codeine are prescribed medicinally as analgesics (pain-killers), they are highly addictive and are also considered drugs of abuse. Additionally, the narcotic heroin is derived from morphine.

Several species of cacti are also known to produce psychoactive drugs. The most notable is peyote (*Lophophora williamsii*), which contains the alkaloid mescaline [15]. It is used in religious ceremonies as well as recreationally and medicinally.

Forensic plant science can be used to taxonomically identify drugs of plant origin in some cases. It can also provide tools to genetically associate plants with a criminal network consisting of growers, producers, couriers, suppliers, and dealers. Drug poisoning

is the leading cause of injury death in the United States according to the DEA's National Drug Threat Assessment (October 2018) [16], and drug trafficking crimes affect the safety of individuals and communities. Additionally, drug trafficking is often associated with other crimes, including corruption, money laundering, and transport of other illicit products.

Environmental Crime

The term environmental crime encompasses a broad array of illicit activities that break laws protecting ecological and wildlife resources. These include illegal wildlife trade, smuggling of ozone-depleting substances, illicit trade of hazardous waste, illegal and unregulated fishing, and illegal logging. These crimes affect the quality of air, water, and soil and threaten the survival of species. Forensic plant science can be used to provide evidence of environmental crimes such as illegal logging, which involves the harvesting and trade of timber products in violation with national laws [17]. It is estimated that 15 to 30 percent of timber traded around the world is obtained through illegal means, and these actions lead to deforestation, habitat destruction, species extinction, and loss of revenue for producer countries [18]. Forensic plant science can provide evidence of environmental crimes by identifying protected species or determining when and where a specimen (such as a tree log) was cut.

Molecular plant biology

Molecular plant biology is the use of DNA techniques to identify species, populations, or individuals. The two most-used techniques are DNA barcoding, which identifies sequences or alleles specific to a species or population, and DNA fingerprinting,

which identifies DNA profiles unique to an individual. Plant DNA techniques are similar to those used for human individualization and ancestry and identification of animal species.

DNA extraction from plant material

The first step in any genetic analysis is extraction of the DNA. Plant cells differ from animal cells in that they are surrounded by a cell wall composed mainly of cellulose, a polysaccharide made of glucose molecules [19]. This cell wall strengthens plant cells against water pressure, and it must be ruptured before DNA can be extracted. Commonly, physical destruction of the cell wall is performed by freezing the sample in liquid nitrogen and grinding with a mortar and pestle or homogenizer [20].

Following disruption of the cell wall, the debris can be centrifuged to remove it from solution. This is important because polysaccharides found in the cell wall may inhibit downstream PCR applications [20]. Other common inhibitors found in plant cells include tannins and other polyphenols, which bind to DNA and inhibit PCR [20]. Due to the presence of these inhibitory compounds, some extraction procedures call for the use of cetyltrimethyl ammonium bromide (CTAB), beta-mercaptoethanol (BME), and polyvinylpyrrolidone (PVP) to assist with their removal. CTAB, a detergent, assists with the separation of polysaccharides; BME breaks down protein bonds; and PVP forms hydrogen bonds with polyphenolics to prevent them from binding to DNA [19].

The CTAB method for DNA extraction [19, 21, 22] is an effective way to purify high-quality DNA from plant cells. However, it requires the use of hazardous solvents, such as phenol or chloroform, to separate proteins and other soluble contaminants from the DNA. Therefore, silica-based methods have become increasingly popular. QIAGEN's DNeasy® Plant Mini Kit [23] is commercially available and can be used to isolate pure

total DNA in under an hour, while effectively removing PCR inhibitors. Silica-based methods work by binding DNA to a silica column and, through multiple wash steps, removing inhibitors by allowing them to pass through the column. Pure DNA is then eluted from the column in a low-salt buffer or water.

DNA has successfully been extracted from many parts of the plant, including leaves, stems, roots, flowers, and seeds [24-27].

Genomes

Plants contain three genomes: nuclear or autosomal DNA, housed in the cell nucleus, and two organelle genomes, housed in the mitochondria and the chloroplasts. Nuclear DNA, which is unique to each individual, is the target of DNA fingerprinting techniques, while organelle DNA is the target of most DNA barcoding techniques.

Nuclear DNA (nDNA)

Nuclear DNA is stored in the nucleus of cells, and it is inherited (usually) from two parents in dioecious plants (plants that have both male and female individuals). Some plants demonstrate polyploidy, in which their cells have more than two homologous sets of chromosomes. Polypoloidy can occur naturally, arise through hybridization, or be induced by chemical means, such as with colchicine [28]. All plant species examined in this dissertation are naturally diploid organisms, which contain two sets of chromosomes (2x), so the following sections will focus on diploid organisms.

Because of biparental inheritance and genetic recombination, each individual contains a unique nuclear genome. In humans, the genome of each individual is 99.7% identical to every other human on earth, with variation occurring only in that remaining 0.3% [29]. Similarly, each individual plant has a genome that is largely conserved among

members of its species. Only specific polymorphic locations (loci), therefore, are useful for individualization, and these loci are the target of a technique called DNA fingerprinting.

Organelle DNA

Organelle DNA exists outside of the cell nucleus, in the mitochondria or chloroplast. Unlike nDNA, which is linear, organelle DNA is circular. Because a single cell usually contains multiple mitochondria and chloroplasts, each cell contains multiple copies of mitochondrial or chloroplast DNA (mtDNA or cpDNA) [1]. Mitochondrial genes are primarily involved in respiration and metabolism [1], similar to animals, and chloroplast genes are primarily involved in photosynthesis [1]. In humans, mtDNA is inherited uniparentally from the mother. In plants, the pattern of mtDNA and cpDNA inheritance varies between species and may be maternal, paternal, or biparental [30, 31].

In human forensics, mtDNA is used for typing of degraded DNA (e.g., bone) since it is present in higher copy numbers than nDNA and its circular structure makes it resistant to degradation [32]. However, because of maternal inheritance, DNA tests cannot differentiate between relatives of a maternal lineage, such as between a mother and her sons and daughters. Additionally, mtDNA is helpful to predict human ancestry. Previous research has shown that single nucleotide polymorphism (SNP) mutations in the mtDNA, which occur only once every 10^8 generations, can become "fixed" in a population. SNPs in the mtDNA, therefore, can be used to predict ancestry [33]. Lao et al. [34] and Novembre et al. [35] were able to pinpoint individuals' geographic location within Europe based on mitochondrial SNP analysis. A similar technique, called DNA barcoding, is useful for determining the genetic population of origin of individual plants, referred to as the biogeographical origin [36-38].

DNA fingerprinting

The nuclear genome of every individual is unique and varies at specific polymorphic loci. Types of polymorphic loci include base substitutions, also known as single nucleotide polymorphisms (SNPs), and microsatellite markers, short repetitive sequences of DNA commonly referred to as simple sequence repeats (SSRs) among botanists or short tandem repeats (STRs) among forensic scientists (hereafter referred to as STRs). Additionally, insertion-deletion markers known as INDELs are less frequently used. The term "DNA fingerprinting" was coined by Jeffreys et al. [39] in 1985 to describe an individual-specific DNA profile in humans resulting from the typing of minisatellite markers (also known as variable number of tandem repeats, or VNTRs), which are repetitive sequences slightly larger than the microsatellites, which are now in common use. DNA fingerprinting techniques in plants closely mirror those used in human identification.

Restriction fragment length polymorphisms (RFLP)

The RFLP technique was first used in the 1980s [39], and it involved cutting DNA samples with various restriction enzymes (restriction endonucleases) and performing a Southern immunoblot, followed by autoradiography to detect fragments of different sizes. Restriction enzymes cut DNA at specific target sequences, also called recognition sites, which differ between enzymes. SNPs or INDELs in the DNA of different individuals results in varying cutting patterns due to creation or abolishment of the recognition sites of different enzymes, and VNTRs (or other repetitive sequences) affect the length of fragments. This results in differing band patterns when the DNA fragments are separated by size. This method was time-consuming and did not always produce enough variation to differentiate individuals, but it was useful for the development of genetic maps in plant

species [40] as well as identifying relationships among different cultivars [41, 42]. Since the advent of polymerase chain reaction (PCR), other, more sensitive and higher-resolution techniques have taken its place.

Random amplified polymorphic DNA (RAPD)

RAPD markers have been used for paternity testing, taxonomic identification, population genetics studies, and genetic diversity [43, 44]. Short, random primers are used in PCR amplification, and resulting fragment patterns are analyzed and compared to known samples. RAPD results, however, are sometimes difficult to interpret and results may vary from laboratory to laboratory based on their specific protocols and random primers used.

An example of the use of RAPD markers in forensic plant science is seen in a murder where seed pods from the Palo Verde tree (*Cercidium floidum*) placed the suspect in the vicinity of the crime scene [45]. RAPD analysis of the seed pods found in the suspect's truck matched exactly the profile of a tree near the murder site.

Inter-simple sequence repeats (ISSR)

ISSR markers are DNA fragments approximately 100-3,000 base pairs (bp) in length that are located between adjacent microsatellite regions. Single primer amplification reactions (SPAR) use a single primer for PCR, designed to target microsatellite motifs. No information about sequence variation is necessary since these primers anneal directly to the microsatellite motifs and amplify multiple microsatellite loci with a single primer, producing variable fragment patterns following electrophoresis [44, 46].

Amplified fragment length polymorphisms (AFLP)

AFLP is a PCR-based technique first reported by Vos et al. [47]. It involves cutting of genomic DNA with restriction enzymes, similar to RFLP, followed by ligation of

adapter sequences to the fragments. PCR is then performed, with primers targeting three locations: the adapters, restriction enzyme recognition sequence, and a few nucleotides of the intended restriction fragment. Following electrophoresis, the presence and absence of bands is noted for genotyping. From a forensic standpoint, AFLP is problematic because the polymorphisms are not randomly distributed and are primarily dominant, resulting in bands that are not necessarily independent of one another [48]. Additionally, the technique requires high quality DNA, which is not always available in forensic samples.

Short tandem repeats (STRs)

STRs are repeating units of 1-6 bp that have a high level of polymorphism and high inter-laboratory reproducibility [49]. They are codominant and capable of being multiplexed, resulting in a high power of discrimination. Since STR analysis targets specific loci, previous knowledge of the genome is required. STRs are considered the gold standard for human identification, and many studies have been conducted to identify informative STR markers in plants of forensic interest.

DNA barcoding

DNA barcoding involves analyzing the sequence of a standardized portion of the genome in order to identify species. Sequences at barcode regions uniquely identify a species in the same way that UPC barcodes at the supermarket are unique identifiers for a particular item, and species identification is accomplished by comparing the sequence to a database. If individuals from different populations of a species exhibit sequence differences at these loci, DNA barcoding techniques can also be used for geographic origin assignment of individuals, as long as an appropriate database of the relevant populations exists. Typically, mtDNA, cpDNA, and ribosomal DNA (rDNA; consisting of nuclear genes

encoding ribosomes) are targeted for this technique. An ideal barcode marker must be variable enough to be informative (unique to a species) but conserved enough for the development of universal primers [50]. In order to be cost-effective, it also must be short enough to sequence in a single reaction.

Species identification

Many plants and fragments of plant material cannot be identified at the species level by morphology alone. DNA barcoding presents a method for identifying the species of unknown plant matter that may be valuable in forensic investigations. In animals, the mitochondrial genome is preferred for barcoding studies due to its small size, high copy number, conserved genetic structure, and rapid evolution. The cytochrome c oxidase subunit 1 (CO1) gene is used almost universally for species identification in animals [51]. In plants, the mitochondrial genome is much larger and evolves much more slowly due to low substitution rates [52]. Due to the low substitution rate, no universal barcode matching the utility of CO1 has been discovered in plants. The plant mitochondrial genome is also subject to intramolecular recombination and does not have a consistent gene order. The chloroplast genome, therefore, is preferred for barcoding studies in plants due to its conserved gene order and lack of heteroplasmy and recombination. The most commonly used barcoding strategies involve a combination of the chloroplast *rbcL*, *matK*, and *trnH-psbA* regions and the ribosomal internal transcribed spacer (ITS) region [37, 52-54].

An example of the usefulness of DNA barcoding can be found in the world of forensic palynology, the study of pollen as it relates to criminal activities. Since pollen is ubiquitous in the environment and pollen signatures vary between environments, pollen

can be used to link a person to a location and, in some cases, a particular time of year by identifying pollen species using DNA barcoding [55].

Origin determination

Often, species level identification of a plant is not enough. Instead, population-level identification may be necessary, such as in cases of illegal logging, where it must be proven that a log was cut from a tree in a particular region or country. Origin determination of illegal drugs is also helpful for law enforcement agencies that investigate drug trafficking. It has been shown that crops planted in different parts of the world, genetically isolated from each other, develop region-specific differences in organelle barcoding regions, which are passed to the offspring within each population. This results in sequence differences in barcoding regions, and analyzing multiple barcoding regions can produce haplotypes that are characteristic of a certain region or population.

In humans, the D-loop of the mitochondrial genome, containing hypervariable regions 1, 2, and 3 (HV1, HV2, and HV3), is the target of lineage studies. Analysis of the HV regions of an individual results in a haplotype, and ancestry can be determined by comparing this haplotype to a database [56]. Construction of a similar genetic database, consisting of organelle haplotypes corresponding to lineages with a defined geographic range, for any species with an appropriate genetic structure would allow for biogeographical origin determination of individuals.

The use of DNA barcoding in determining the biogeographical origin of *Cannabis sativa* samples has been explored, with implications in drug trafficking investigations. Gilmore et al. [36] found differing organelle haplotypes from *C. sativa* cultivars from several different countries. Since hemp and marijuana cultivars are genetically isolated

from each other, barcoding haplotypes may also be capable of distinguishing marijuana from hemp, which was supported by Gilmore et al.'s study. However, an extensive database has yet to be created, and there is no agreement on the most informative genomic regions to use for this purpose in plants.

Massively parallel sequencing (MPS)

Sanger sequencing (pre-MPS)

The first generation of sequencing began with Sanger et al.'s method [57] published in 1977. Sanger sequencing, as the method and its variations came to be known, is a method based on the incorporation of chain-terminators during PCR. In current Sanger sequencing techniques, each PCR reaction contains the commonly-used deoxynucleotide triphosphates (dNTPs) as well as dideoxynucleotide triphosphates (ddNTPs), which lack the 3' hydroxyl group necessary for extension. Each ddNTP is labeled with a different fluorescent dye to indicate the nucleotide base (adenine, thymine, cytosine, or guanine). During PCR, incorporation of a dNTP allows for extension of the fragment; however, when a ddNTP is randomly incorporated, no additional nucleotides can be added, terminating extension of that fragment. This results in fragments of varying lengths, each labeled with a fluorescent chain terminator. Separating these fragments by size via capillary electrophoresis (CE) results in a base-by-base sequence of the template DNA.

Overview & advantages of MPS

MPS, also known as next generation sequencing, allows users to sequence hundreds of targets in hundreds of samples simultaneously [58, 59]. Forensic DNA is moving toward MPS for simultaneous analysis of human STR markers and SNPs for identity, ancestry, and phenotype prediction. In addition to allowing analysis of more markers than traditional

CE-based methods, MPS provides more information about STR alleles by allowing the detection of sequence-variants, or isoalleles. For example, a tetranucleotide STR allele with the sequence $(ATTC)_7$ and an allele with the sequence $(ATTC)_4(ATTT)(ATTC)_2$ would look identical on CE but can be distinguished by their sequences using MPS, providing a greater power of discrimination. There are two major MPS chemistries used in forensic science: semi-conductor (Thermo Fisher Scientific's Ion Torrent™) and reversible dye terminator (Verogen MiSeq FGx®).

Sequencing platforms/chemistries

Semi-conductor

During extension, when a nucleotide is added, cleavage of the pyrophosphate results in the release of a proton, which causes a pH change. Semi-conductor sequencing instruments, such as the Ion S5 System from Thermo Fisher Scientific, work by detecting changes in voltage due to these pH changes and translate that signal to a base call [60].

The first step in semi-conductor sequencing is library preparation, which results in DNA fragments (targeted or random) containing adapter sequences and barcodes. The adapter sequences are complementary to those found on the Ion Sphere™ particle and facilitate clonal amplification. Barcodes are unique sequences ligated to the ends of all template fragments from a given sample, and they allow sequencing results to be associated with that particular sample. Following clonal amplification by emulsion PCR, the Ion Sphere™ particles are loaded onto an ion chip, which has millions of wells into which single particles fit. The sequencer then floods the entire chip with a particular nucleotide; if that nucleotide is incorporated in a fragment, a pH change is detected in that well and is recorded as a base call. Then a different nucleotide is washed across the chip and the

process is repeated. If multiple of the same nucleotides are incorporated during a step, the voltage change corresponds to the number of nucleotides, allowing the sequencer to analyze homopolymeric stretches.

Reversible terminator

The Verogen MiSeq FGx™ is a reversible terminator sequencing instrument commonly used in forensics. It uses ddNTPs with different fluorescent labels, similar to Sanger sequencing, but they contain a removable blocking group [61]. This allows the instrument to record the fluorescence of a single base incorporation, followed by removal of the blocking group and incorporation of the next nucleotide.

The process begins with library preparation, during which template DNA is prepared by targeted amplification or random fragmentation, followed by ligation of adapters and indices (similar to the barcodes used in semiconductor sequencing). The adapters on the templates hybridize to complementary sequences on the flow cell, and clusters are generated via bridge amplification. During each sequencing cycle, fluorescently tagged nucleotides are flowed across the cell, resulting in one being incorporated to the growing strand and a corresponding fluorescent signature being recorded by the instrument. The dye and terminator group are then cleaved, and the next sequencing cycle begins.

Cannabis sativa

Botany and Taxonomy

Cannabis sativa Linnaeus (*C. sativa* L.) is an annual, herbaceous angiosperm (flowering plant). It is dioecious, meaning that there are distinct male and female plants [62]. Male plants are taller, shorter-lived, and have stamen, which are responsible for pollen

production [62]. Female plants have pistils, which contain eggs, and are the preferred sex for marijuana (drug-type cannabis) growers due to the high concentration of cannabinoids in the buds [63]. The phloem, or bast, of the plants are used for fiber production, while the leaves and buds are selected for drug use [63, 64]. In addition to the drug marijuana, which refers to dried flowers and leaves of the plant, hashish (hash) and hash oil are derived from *C. sativa*. Hash is produced by collecting the dried resin and resin glands (trichomes), and hash oil is a distilled form of the resin. Typically, both marijuana and hash are inhaled through smoking or vaping, but edibles are also increasing in use. A few drops of hash oil may be applied to a cigarette or joint or used for vaping. Additionally, hemp seeds and seed oil are sold as dietary supplements rich in fatty acids, protein, and magnesium.

The taxonomic classification of *C. sativa* has been widely debated [65]. Currently, a monotypic classification is preferred, with *sativa* recognized as the only species in the genus *Cannabis*. Historically, however, some experts divided the genus into three separate species: *sativa, indica,* and *ruderalis*. Cannabis was first classified by Carl Linnaeus in 1753 as a single species, *C. sativa*. Later, in 1785, Jean-Baptiste Lamarck classified an Indian variety of cannabis as a separate species, *C. indica*, on the basis of several morphological characteristics that distinguished the plant from *C. sativa*. Finally, a Russian variety of cannabis was found to have differences in size, shape, and seed morphology and was deemed to be a third species, *C. ruderalis*.

Originally, the Cannabaceae family contained only *Cannabis* and *Humulus* genera. However, it has been expanded and now contains ten genera [66]. *Humulus lupulus* (hops) is the species most closely related to *C. sativa*. Table 1.1 shows the taxonomic classification of *C. sativa* [67].

Table 1.1 Taxonomic classification of *C. sativa*

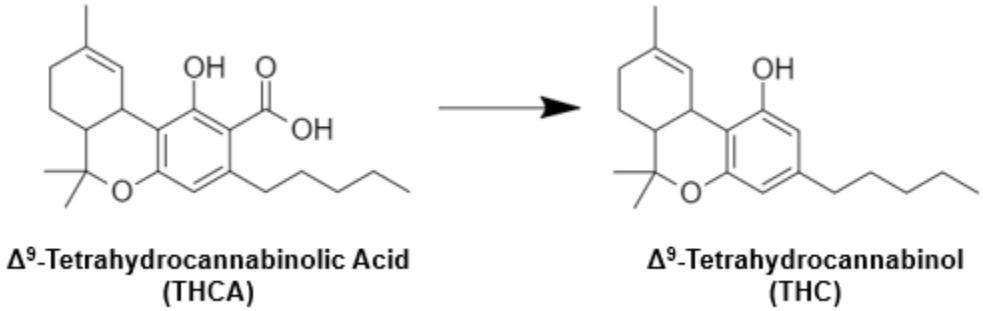
Domain	Eukaryota
Kingdom	Plantae
Subkingdom	Viridiplantae (green plants)
Infrakingdom	Streptophyta (land plants)
Superdivision	Embryophyta
Division	Tracheophyta (vascular plants)
Subdivision	Spermatophytina (seed plants)
Class	Magnoliopsida
Superorder	Rosanae
Order	Rosales
Family	Cannabaceae
Genus	Cannabis
Species	*Cannabis sativa* L.

Chemistry

Cannabinoids

C. sativa contains over 400 chemical compounds, including more than 60 cannabinoids, which act on cannabinoid receptors to produce effects in the human body. The cannabinoid most associated with psychoactive properties in marijuana is delta-9-tetrahydrocannabinol (THC). It is present in its acidic form, delta-9-tetrahydrocannabinolic acid (THCA), in the flowering tops (buds), leaves, and resin of the plant. THCA is converted to THC via non-enzymatic decarboxylation through heating or drying of the plant material. The structure of THC was reported in 1964 [68] (Fig. 1.1). THC content (measured as a percentage of dry weight) varies widely between sources, ranging from less than 1% to 37.2% [69], and the average THC content has risen the last two decades [70]. Whereas the average THC concentration in marijuana prior to 1990 was 2%, now the average THC content is closer to 15%, and some strains available in marijuana dispensaries

boast concentrations as high as 28% [71]. Cannabidiol (CBD) is the cannabinoid highly abundant in fiber-type, or hemp, plants. Other cannabinoids present in *C. sativa* include cannabinol (CBN), cannabigerol (CBG), and cannabichromene (CBC).



Δ⁹-Tetrahydrocannabinolic Acid (THCA) **Δ⁹-Tetrahydrocannabinol (THC)**

Fig. 1.1 Chemical structures of THCA and THC. Decarboxylation of THCA occurs during heating, producing the psychoactive chemical THC

Cannabinoids act on endogenous cannabinoid receptors (CB_1 and CB_2), which are abundant in the striatum, cerebral cortex, and hippocampus of the brain and in the peripheral nervous system, with CB_2 being mainly associated with immune cells [11, 72, 73]. THC binds to these cannabinoid receptors, disrupting neural communication and causing affects such as euphoria, relaxation, altered time perception and mood, increased appetite, impaired concentration and learning, paranoia, and, at high doses, hallucinations. Researchers have also shown that THC has activity on dopamine receptors, increasing dopamine release from the nucleus accumbens and prefrontal cortex [74], which is common for many drugs of abuse.

Chemotypes

Five chemical phenotypes, or chemotypes, exist based on the relative amounts of the major cannabinoids [75, 76]. Chemotype I plants, considered ‘drug type,’ have a low CBD/THC ratio; chemotype II plants, considered ‘intermediate,’ have a roughly equal

CBD/THC ratio; chemotype III plants, considered ‘fiber type,’ display a high CBD/THC ratio; chemotype IV plants are high in CBG; and chemotype V plants lack any detectable cannabinoid content.

Cultivation

C. sativa is cultivated around the world for its use as a hallucinogen, medicine, and industrial fiber. It has a long history of human use and is purported to have originated in central Asia [77]. Archaeological findings indicate that the plant was cultivated in China as early as 4,000 B.C. to make ropes, textiles, and paper [78]. Early Chinese medicine indicated the use of cannabis as an analgesic and anti-inflammatory drug for rheumatic pain, a laxative to relieve constipation, and as a surgical anesthetic. In India, cannabis was used as a medicine and in religious rituals beginning around 1,000 B.C. Tibetan culture also used cannabis to facilitate meditation.

Cultivation of the plant in America dates back to the early colonial period, when it was used to make textiles and rope. In the 1830s, Irish physician Sir William Brooke O’Shaughnessy helped to popularize medicinal cannabis in Europe, following his study of the drug in India. He found that it could be used to treat rheumatic diseases and convulsive disorders and reduce stomach pain and vomiting in cholera patients [79]. By the 1850s, cannabis preparations were found in American pharmacies and were considered mainstream.

Legal status

The Controlled Substances Act (CSA)

The CSA, originally passed in 1970, regulates the possession and use of certain substances and classifies them as belonging to one of five “schedules” based on their

potential for abuse, safety, and current medical uses. Drugs classified as Schedule I by the Food and Drug Administration (FDA) and Drug Enforcement Administration (DEA) have a high potential for abuse and no accepted medical uses under federal law. Use or possession of these drugs is punishable under federal law. Currently, marijuana, and the cannabinoids derived from it are listed under Schedule I [80]. However, dronabinol, a synthetic form of THC, has been approved for medical use in the drugs Marinol (Schedule III) and Syndros (Schedule II). Cesamet (nabilone) contains another synthetic cannabinoid and is currently schedule II.

2018 Farm Bill

The Agriculture Improvement Act of 2018 (colloquially known as the Farm Bill) modified the definition of marijuana in the CSA, making hemp legal at the federal level [81, 82]. Hemp is defined as cannabis and its derivatives with no more than 0.3% THC. The production of hemp, however, remains under tight regulation by the U.S. Department of Agriculture (USDA), and hemp products are subject to regulation by the FDA.

State laws

Though marijuana is illegal at the federal level, at the time of this writing, 33 states and the District of Columbia (D.C.) have passed laws legalizing medicinal or recreational marijuana (Fig. 1.2). The first state to legalize medicinal marijuana was California, in 1996. Eleven states – Alaska, California, Colorado, Illinois, Maine, Massachusetts, Michigan, Nevada, Oregon, Vermont, and Washington – and D.C. allow recreational use of the drug for adults age 21 and over.

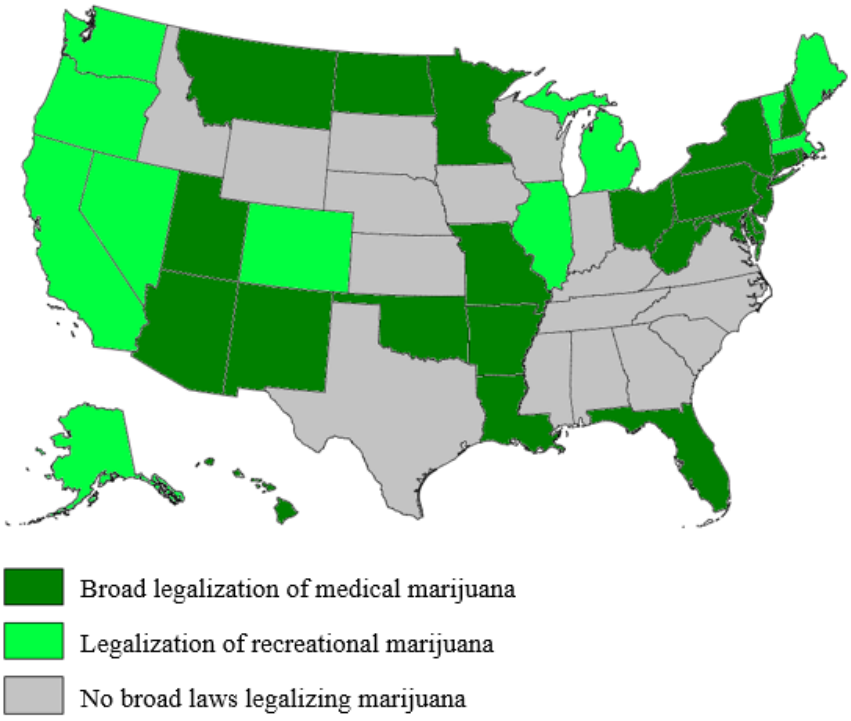


Fig. 1.2 Map showing state marijuana laws as of March 2020

Cannabis research

The CSA and the 1961 International Single Convention on Narcotic Drugs give the federal government sole power in authorizing marijuana growing and distribution [83]. Currently, the DEA has authorized only one grower, the University of Mississippi [84]. The National Institute on Drug Abuse (NIDA) contracts with the University of Mississippi to grow marijuana for research studies.

Forensic identification

Forensic identification of *C. sativa* involves microscopic observation of cystolithic hairs on plant material, a positive (blue-violet) Duquenois-Levine color test, thin layer chromatography (TLC), and identification of THC by gas chromatography-mass spectrometry (GC-MS) [85]. It is necessary to confirm the presence of THC using GCMS

because of the possibility of false positives with other testing methods. For example, over 80 plant species contain cystolithic hairs similar in morphology to *C. sativa*, making false identification of marijuana possible by analysts without an advanced degree in botany [86]. The Duquenois-Levine color test [87] also has the potential for false positives due to other plant substances and the subjective nature of color test interpretation [88, 89].

With the passing of the 2018 Farm Bill, which legalized the possession of hemp, identification of a substance as *C. sativa* is no longer enough for legal purposes. A validated method that quantifies THC in a sample must be used to determine whether that sample is marijuana (>0.3% THC content). This test is not currently common practice in most crime laboratories, which have typically only tested for the presence of THC. Similarly, microscopic, Duquenois-Levine, and TLC testing cannot distinguish between marijuana and hemp.

Additionally, difficult sample types (such as trace amounts, burned samples, young plants, and seeds) complicate identification of marijuana. The THC content of *C. sativa* typically increases as the plant matures [90], making testing on juvenile plants inconclusive as to whether they are marijuana or hemp. Sample types such as seeds, pollen, and root material have limited amounts of THC, lack cystolithic hairs, and do not yield reliable results for determining marijuana or hemp. Small amounts of material may not be sufficient for chemical analysis. For these reasons, sensitive and specific DNA tests to identify *C. sativa* are necessary, and genetic markers for distinguishing between hemp and marijuana have been extensively studied [36, 91-93]. Additionally, genetic testing opens the possibility of providing investigative leads for law enforcement based on the

biogeographical origin of samples and suggesting case linkage due to genetic relatedness of samples.

Genetic studies

Genome

C. sativa is a diploid organism (2n=20) with nine autosome pairs and one sex chromosome pair [94]. Similar to humans, sex is determined by X and Y chromosomes, with females having two X chromosomes and males having one X and one Y chromosome. Diploid female and male plant genome sizes are 1,636 and 1,683 Mbp, respectively [94]. The nuclear genome contains 30,000 genes and has been fully sequenced and mapped for the cultivar Purple Kush (GenBank accession AGQN00000000) [95]. Additionally, the draft chloroplast genome has been reported for several fiber-type and drug-type varieties [96-98], and the draft mitochondrial genome has been reported for hemp [99]. The chloroplast genome is 153,871 bp, AT-rich (63%), and contains 83 genes [97]. The mitochondrial genome is 415,499 bp and contains 54 genes [99].

Species identification

As previously discussed, DNA barcoding techniques have the ability to identify species. To this end, several barcoding markers have been examined in *C. sativa*, and researchers have found that the chloroplast intergenic spacer trnL-trnF and the nuclear internal transcribed spacer regions (ITS1 and ITS2) are effective for identifying *C. sativa* and distinguishing it from its close relatives, including *H. lupulus* [100-107]. Markers with sequence variation among *C. sativa* accessions may also be informative for distinguishing between populations of *C. sativa* or determining crop type (marijuana or hemp).

Determination of crop type and biogeographical origin

Several genetic methods have been proposed to distinguish between drug-type (marijuana) and fiber-type (hemp) crops and determine the biogeographical origin, including analysis of nuclear and organelle polymorphisms.

Nuclear DNA polymorphisms

When analyzing nuclear DNA polymorphisms, allele frequencies are expected to vary between populations. Therefore, allelic variation could indicate a particular biogeographical origin for *C. sativa* samples. Since marijuana and hemp are cultivated separately (in order to maintain their respective chemotypes for their intended purposes), it is expected that allele frequencies between the two crop types would differ as well. Phylogenetic analysis of hemp and marijuana samples genotyped with a 13-plex STR assay support this hypothesis, showing a close genetic relatedness between marijuana samples from different countries and a differentiation from hemp based on autosomal genotypes [108]. Additionally, an STR genotype database for Australian cannabis seizures identified fourteen alleles that were unique to drug-type cannabis and thirty that were unique to fiber-type cannabis [109]. Gilmore et al. [93] genotyped 98 individuals using five polymorphic STR markers and concluded that drug-type cultivars had less genetic variability than fiber-type and that with a DNA database, STR genotypes could be used to provide information about the geographic origin of samples.

Similarly, AFLP analysis was used to analyze polymorphisms in the autosomal DNA of marijuana and hemp [110-112]. Using ten primer pairs and producing 1,206 bands per sample, researchers found eighteen bands representing fixed differences between fiber and drug cultivars [110]. The study was limited, however, in that it used only three varieties

of hemp and a single drug-type variety. Kojoma et al. [111] also investigated the use of ISSR markers and was able to differentiate a single drug-type sample from two hemp cultivars, though many more samples are needed to make any conclusions about the ability of these markers to differentiate crop types. Hakki et al. used ISSR markers to discriminate between Turkish marijuana and hemp [112]. Another promising set of markers was reported by Pinarkara et al. [113], which used RAPD analysis to separate marijuana samples into two groups by their geographic origin: eastern or western Turkey. RAPD markers were also used in studies by Jagadish et al., Faeti et al., Shirota et al., and Forapani et al. [114-117], which all found differences in genotypes in samples from different geographic origins, though in all three studies, the sampling area and number of samples were limited.

DNA barcoding

Kohjyouma et al. [106] identified two sequences (differing by a one bp deletion) in the trnL-trnF intergenic spacer region in hemp and predicted that *C. sativa* had differentiated into many local variations. Mello et al. [118] compared a limited number of cultivars from Brazil to database sequences from China, the United States, and the United Kingdom and found two SNPs in the rbcL gene which could be informative for determining biogeographical origin. Gigliano and Caputo [102] identified three SNPs in the ITS2 region between different cultivars. Alone, none of these markers appeared to be sufficient for distinguishing between populations of *C. sativa*, but they point to the possibility that organelle genome polymorphisms could provide important population information.

Based on the prediction that genetically isolated populations will produce different organelle haplotypes, seven organelle DNA polymorphisms (five cpDNA and two

mtDNA) have been evaluated for their use in distinguishing between marijuana and hemp [36, 108]. Gilmore et al. found that in comparing haplotypes of wild, fiber, and drug type *C. sativa* samples, haplotype frequencies differed between crops with different uses and from different geographic regions [36]. Houston et al. genotyped hemp and marijuana samples from North and South America using the same seven organelle loci and found that the hemp samples produced a unique haplotype that differentiated them completely from marijuana samples [108]. However, the majority of cultivars could not be differentiated from each other, and a larger set of markers is likely required to identify differences between *C. sativa* grown in countries throughout the world.

In addition to STR, RAPD, ISSR, AFLP, and organelle markers, a large number of studies have sought to differentiate between marijuana and hemp using genes coding for enzymes responsible for cannabinoid production.

Cannabinolic acid synthases

The genetic inheritance of chemotype has been extensively studied, with genes controlling the THCA and cannabidiolic acid (CBDA) synthase enzymes at the forefront. These enzymes are involved in the final step in the biosynthesis of the major cannabinolic acids. A common precursor, cannabigerolic acid (CBGA), is the target of both THCA synthase and CBDA synthase enzymes, resulting in the production of THCA or CBDA, respectively [119].

Two chemotype inheritance models are described in the literature. The first was proposed by de Meijer et al. [75], which reported based on inheritance patterns that chemotype is influenced by a single gene at locus B with two codominant alleles, B_D and B_T. The allele B_D codes for CBDA synthase, while the allele B_T codes for THCA synthase.

In this model, genotypes at the B locus correspond to three chemical chemotypes: chemotype I (B_T/B_T), chemotype II (B_D/B_T), chemotype III (B_D/B_D) [75]. This model was expanded by Pacifico et al. [76], which proposed the addition of two additional chemotypes: chemotype IV with high CBG due to the presence of a new allele at the B locus (B_0) and chemotype V with a lack of detectable cannabinoids, which may be caused by a gene upstream of the B locus preventing formation of CBG.

A second genetic model describes two separate but tightly linked genes for THCA and CBDA synthase. In this model, differences in expression or allelic variation at these loci affects enzymatic activity. Since THCA and CBDA synthase compete for a substrate (CBGA), the genotypes at both loci influence cannabinoid content. This model is supported by several studies [92, 95, 120-127]. Since the two loci are linked (inherited together), inheritance patterns are identical to single locus models [127].

Kojoma et al. [92] reported the sequences of two varieties of the THCA synthase gene. An active form of the gene was found only in drug-type plants, while an inactive form was present in hemp. The isoforms varied by 62 base substitutions, or single nucleotide polymorphisms (SNPs), which resulted in 37 amino acid substitutions. Based on these sequences, several research groups designed assays to differentiate between drug and non-drug samples [120-122]. Other researchers reported additional polymorphisms in the THCA synthase gene, leading to less active isoforms of the enzyme [123, 125].

However, two recent studies have suggested that no “inactive” (or less active) THCA synthase gene exists [124, 128]. Rather, Weiblen et al. [124] suggested that chemotype was due to a less functional CBDA synthase gene expressed in marijuana (chemotype I plants), resulting in less competition for the substrate CBGA. The THCA

synthase-like sequence reported by Kojoma et al. [92] in hemp was later identified as the cannabichromene acid (CBCA) synthase gene [128].

In addition to allelic variation at the THCA and CBDA synthase genes, copy number variation or the influence of other genes have also been proposed as potential factors affecting chemotype [124, 128-131].

Individualization

Individualization of marijuana samples can provide investigative leads by linking related cases and providing associations between distributors and growers. Marijuana may be bred by pollination, resulting in seeds that produce a new, genetically distinct plant, or propagated clonally, involving cutting off a portion of the plant and placing it into soil to grow into a new plant. Clonal propagation results in a new plant that is genetically identical to the parent plant, and this method is often preferred by breeders, as it preserves desirable qualities of the parent plant such as cannabinoid content (or fiber quality in hemp). Case linkage with STR data may be accomplished with either clonally propagated plants (which will have identical STR profiles) or pollinated plants (which will be genetically related), providing important leads for law enforcement agencies.

RAPD

RAPD analysis of *C. sativa* was first reported by Gillan et al. in 1995, which concluded that RAPD analysis provided better individualization potential compared to high performance liquid chromatography (HPLC) [132]. Jagadish et al., Faeti et al., Forapani et al., Pinarkara et al., and Kayis et al. further showed that RAPD analysis could be used to group samples by cultivar or geographic origin [113-115, 117, 133]. Shirota et al. used

RAPD analysis in combination with RFLP to distinguish between samples with different chemotypes (drug, fiber, and intermediate).

ISSR

ISSR analysis of *C. sativa* for individualization and distinguishing between hemp and marijuana cultivars has been reported by several studies [111, 112, 133]. Kojoma et al. found that ISSR provided better discrimination of hemp samples compared to HPLC [111], and Kayis et al. reported that ISSR analysis had slightly better resolving power than RAPD for individualization of marijuana [133]. Hakki et al. was able to discriminate between marijuana and hemp from Turkey using principal coordinate analysis (PCoA) [112].

AFLP

Another genetic technique for individualization and genetic diversity studies of *C. sativa* is AFLP [134, 135]. Additionally, Datwyler et al. used AFLP markers to identify the geographic sources of *C. sativa* and discriminate between crop types [110]. AFLP analysis was also useful for sex determination of cannabis plants [136, 137].

STRs

STR analysis is the most common method for human identification in forensic laboratories, and STRs are found in many organisms, including plants [138]. This type of marker has desirable characteristics such as being highly variable between individuals, multiallelic, codominant, abundant and distributed throughout the genome, reproducible, and amenable to high throughput genotyping and automation. As such, multiple studies have sought to identify and evaluate STR markers in *C. sativa* for individualization.

The first *C. sativa* STR markers were reported in 2003 by Gilmore and Peakall (fifteen loci) [139], Alghanim and Almirall (eleven loci) [140], and Hsieh (one highly

polymorphic hexanucleotide locus) [141]. Additionally, seven loci were proposed by Valverde et al. [142], which also sought to standardize the allele nomenclature for all *C. sativa* loci. Gao et al. [143] also reported the existence of over 3,400 expressed sequence tag (EST)-derived SSR markers, which are expressed in mRNA transcripts and associated with genes, and developed 56 loci. The utility and diversity of these STR markers has been evaluated by subsequent studies [93, 144-146], and several groups created STR multiplexes to analyze multiple loci simultaneously, giving a high power of discrimination for forensic analyses [147-151].

Locating highly variable STR loci in marijuana is only the first step. Before these markers can be meaningfully used in forensic applications, an extensive and representative population database must be created, followed by testing for independence of loci and other population genetic statistics, including allele frequencies, heterozygosity, and conformance with Hardy-Weinberg expectations. Several studies have acknowledged the need for these databases and begun compiling *C. sativa* genotypes for a variety of the previously reported loci [108, 109, 152].

One study by Houston et al. created an MPS assay to simultaneously sequence twelve STR loci in a high-throughput manner [153], which resulted in additional variation compared to traditional capillary electrophoresis-based genotyping.

Papaver somniferum

Taxonomy

Papaver somniferum, commonly known as the opium poppy, is classified according to Table 1.2 [67]. The term “poppy” refers to all approximately 800 species within the

family *Papaveraceae*, and these plants are characterized by the production of a watery or milky latex by laticifers (secretory cells found in some plants).

Table 1.2 Taxonomic classification of *P. somniferum*

Domain	Eukaryota
Kingdom	Plantae
Subkingdom	Viridiplantae (green plants)
Infrakingdom	Streptophyta (land plants)
Superdivision	Embryophyta
Division	Tracheophyta (vascular plants)
Subdivision	Spermatophytina (seed plants)
Class	Magnoliopsida
Superorder	Ranunculanae
Order	Ranunculales
Family	Papaveraceae (poppies)
Genus	Papaver
Species	*Papaver somniferum* L.

Chemistry

Opiates and opioids

As its nickname suggests, the opium poppy is the source of opium, a dried latex containing chemical compounds called opiates. The most abundant opiates are morphine, codeine, thebaine, papaverine, and noscapine [154]. Morphine and codeine are prized for their analgesic effects and are commonly prescribed for management of moderate to severe pain. Other analgesic medications, such as hydrocodone and oxycodone, are derived from thebaine [155]. Morphine was first isolated from the poppy plant in 1803 by Friedrich Serturner, a German chemist (Fig. 1.3) [156]. Unfortunately, opiates are also one of the most highly abused drug classes worldwide.

The term opiate refers to the naturally-occurring alkaloids found in the latex of the *P. somniferum* plant [14, 157]. The term opioid (or narcotic) is broader and refers to all

compounds or drugs with opiate-like pharmacological effects, including semi- and fully-synthetic opioids like hydrocodone and fentanyl. These drugs create their effects by acting on opioid receptors in the central and peripheral nervous systems. There are three major subtypes of opioid receptors: delta, kappa, and mu [158, 159]. The most potent analgesic effects are caused by opiates that bind to the mu receptors, but side effects include euphoria, constipation, dependence (psychological and physical), and respiratory depression, which results in death when large amounts of the drug are taken. Mu antagonists, such as naloxone (Narcan), can reverse respiratory depression quickly in cases of morphine overdose [160]. Opioids that act on the kappa and delta receptors are considered to be safer than drugs which act on mu receptors.

Routes of administration for licit and illicit opioids varies, with oral administration (swallowing) of pills being the most common [161]. Hydrocodone, methadone, and oxycodone are commonly snorted, and morphine and hydromorphone are commonly injected.

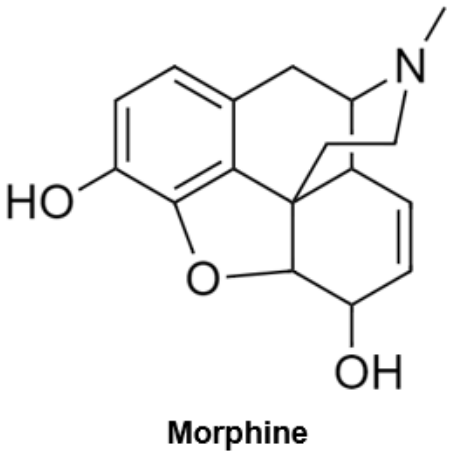


Fig. 1.3 Chemical structure of morphine

Heroin

Heroin is a semi-synthetic opioid derived from morphine. It acts on the mu opioid receptors and is twice as potent as morphine [162]. It has no accepted medical uses in the United States and a high abuse potential, earning it a Schedule I classification by the Controlled Substances Act. Additionally, as an illicitly produced substance, heroin is not regulated; therefore, samples from different producers have varying purity and additives. Currently, the most common additive is fentanyl, a synthetic opioid [163]. Fentanyl and its analogs are hundreds to thousands of times more potent than morphine, making heroin samples containing these additives even more potent and dangerous and leading to accidental overdoses [164]. Heroin comes in three main forms: white powder, brown powder, and black tar heroin [165, 166]. White powder mixed with fentanyl, an especially deadly combination, is known as China White. White powder heroin is the most pure form of the drug, though drug dealers mix it with cutting agents to maximize supply and profits. It is commonly injected or snorted. Black tar heroin may be found in gooey or rock-hard consistencies, and the dark color comes from contaminants from the refinement process. This form is usually smoked or injected but may also be ground down and snorted. Brown powder heroin is created by crushing black tar heroin and cutting it with other additives, and it can be smoked or snorted. Popularity and availability of the heroin types varies by location [165]. East of the Mississippi River, most heroin is powdered and sourced from Colombia. To the west, much of the heroin is black tar sourced from Mexico.

Poppy seed tea

Seeds from the opium poppy are legal, easy to obtain, and are a source of opiates due to their contact with the opiate-rich latex inside the opium poppy pod. Consumption of

poppy seeds can lead to positive drug tests for opiates and may be detected in urine for up to 48 hours [167-169]. However, consuming poppy seeds in food items does not normally produce a “high” due to the low content of opiate alkaloids. Poppy seeds bought at the grocery store are typically washed prior to sale, reducing the opiate content of the seeds, and heating during the baking process causes further loss of the opiates [170]. However, a “home-brewed high” obtained from tea brewed from unwashed poppy seeds has become popularized by internet culture in recent years [171, 172]. Poppy seed tea, or opium tea, is brewed by combining a substantial amount of poppy seeds (100 g to 1.4 kg) with water, either at room temperature or heated, allowing the extraction of the opiates from the surface of the seeds into water. Additives such as lemon juice or alcoholic beverages may also be called for in these recipes to mask the bitter taste.

The onset of effects from the tea occurs quickly, within 15-30 min, and these effects have been reported to last for 8-24 hours, depending on the alkaloid content of the seeds [171, 173-175]. Opiates in the tea have the potential to be abused, similar to other opioid drugs [176], and individuals may turn to poppy seed tea to reduce their dependence on more potent drugs, such as heroin [177]. Deaths from accidental overdoses of poppy seed tea have made the news on several occasions [178-181], and two overdose cases (one fatal, one nearly fatal) have been published in the scientific literature [173, 175]. In reaction to the deaths, the U.S. Department of Justice has issued a warning about poppy seed tea [171]. Accidental overdoses from the tea occur due to a high variation in the alkaloid content of seeds purchased from various sources [182]. As a response to the growing trend and dangers associated with poppy seed tea, the DEA released a statement in 2019 that unwashed poppy seeds violate the CSA due to the presence of opium alkaloids [183].

Cultivation

The opium poppy most likely evolved from a wild Asian species and was native to the Mediterranean [184]. Now, it is cultivated around the world in subtropical and temperate zones for its use as an analgesic (painkiller) and for its seeds and seed oil, which are popularly used in baking [185]. *P. somniferum*, characterized by bright, colorful flowers (red, pink, purple, or white), is also used for ornamental purposes and can be kept as a houseplant. The largest producers of opium poppy are India, Burma (Myanmar), Afghanistan, Turkey, and the former Soviet Union, which together produce more than two-thirds of the world's opium poppy. India alone supplies half of the opium used by pharmaceutical industries around the world [14, 157].

A milky latex is created inside the poppy pods by laticifers, specialized secretory cells associated with the phloem (lymph tissue) [14]. It is collected through manual lancing of the unripe poppy pods, resulting in a white-brown exudate with a complex chemical composition including morphine, codeine, thebaine, narcotine, and papaverine. If collected too early, morphine content is low, and if harvested too late, a large percentage of morphine in the latex biodegrades to codeine [14].

History

P. somniferum has been used for thousands of years, with its earliest documented use occurring in ancient Mesopotamia around 3,400 B.C. [186, 187]. The Sumerians called it *hul gil*, or "the joy plant," and its use as a sleep aid and pain reliever has also been documented by the ancient Greeks, Persians, and Egyptians. Use of opium spread throughout Europe and Asia by 600 A.D. via trade along the Silk Road [186]. Paracelsus developed laudanum by dissolving opium in alcohol, which helped to mask the bitter taste

of raw opium. Laudanum quickly took root in Europe and became the basis for numerous medications [187].

Following its conquest of India in the 1700s, Britain's East India Trading Company began to smuggle opium into China, a lucrative business. As a result, opium addiction became widespread in China, and the Qing Dynasty outlawed the import of opium. The first Opium War began in 1839 when Britain used "gunboat diplomacy" to force China to open its ports to trade [186, 187]. The conflict ended in 1842 with the Treaty of Nanking, which gave Britain control of Hong Kong. The Second Opium War began in 1856 with Britain and France allied against China to force legalization of opium trade. It ended in 1860 with another European victory. During the Gold Rush of 1849, Chinese immigrants brought the habit of opium smoking to America and established opium dens [186, 187].

Isolation of morphine from opium was accomplished in 1803 by Friedrich Serturner [156]. The German chemist named the substance after Morpheus, the Greek god of dreams. Pure morphine is ten times as potent as opium and was used as a painkiller during the U.S. Civil War. Liberal use of the drug resulted in hundreds of thousands of soldiers becoming addicted, causing scientists to search for a less addictive drug similar to morphine [186]. In 1874, Alder Wright, an English chemist, synthesized heroin, intended to be a safer alternative, from morphine. It was marketed as a cough suppressant by the pharmaceutical company Bayer. Unfortunately, heroin has an even higher addiction potential than morphine, and heroin addiction skyrocketed in the U.S. and Western Europe. The first major legislation controlling sale and use of opiates in the U.S. came in 1914 with the Harrison Narcotics Tax Act [186].

Legal status

Under the Controlled Substances Act, opioids (or narcotics) vary from Schedule I to Schedule V depending on medicinal uses, potential for abuse, and safety [80]. Notably, heroin is Schedule I and has no accepted medical use in the U.S. Opium and the natural opiates morphine, codeine, and thebaine fall under Schedule II. Fentanyl, hydromorphone, and oxycodone are also Schedule II. Technically, the opium poppy itself (except for the seeds) is also considered a Schedule II controlled substance, though this is not enforced unless other illicit activities are suspected (e.g., producing opium), and *P. somniferum* is a popular ornamental flower with seeds that are readily available [188].

Opioid epidemic

Widespread use of opioids has caused an opioid epidemic in the U.S. This refers to a growing number of deaths and hospitalizations due to opioid use, including prescription opioids, illicit drugs such as heroin, and opioid analogs [189]. Between 2000 and 2014, the number of overdose deaths involving opioids doubled [190]. In 2018, 10.3 million people misused prescription opioids, and 47,600 people died from opioid overdoses [191]. Drug overdose is the number one cause of accidental death, and about 68% of overdose deaths involved opioids (1999-2017) [192]. Many opioid abusers first become addicted after receiving a prescription. Once their prescription ends, people must either endure withdrawal symptoms or turn to other methods of getting high, including heroin, which is cheaper and more potent than prescription medications, or poppy seed tea. In fact, about 80% of heroin users initially became dependent due to using a prescription opioid [189].

Forensic identification

Several color tests are available for presumptive identification of opioids, including Marquis (violet), Liebermann's (black), Mecke (blue-green), and Froehde (color dependent on specific opioid) [193]. To confirm the presence of morphine, codeine, and other opioids or opioid metabolites, GC-MS, liquid chromatography-mass spectrometry (LC-MS), and other chemical methods may be utilized [194-198].

Heroin trafficking

The DEA runs two programs to provide information about trends in trafficking of heroin in the United States. The Heroin Signature Program (HSP) provides information about the geographic origin of manufacture for wholesale-level heroin [199], and the Heroin Domestic Monitoring Program (HDMP) determines the geographic origin, purity and price of heroin sold on the street and purchased by undercover law enforcement (retail-level) [200]. The HSP performs chemical analysis of sample constituents, allowing association of samples with a particular geographic source area and processing method. According to the 2016 Heroin Signature Program Report (published in 2018) [199], 86% (by weight) of heroin came from Mexico, 4% came from South America, less than 1% came from southwest Asia, and 10% was inconclusively assigned to either Mexico or South America. The HDMP indicated that the average purity of Mexican heroin was 31.5% and cost $0.84 per milligram of pure heroin. Caffeine, diphenhydramine, acetaminophen, quinine, procaine, and lidocaine were commonly observed adulterants, and fentanyl, one of the most potent and dangerous opioids available, was identified in 11% of Mexican heroin samples tested [200].

The ability to determine geographic origins of manufacture aid law enforcement in identifying entry points into the country and blocking the entry of more drugs. However, the utility of chemical analysis is limited, and including genetic methods of tracking samples could strengthen the program and help to elucidate trafficking and distribution routes. Until recently, it was thought that DNA could not survive the temperature and pH ranges involved in heroin production. In 2018, however, Marciano et al. [201] developed a method to extract *P. somniferum* DNA from heroin and showed that it was of sufficient quantity and quality to be used for genetic analysis.

Genetic studies

Genome

The *P. somniferum* genome is diploid (2N=22) [202] and has been recently fully sequenced and annotated [203]. Despite its forensic importance, there have been few genetic studies on the opium poppy.

Individualization

Several studies have investigated the utility of genetic markers for identification and genetic diversity studies in *P. somniferum* [204-212]. Random amplified polymorphic DNA (RAPD) [204], inter-simple sequence repeat (ISSR) [204, 205], and amplified fragment length polymorphism (AFLP) [206] techniques have been used to study the genetic diversity. More recent techniques include the development of STR and expressed sequence tag derived SSR (EST-SSR) markers [207-212]. Several polymorphic STR loci have been identified. However, to date a successful multiplex of SSR markers for the forensic identification of opium poppy material has not been developed.

STRs

Several studies have found that trinucleotide repeats are the most common STRs found in opium poppy, with the AAG repeat being the most common [208-211]. Many of the markers that have been discovered are highly transferable within the *Papaver* genus [209, 210] as well as the genus *Eschscholzia* [208].

A forensically relevant study conducted by Lee et al. [208] developed primers for 22 EST-SSR loci, named psom(#). Six of the markers were polymorphic (2-5 alleles per locus) and proved capable of individualizing opium poppy plants. However, EST-SSRs are highly conserved compared to genomic STRs [213], and the authors stressed the need for more variable markers to enhance the discrimination power for individual identifications. Additionally, six of the reported markers were dinucleotide repeats, which are prone to elevated stutter and result in profiles that are difficult to interpret. For this reason, the International Society of Forensic Genetics (ISFG) recommends using tetra- or pentanucleotide STRs [214]. Another study by Lee et al. [207] reported STRs located in genes for enzymes involved in morphine biosynthesis, which may be able to identify plants grown for edible poppy seeds and plants intended for pharmaceutical purposes.

Şelale et al. [209], Celik et al. [210], and Masárová et al. [211] mined the available sequence data and identified tens of thousands of STR loci, though only a few primer sets were developed. Celik et al. [210] tested the ability of 53 novel genomic STRs to differentiate 37 Turkish poppy samples and discovered that 32 of the markers were polymorphic. Şelale et al. [209] reported on the development 67 EST-SSRs, which were tested in a limited number of samples, and 53 were polymorphic. Mičianová et al. [212] tested several of the previously reported loci using poppy samples from Slovakia, the

Czech Republic, and Austria. Of the twelve loci tested, eight were polymorphic, and all thirteen samples had unique genotypes. Only 2-4 alleles were observed at each locus, which is lower than ideal for individualization purposes, and four of the loci tested were mono- or dinucleotide STRs.

Kati et al. [215] reported 30 STR loci in *Papaver rhoeas* (corn poppy), a close relative of *P. somniferum*. Though the markers were not tested in opium poppy, many SSR markers are transferable across species within the same family. Of the eleven polymorphic loci, eight were dinucleotide STRs and the remaining three were trinucleotide. Though tetranucleotides are preferred by the forensic community, the three trinucleotide STRs may be worth testing for variability in opium poppy. Additionally, a full, annotated genome for *P. somniferum* was recently reported [203], providing the ability for more STRs to be discovered.

Several of the reported markers have shown promising preliminary results, but much more testing is needed to assess the diversity of the loci and choose informative markers. Once markers are chosen, a multiplex assay should be developed to quickly genotype samples and provide a high power of discrimination. Additionally, little information can be gained from STR results without a representative database with allele frequencies, and the loci must be shown to be unlinked and in Hardy-Weinberg equilibrium. An STR multiplex for *P. somniferum* should follow the development and validation guidelines for non-human analyses set by ISFG [214]. A validated genetic method for *P. somniferum* would prove valuable to the forensic community by linking cases and providing leads for investigations of illegal operations.

Eucalyptus

Taxonomy

Eucalyptus is a genus of flowering trees and shrubs in the myrtle family (Myrtaceae). There are over 700 species in the genus [216], and they are commonly referred to as eucalypts. They are native to the Southern hemisphere, mostly Australia, but eucalypts are highly adaptable to environmental conditions and are grown in plantations world-wide for their important industrial uses [217]. Two of the most important species around the world are *Eucalyptus globulus* (blue gum) and *Eucalyptus nitens* (shining gum) [218]. Their taxonomic classifications are shown in Table 1.3 [67].

Table 1.3 Taxonomic classification of *E. globulus* and *E. nitens*

Domain	Eukaryota
Kingdom	Plantae
Subkingdom	Viridiplantae (green plants)
Infrakingdom	Streptophyta (land plants)
Superdivision	Embryophyta
Division	Tracheophyta (vascular plants)
Subdivision	Spermatophytina (seed plants)
Class	Magnoliopsida
Superorder	Rosanae
Order	Myrtales
Family	Myrtaceae (myrtles)
Genus	Eucalyptus L'Heritier (gum)
Species	*Eucalyptus globulus* Labill., *Eucalyptus nitens*

Industrial uses

Most plantations of eucalyptus support the paper production and solid wood and veneer industries [219]. Other uses include providing charcoal to supply energy for the manufacture of steel and for firewood, shade, and shelter [220]. Eucalyptus oil is also highly sought after for its health effects and as an industrial solvent, antiseptic, and insect

repellant [221-224]. However, compounds in the oil are toxic to most animals, including humans [225], though koalas and other marsupials are resistant to these toxins and rely on eucalypts as a major food source.

Eucalyptus oil is used as a repellent chemical in insecticides and herbicides [221, 226]. It is considered safer for the environment since it is biodegradable and less toxic than many alternative repellents. The oil also possesses a range of useful medicinal properties. Cineol (1,8-cineol; also known as eucalyptol), a monoterpene, is the major component of eucalyptus oil (over 70%). Other compounds include macrocarpals, other monoterpenes, oxygenated monoterpenes, oxygenated sesquiterpenes, eucalyptin, phenols, flavonoids, oleanolic acid, tannins, terpenoid phenolaldehydes, and verbenone [226]. Eucalyptus essential oils have demonstrated antimicrobial activity and are effective against bacteria, fungi, some viruses, and other pathogens [222-224, 226-228]. The oil has also shown antidiabetic and antioxidant properties through the reduction of oxidative stress and lowering of plasma glucose levels in diabetic rats [229-231]. The oil also has antihistaminic [232], anti-inflammatory [233, 234], antimalarial [235], antioxidant [236, 237], cytotoxic (antitumor) [238-240], nociceptive (pain-killing) [241, 242], and wound-healing [243] properties.

Illegal logging

Illegal logging is a worldwide problem [244] that costs billions of dollars annually. The U.S. Forest Service estimates that up to ten percent of trees cut in national forests are taken illegally, and tree theft accounts for approximately one billion dollars annually in the U.S [245]. Despite these numbers, arrests and prosecutions for illegal logging activities are uncommon [244, 245]. Unfortunately, state and federal investigators are too busy

investigating other crimes, and forestry operations take place in large, remote areas where monitoring is nearly impossible. Additionally, no routine forensic testing is available to provide courtroom evidence of illegal logging. Individualization techniques, similar to those used in human identification, could provide evidence to match an illegally obtained log to its remaining tree stump, or a genetic similarity to nearby trees could help to identify the forest or population a log was cut from.

Genetic studies

Genome

Eucalypts are diploid with eleven chromosome pairs (2n=22) [246-248]. The draft genome of *E. grandis* has been reported [249]. It is 640 Mb with 36,375 protein-coding genes. While much of the genome is conserved, the genome sizes of eucalypts vary considerably; *E. globulus* has a genome of 530 Mb [246]. The size difference is almost completely accounted for (88.7 Mb) by small, non-transposable element-derived changes throughout the genome [249]. Additionally, comparison of *Eucalyptus* and *Corymbia*, a closely related genus, reveal that genome structure is largely conserved between the two genera [250].

Individualization

While there is a lack of forensic studies, forestry and conservation geneticists have identified over 1,200 STR markers in eucalyptus and the closely related genus *Corymbia* [251-277]. Over 800 genomic STR and EST-SSR markers with the prefix EMBRA have been developed in *E. grandis* and *E. urophylla* [253, 258, 261, 267, 269, 276]. Faria et al. [269] reported 21 EMBRA loci with tetra-, penta-, and hexanucleotide repeat motifs. These loci were incorporated into two multiplex STR assays which could be used for

individualization of four eucalyptus species. However, the assays did not include sequenced allelic ladders or comply with other ISFG recommendations for non-human DNA analysis [214], and a limited number of individuals were genotyped.

Additionally, 40 EST-SSR markers with the prefix EMCRC were developed in *E. globulus* [256], *Corymbia variegata* [255], and *Corymbia citriodora* subspecies *variegata* [262]. Steane et al. [278] reported a set of twelve dinucleotide STR loci (EMCRC 1-12) which were highly polymorphic in *E. nitens* and other eucalyptus species (14-21 alleles per locus). Another large source of EST-SSR markers is the 240 loci reported by Zhou et al. [275], 218 of which were found to be polymorphic. Bradbury et al. [274] developed 17 EST-SSRs in *E. gomphocephala* with the prefix EGM. He et al. reported seven genomic STRs (EUCgSSR). Glaubitz et al. [257] reported eight dinucleotide STRs in *E. sieberi* with the prefix Es. Ottewell et al. [260] reported eight polymorphic di- and trinucleotide loci in *E. leucoxylon* with the prefix El. Da Silva et al. [265] reported 13 dinucleotide and one pentanucleotide STRs (ECc) in *E. camaldulensis* which proved highly variable (4-13 alleles per locus). Van der Nest et al. [254] developed five STRs (FMRSA) with mono-, di-, tri-, hexa-, and nonanucleotide repeats but did not assess their variability. Nevill et al. [273] developed ten polymorphic loci (KPEV) with 5-25 alleles per locus (all di- and trinucleotide repeats) in *E. victrix*. Acuna et al. [279] proposed six EST-SSR markers in *E. globulus*, as well as a set of gene-specific STRs for wood quality [270, 271]. Thamarus et al. [259] also developed 40 STRs for wood, fiber, and floral traits (quantitative trait loci, QTL).

Many of these markers were found to be transferable between various *Eucalyptus* and *Corymbia* species [256, 257, 260, 262, 267, 269, 271, 272, 274, 280-287]. However,

several researchers reported the occurrence of null alleles when testing loci in species other than the one in which they were developed [251, 261, 278].

There is a plethora of STR markers that could be used for individualization of eucalypt species. Many of the markers have been used in conservation, breeding, diversity, parentage, mapping, and QTL studies [251]. However, to date, no publications reference the use of eucalyptus STRs in forensic casework, and no multiplex assays have been developed in accordance with ISFG guidelines [214]. Additionally, the variability of many of the markers in the literature remains unknown. Since a lack of forensic evidence for illegal logging results in many cases being dismissed, the forensic community needs a validated STR individualization method to identify trees which may have been illegally cut down and link them to a particular tree stump or location.

Standardization of non-human forensic genetics

SWGDAM

The Scientific Working Group on DNA Analysis Methods (SWGDAM) was preceded by the Technical Working Group on DNA Analysis Methods (TWGDAM), which first met in 1988. The purpose was to bring forensic scientists together to share protocols and establish validation guidelines. The guidelines became standards which are recognized by courts as minimum requirements to ensure quality forensic DNA analysis. . SWGDAM has published standards and guidelines for developmental and internal validations of DNA analysis methods to ensure quality of results. The group also recommends revisions to the FBI's Quality Assurance Standards (QAS) for DNA Analysis [288]. Laboratories must adhere to these QAS in order to participate in the National DNA Index System (NDIS). Publications by SWGDAM include Interpretation Guidelines for

Autosomal STR Typing by Forensic DNA Laboratories, Contamination Prevention and Detection Guidelines for Forensic DNA Laboratories, and Guidelines for the Validation of Probabilistic Genotyping Systems, to name a few, and their publications can be accessed online at https://www.swgdam.org/publications. The majority of SWGDAM guidelines are meant for analysis of human DNA, but following these guidelines as closely as possible for non-human DNA analyses guarantees robust and reliable results.

OSAC

The Organization of Scientific Area Committees for forensic science (OSAC) strengthens forensic science by publishing standards and guidelines and encouraging their use by the forensic science community. OSAC is made up of five scientific area committees (SACs) specializing in different areas of forensic science (Biology/DNA, Chemistry/Instrumental Analysis, Physics/Pattern Interpretation, Crime Scene/Death Investigation, and Digital Multimedia). Within the Biology/DNA SAC, the Wildlife Forensics subcommittee develops standards related to the use of public databases, methods for geographic assignment of animals, best practices for developing new STR panels, and other documents. Their published standard on validations of new STR panels (ANSI/ASB Standard 046) includes requirements for testing species specificity, sensitivity and stochastic effects, repeatability and reproducibility, case-type samples, population studies, and mixture studies, for example.

ISFG

The International Society for Forensic Genetics (ISFG), founded in 1968, promotes knowledge in forensic genetics and holds regular meetings at regional and international levels. They also publish recommendations on relevant issues in forensic genetics. ISFG

published specific recommendations for non-human DNA analyses in 2011 [214], which include validation parameters, the use of tetranucleotide STR loci, development and use of sequenced allelic ladders, and the estimation of relevant population and forensic genetic parameters, including allele frequencies. Though these recommendations apply specifically to animals, the same principles should apply to plant forensic analyses.

Statement of the problem

Forensic plant science is an underutilized discipline due to a lack of research and standardization within the field. It can be applied to cases involving drugs of plant origin, including marijuana (*Cannabis sativa*) and heroin (*Papaver somniferum*), allowing association of cases and potentially providing leads, such as the biogeographical origin of a sample or its crop type. Additionally, it can be used to provide evidence of environmental crimes, such as illegal logging, which currently goes unpunished due to a lack of evidence to support claims of wood theft. The purpose of this work was to develop forensic genetic techniques to provide investigative leads or evidence of criminal activity for trafficking of marijuana and heroin and illegal logging of eucalyptus.

C. sativa can be classified as marijuana (a drug with psychoactive properties) or hemp (the form which lacks the psychoactive chemical THC and is used for fiber production). Marijuana is a Schedule I controlled substance and the most commonly used drug in the United States. However, its use has been legalized in 33 states, creating a new challenge for law enforcement: investigating the diversion of legal cannabis to states where its possession is unlawful. Additionally, cannabis trafficking into the U.S. remains a problem, especially at the U.S.-Mexico border. Differentiation between marijuana and hemp is important for forensic purposes because hemp is legal in the U.S. Additionally,

discovering the biogeographical origin (the country or geographic region in which the marijuana strain originated) could provide important leads for law enforcement investigating marijuana trafficking.

One purpose of this work was to investigate genetic differences in *C. sativa* nuclear and chloroplast DNA to associate samples with a particular biogeographic origin, discriminate between marijuana and hemp, and provide a method for linking seizures from common growers and distributors. DNA barcoding markers in the chloroplast of *C. sativa* have previously shown potential for differentiating between samples from different biogeographical origins and determining crop type [36, 108]. However, a larger set of polymorphic markers is needed to provide discriminatory results and differentiate between closely related samples. There is no consensus on the most discriminatory barcoding markers for plants [52, 289], so this work sought to discover "hotspots" of polymorphisms in the chloroplast genome of *C. sativa* and assess their variability in hemp and marijuana samples from different countries. Additionally, a massively parallel sequencing (MPS) assay was developed for simultaneous analysis of seven chloroplast DNA hotspots in a high-throughput manner. The creation of a database with representative samples from around the world is necessary for determination of biogeographical origin and crop type using this method, and the developed MPS assay will be useful in creating a database and providing more discriminatory results.

Another method of determining crop type that has been investigated is analysis of the THCA and CBDA synthase genes, which control the THC content of *C. sativa* plants, resulting in their classification as either hemp or marijuana [92]. However, several studies have called into question the viability of these genetic tests, as chemotype inheritance may

be more complicated than originally proposed [127, 128]. To contribute to this growing body of literature and test the chemotype inheritance model proposed by Kojoma et al. [92], a single nucleotide polymorphism (SNP) assay previously published by Rotherham and Harbison [120] was optimized and evaluated for determination of crop type using hemp and marijuana samples from North and South America.

Papaver somniferum is the source of opiates including morphine and codeine, which have analgesic properties. Several pharmaceutical drugs contain opiates or their derivatives (opioids), and many of these drugs have a high abuse potential. Individuals can become dependent on opioids during the course of a prescription treatment and turn to illegal alternatives once the prescription ends. Heroin is a Schedule I drug derived from morphine that is twice as potent and more addictive than the natural opiate. The DEA has two programs to address heroin trafficking in the U.S.: the Heroin Signature Program (HSP) and the Heroin Domestic Monitoring Program (HDMP), which use chemical signatures of the drugs to determine the geographic area where they were manufactured. Addition of a genetic method to such programs could provide a more sensitive and discriminatory method of individualizing heroin samples and determining their biogeographical origin. Recently, *P. somniferum* DNA suitable for genetic analyses has been successfully extracted from heroin [201]. Only a few studies have proposed STR markers for individualization of *P. somniferum* [208-212], and few of these markers have been evaluated for their variability in a diverse sample set. This work sought to evaluate previously published STR markers and develop a forensically-relevant STR panel following ISFG guidelines for non-human DNA analyses [214] for individualization and

biogeographical origin determination of *P. somniferum* samples, including heroin and poppy seeds.

Eucalyptus (gum trees) are grown around the world for production of paper pulp, solid wood, shade, and other uses. Illegal logging, consisting of wood theft or illegal cutting of the trees, costs the world economy billions of dollars each year and brings hardship to plantation owners and countries whose economies rely on the export of timber. Unfortunately, many of these crimes go unpunished due to difficulty in monitoring forests and lack of forensic evidence. Over 1,200 STR markers have previously been identified in eucalyptus and related species [251], but their viability in forensic individualization has not yet been explored. This work sought to evaluate available STR markers and demonstrate their use as forensic evidence by applying them to a case study involving suspected wood theft in Chile.

References

1. Bock JH, Norris DO. (2015) Forensic plant science. Academic Press
2. Coyle HM, Ladd C, Palmbach T, Lee HC (2001) The green revolution: botanical contributions to forensics and drug enforcement. Croat Med J 42: 340-5
3. Chandra R, Sharma V (2014) Forensic botany: an emerging discipline of plant sciences. Indian Botanists Blog-o-Journal: 1-8
4. Hall DW (1988) The contributions of the forensic botanist to crime scene investigations. The Prosecutor 22: 35-8
5. Graham SA (1997) Anatomy of the Lindbergh kidnapping. J Forensic Sci 42: 368-77
6. Federal Bureau of Investigation. Lindbergh Kidnapping

7. Horrocks M, Walsh KA (1998) Forensic palynology: assessing the value of the evidence. Rev Palaeobot Palyno 103: 69-74
8. Yoon CK (1993) Botanical witness for the prosecution. Science 260: 894-6
9. Siver PA, Lord WD, McCarthy DJ (1994) Forensic limnology: the use of freshwater algal community ecology to link suspects to an aquatic crime scene in southern New England. J Forensic Sci 39: 847-53
10. Bock JH, Norris DO (1997) Forensic botany: an under-utilized resource. J Forensic Sci 42: 364-7
11. Devane WA, Dysarz Fr, Johnson MR, Melvin LS, Howlett AC (1988) Determination and characterization of a cannabinoid receptor in rat brain. Mol Pharmacol 34: 605-13
12. Hill KP (2015) Medical marijuana for treatment of chronic pain and other medical and psychiatric problems: a clinical review. Jama 313: 2474-83
13. Substance Abuse and Mental Health Services Administration (2018) Key substance use and mental health indicators in the United States: Results from the 2017 National Survey on Drug Use and Health (HHS Publication No. SMA 18-5068, NSDUH Series H-53). Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration Rockville, MD
14. Labanca F, Ovesnà J, Milella L (2018) Papaver somniferum L. taxonomy, uses and new insight in poppy alkaloid pathways. Phytochem Rev 17: 853-71. doi: 10.1007/s11101-018-9563-3
15. Kapadia GJ, Fayez M (1970) Peyote constituents: chemistry, biogenesis, and biological effects. J Pharm Sci 59: 1699-727

16. Drug Enforcement Administration (2018) 2018 National Drug Threat Assessment. In: Justice USDo, ed.
17. World Wide Fund for Nature (2019) Illegal Logging. https://wwf.panda.org/our_work/forests/deforestation_causes2/illegal_logging/
18. Koch GT (2019) Best Practices on Tools for Combating and Preventing Illegal Logging Activities and Associated Trade. Asia-Pacific Economic Cooperation
19. Porebski S, Bailey LG, Baum BR (1997) Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Mol Biol Rep 15: 8-15
20. Henry RJ (2001) Plant DNA extraction. Plant Genotyping: the DNA fingerprinting of plants: 239-49
21. Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue
22. Doyle J. (1991) DNA protocols for plants. Molecular techniques in taxonomy. Springer. pp. 283-93
23. QIAGEN. (2018) DNeasy® Plant Handbook. Hilden, Germany
24. Gitzendanner MA (2012) Use and guidelines for plant DNA analyses in forensics. Forensic Botany: A Practical Guide: 93-106
25. Csaikl U, Bastian H, Brettschneider R et al (1998) Comparative analysis of different DNA extraction protocols: a fast, universal maxi-preparation of high quality plant DNA for genetic evaluation and phylogenetic studies. Plant Mol Biol Rep 16: 69-86
26. Kang HW, Cho YG, Yoon UH, Eun MY (1998) A rapid DNA extraction method for RFLP and PCR analysis from a single dry seed. Plant Mol Biol Rep 16: 90

27. Xin Z, Chen J (2012) A high throughput DNA extraction method with high yield and quality. Plant Methods 8: 26

28. Dermen H (1940) Colchicine polyploidy and technique. Bot Rev 6: 599-635

29. Butler JM. (2011) Advanced topics in forensic DNA typing: methodology. Academic Press

30. Birky CW (1995) Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. Proc Natl Acad Sci 92: 11331-8

31. Wang D-Y, Zhang Q, Liu Y, Lin Z-F, Zhang S-X, Sun M-X (2010) The levels of male gametic mitochondrial DNA are highly regulated in angiosperms with regard to mitochondrial inheritance. Plant Cell 22: 2402-16

32. Hagelberg E, Clegg JB (1991) Isolation and characterization of DNA from archaeological bone. Proceedings of the Royal Society of London Series B: Biol Sci 244: 45-50

33. Royal CD, Novembre J, Fullerton SM et al (2010) Inferring genetic ancestry: opportunities, challenges, and implications. Am J Hum Genet 86: 661-73

34. Lao O, Lu TT, Nothnagel M et al (2008) Correlation between genetic and geographic structure in Europe. Curr Biol 18: 1241-8

35. Novembre J, Johnson T, Bryc K et al (2008) Genes mirror geography within Europe. Nature 456: 98-101

36. Gilmore S, Peakall R, Robertson J (2007) Organelle DNA haplotypes reflect crop-use characteristics and geographic origins of Cannabis sativa. Forensic Sci Int 172: 179-90

37. Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S (2015) Plant DNA barcoding: from gene to genome. Biol Rev 90: 157-66

38. Ogden R, Linacre A (2015) Wildlife forensic science: a review of genetic geographic origin assignment. Forensic Sci Int Genet 18: 152-9

39. Jeffreys AJ, Wilson V, Thein SL (1985) Hypervariable 'minisatellite'regions in human DNA. Nature 314: 67-73

40. Tanksley S, Young N, Paterson A, Bonierbale M (1989) RFLP mapping in plant breeding: new tools for an old science. Biotechnol 7: 257-64

41. Nybom H, Weising K, Rotter B (2014) DNA fingerprinting in botany: past, present, future. Investig Genet 5: 1

42. Hubbard M, Kelly J, Rajapakse S, Abbott A, Ballard R (1992) Restriction fragment length polymorphisms in rose and their use for cultivar identification. HortScience 27: 172-3

43. Hallerman EM (2003) Population genetics: principles and applications for fisheries scientists

44. Ishfaq M, Saleem M (2016) Application of molecular markers in forensic botany. Bull Biol Allied Sci Res 1: 1-7

45. Mestel R (1993) Murder trial features tree's genetic fingerprint. New Scientist 138: 6.

46. Godwin ID, Aitken EA, Smith LW (1997) Application of inter simple sequence repeat (ISSR) markers to plant genetics. Electrophoresis 18: 1524-8

47. Vos P, Hogers R, Bleeker M et al (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23: 4407-14

48. Mueller UG, Wolfenbarger LL (1999) AFLP genotyping and fingerprinting. Trends Ecol Evol 14: 389-94

49. Fan H, Chu J-Y (2007) A brief review of short tandem repeat mutation. Genomics, Proteomics & Bioinformatics 5: 7-14

50. Ferri G, Alù M, Corradini B, Beduschi G (2009) Forensic botany: species identification of botanical trace evidence using a multigene barcoding approach. Int J Legal Med 123: 395-401

51. Hebert PD, Ratnasingham S, De Waard JR (2003) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc R Soc B Biol Sci 270: S96-S9

52. Hollingsworth PM, Graham SW, Little DP (2011) Choosing and using a plant DNA barcode. PLOS One 6: e19254

53. Hollingsworth PM, Li D-Z, van der Bank M, Twyford AD (2016) Telling plant species apart with DNA: from barcodes to genomes. Phil Trans R Soc B 371: 20150338

54. Ferri G, Corradini B, Ferrari F, Santunione AL, Palazzoli F, Alu M (2015) Forensic botany II, DNA barcode for land plants: Which markers after the international agreement? Forensic Sci Int Genet 15: 131-6

55. Bell KL, De Vere N, Keller A et al (2016) Pollen DNA barcoding: current applications and future prospects. Genome 59: 629-40

56. Röhl A, Brinkmann B, Forster L, Forster P (2001) An annotated mtDNA database. Int J Legal Med 115: 29-39

57. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci 74: 5463-7

58. Moorthie S, Mattocks CJ, Wright CF (2011) Review of massively parallel DNA sequencing technologies. HUGO J 5: 1-12

59. Bruijns B, Tiggelaar R, Gardeniers H (2018) Massively parallel sequencing techniques for forensics: a review. Electrophoresis 39: 2642-54

60. Rothberg JM, Hinz W, Rearick TM et al (2011) An integrated semiconductor device enabling non-optical genome sequencing. Nature 475: 348-52

61. Bentley DR, Balasubramanian S, Swerdlow HP et al (2008) Accurate whole human genome sequencing using reversible terminator chemistry. nature 456: 53-9

62. Moliterni VC, Cattivelli L, Ranalli P, Mandolino G (2004) The sexual differentiation of Cannabis sativa L.: a morphological and molecular study. Euphytica 140: 95-106

63. Ashton CH (2001) Pharmacology and effects of cannabis: a brief review. Br J Psychiatry 178: 101-6

64. De Pauw MA, Vidmar JJ, Collins J, Bennett RA, Deyholos MK (2007) Microarray analysis of bast fibre producing tissues of Cannabis sativa identifies transcripts associated with conserved and specialised processes of secondary wall development. Funct Plant Biol 34: 737-49

65. Watts G (2006) Cannabis confusions. Bmj 332: 175-6

66. Yang M-Q, van Velzen R, Bakker FT, Sattarian A, Li D-Z, Yi T-S (2013) Molecular phylogenetics and character evolution of Cannabaceae. Taxon 62: 473-85

67. Retrieved February 29, 2020 from the Integrated Taxonomic Information System on-line database, http://www.itis.gov

68. Gaoni Y, Mechoulam R (1964) Isolation, structure, and partial synthesis of an active constituent of hashish. J Am Chem Soc 86: 1646-7

69. Mehmedic Z, Chandra S, Slade D et al (2010) Potency trends of Δ9-THC and other cannabinoids in confiscated cannabis preparations from 1993 to 2008. J Forensic Sci 55: 1209-17

70. ElSohly MA, Mehmedic Z, Foster S, Gon C, Chandra S, Church JC (2016) Changes in cannabis potency over the last 2 decades (1995–2014): analysis of current data in the United States. Biol Psychiatry 79: 613-9

71. Stuyt E (2018) The problem with the current high potency THC marijuana from the perspective of an addiction psychiatrist. Missouri Med 115: 482

72. Herkenham M (1995) Localization of cannabinoid receptors in brain and periphery. Cannabinoid receptors

73. Munro S, Thomas KL, Abu-Shaar M (1993) Molecular characterization of a peripheral receptor for cannabinoids. Nature 365: 61

74. Tanda G, Loddo P, Di Chiara G (1999) Dependence of mesolimbic dopamine transmission on Δ9-tetrahydrocannabinol. Eur J Pharmacol 376: 23-6

75. de Meijer EP, Bagatta M, Carboni A et al (2003) The inheritance of chemical phenotype in Cannabis sativa L. Genetics 163: 335-46

76. Pacifico D, Miselli F, Micheler M, Carboni A, Ranalli P, Mandolino G (2006) Genetics and Marker-assisted Selection of the Chemotype in Cannabis sativa L. Mol Breed 17: 257-68

77. Li H-L (1974) The origin and use of Cannabis in eastern Asia linguistic-cultural implications. Econ Bot 28: 293-301

78. Zuardi AW (2006) History of cannabis as a medicine: a review. Braz J Psychiatry 28: 153-7

79. O'Shaughnessy WB (1843) On the preparations of the Indian hemp, or Gunjah: Cannabis indica their effects on the animal system in health, and their utility in the treatment of tetanus and other convulsive diseases. Provincial Medical Journal and Retrospect of the Medical Sciences 5: 363

80. Drug Enforcement Administration (2010) Pharmacist's manual: an informational outline of the Controlled Substances Act. Springfield: United States Department of Justice, Drug Enforcement Administration, Office of Diversion Control

81. Pal L, Lucia LA (2019) Renaissance of industrial hemp: A miracle crop for a multitude of products. BioResources 14: 2460-4

82. (2018) Agriculture Improvement Act of 2018 (115-334)

83. United Nations (1961) Single Convention on Narcotic Drugs, 1961

84. National Institute on Drug Abuse (2019) NIDA's role in providing marijuana for research

85. Thornton JI, Nakamura GR (1972) The identification of marijuana. J Forensic Sci Soc 12: 461-519

86. Mitosinka GT, Thornton JI, Hayes TL (1972) The Examination of Cystolithic Hairs of Cannabis and other Plants by means of the Scanning Electron Microscope. J Forensic Sci Soc12: 521-9

87. Jacobs AD, Steiner RR (2014) Detection of the Duquenois–Levine chromophore in a marijuana sample. Forensic Sci Int 239: 1-5

88. F Kelly J, Addanki K, Bagasra O (2012) The non-specificity of the Duquenois-Levine field test for marijuana. Open Forensic Sci J 5

89. Hughes R, Warner V (1978) A study of false positives in the chemical identification of marihuana. J Forensic Sci 23: 304-10

90. Pacifico D, Miselli F, Carboni A, Moschella A, Mandolino G (2008) Time course of cannabinoid accumulation and chemotype development during the growth of Cannabis sativa L. Euphytica 160: 231-40

91. Vergara D, Baker H, Clancy K et al (2016) Genetic and genomic tools for Cannabis sativa. Crit Rev Plant Sci 35: 364-77

92. Kojoma M, Seki H, Yoshida S, Muranaka T (2006) DNA polymorphisms in the tetrahydrocannabinolic acid (THCA) synthase gene in “drug-type” and “fiber-type” Cannabis sativa L. Forensic Sci Int 159: 132-40

93. Gilmore S, Peakall R, Robertson J (2003) Short tandem repeat (STR) DNA markers are hypervariable and informative in Cannabis sativa: implications for forensic investigations. Forensic Sci Int 131: 65-74

94. Sakamoto K, Akiyama Y, Fukui K, Kamada H, Satoh S (1998) Characterization; genome sizes and morphology of sex chromosomes in hemp (Cannabis sativa L.). Cytologia 63: 459-64

95. Van Bakel H, Stout JM, Cote AG et al (2011) The draft genome and transcriptome of Cannabis sativa. Genome Biol 12: R102

96. Oh H, Seo B, Lee S et al (2016) Two complete chloroplast genome sequences of Cannabis sativa varieties. Mitochondrial DNA Part A 27: 2835-7

97. Vergara D, White KH, Keepers KG, Kane NC (2016) The complete chloroplast genomes of Cannabis sativa and Humulus lupulus. Mitochondrial DNA Part A 27: 3793-4

98. Matielo CBDO, Lemos RPM, Sarzi DS et al (2020) Whole Plastome Sequences of Two Drug-Type Cannabis: Insights Into the Use of Plastid in Forensic Analyses. J Forensic Sci 65: 259-65

99. White KH, Vergara D, Keepers KG, Kane NC (2016) The complete mitochondrial genome for Cannabis sativa. Mitochondrial DNA Part B 1: 715-6

100. Coyle HM, Palmbach T, Juliano N, Ladd C, Lee HC (2003) An overview of DNA methods for the identification and individualization of marijuana. Croat Med J 44: 315-21

101. Dias V, Ribeiro A, Mello I et al (2015) Genetic identification of Cannabis sativa using chloroplast trnL-F gene. Forensic Sci Int Genet14: 201-2

102. Gigliano GS, Caputo P, Cozzolino S (1997) Ribosomal DNA analysis as a tool for the identification of Cannabis sativa L. specimens of forensic interest. Sci Justice 37: 171-4

103. Gigliano SS (1999) Preliminary data on the usefulness of internal transcribed spacer I (ITS1) sequence in Cannabis sativa L. identification. J Forensic Sci 44: 475-7

104. Gigliano GS (1998) Identification of Cannabis sativa L.(Cannabaceae) using restriction profiles of the Internal Transcribed Spacer II (ITS2). Sci Justice 38: 225-30

105. Tsai L-C, Hsieh H-M, Wang J-C, Huang L-H, Linacre A, Lee JC-I (2006) Cannabis seed identification by chloroplast and nuclear DNA. Forensic Sci Int 158: 250-2

106. Kohjyouma M, LEE I-J, Iida O et al (2000) Intraspecific variation in Cannabis sativa L. based on intergenic spacer region of chloroplast DNA. Biol Pharm Bull 23: 727-30

107. Linacre A, Thorpe J (1998) Detection and identification of cannabis by DNA. Forensic Sci Int 91: 71-6

108. Houston R, Birck M, LaRue B, Hughes-Stamm S, Gangitano D (2018) Nuclear, chloroplast, and mitochondrial data of a US cannabis DNA database. Int J Legal Med: 1-13

109. Howard C, Gilmore S, Robertson J, Peakall R (2009) A Cannabis sativa STR genotype database for Australian seizures: forensic applications and limitations. J Forensic Sci 54: 556-63

110. Datwyler SL, Weiblen GD (2006) Genetic variation in hemp and marijuana (Cannabis sativa L.) according to amplified fragment length polymorphisms. J Forensic Sci 51: 371-5

111. Kojoma M, Iida O, Makino Y, Sekita S, Satake M (2002) DNA fingerprinting of Cannabis sativa using inter-simple sequence repeat (ISSR) amplification. Planta Medica 68: 60-3

112. Hakki EE, Kayis SA, Pinarkara E, Sag A (2007) Inter simple sequence repeats separate efficiently hemp from marijuana (Cannabis sativa L.). Electron J Biotechn 10: 570-81

113. Pinarkara E, Kayis SA, Hakki EE, Sag A (2009) RAPD analysis of seized marijuana (Cannabis sativa L.) in Turkey. Electron J Biotechn 12: 5-6

114. Jagadish V, Robertson J, Gibbs A (1996) RAPD analysis distinguishes Cannabis sativa samples from different sources. Forensic Sci Int 79: 113-21

115. Faeti V, Mandolino G, Ranalli P (1996) Genetic diversity of Cannabis sativa germplasm based on RAPD markers. Plant Breed 115: 367-70

116. Shirota O, Watanabe A, Yamazaki M et al (1998) Random amplified polymorphic DNA and restriction fragment length polymorphism analyses of Cannabis sativa. Natural medicines= 生薬學雜誌 52: 160-6

117. Forapani S, Carboni A, Paoletti C, Moliterni V, Ranalli P, Mandolino G (2001) Comparison of hemp varieties using random amplified polymorphic DNA markers. Crop Sci 41: 1682-9

118. Mello I, Ribeiro A, Dias V et al (2016) A segment of rbcL gene as a potential tool for forensic discrimination of Cannabis sativa seized at Rio de Janeiro, Brazil. Int J Legal Med 130: 353-6

119. Taura F, Morimoto S, Shoyama Y, Mechoulam R (1995) First direct evidence for the mechanism of. DELTA. 1-tetrahydrocannabinolic acid biosynthesis. J Am Chem Soc 117: 9766-7

120. Rotherham D, Harbison S (2011) Differentiation of drug and non-drug Cannabis using a single nucleotide polymorphism (SNP) assay. Forensic Sci Int 207: 193-7

121. Cirovic N, Kecmanovic M, Keckarevic D, Markovic MK (2017) Differentiation of Cannabis subspecies by THCA synthase gene analysis using RFLP. J Forensic Legal Med 51: 81-4

122. Staginnus C, Zörntlein S, de Meijer E (2014) A PCR marker Linked to a THCA synthase Polymorphism is a Reliable Tool to Discriminate Potentially THC-Rich Plants of C annabis sativa L. J Forensic Sci 59: 919-26
123. Cascini F, Farcomeni A, Migliorini D et al (2019) Highly Predictive Genetic Markers Distinguish Drug-Type from Fiber-Type Cannabis sativa L. Plants 8: 496
124. Weiblen GD, Wenger JP, Craft KJ et al (2015) Gene duplication and divergence affecting drug content in Cannabis sativa. New Phytologist 208: 1241-50
125. Onofri C, de Meijer EP, Mandolino G (2015) Sequence heterogeneity of cannabidiolic-and tetrahydrocannabinolic acid-synthase in Cannabis sativa L. and its relationship with chemical phenotype. Phytochem 116: 57-68
126. Borna T, Salami SA, Shokrpour M (2017) High resolution melting curve analysis revealed SNPs in major cannabinoid genes associated with drug and non-drug types of cannabis. Biotechnol Biotechnol Equip 31: 839-45
127. Toth JA, Stack GM, Cala AR et al Development and validation of genetic markers for sex and cannabinoid chemotype in Cannabis sativa L. GCB Bioenergy
128. Laverty KU, Stout JM, Sullivan MJ et al (2019) A physical and genetic map of Cannabis sativa identifies extensive rearrangements at the THC/CBD acid synthase loci. Genome Res 29: 146-56
129. Vergara D, Huscher EL, Keepers KG et al (2019) Gene copy number is associated with phytochemistry in Cannabis sativa. AoB Plants 11: plz074
130. McKernan KJ, Helbert Y, Tadigotla V et al (2015) Single molecule sequencing of THCA synthase reveals copy number variation in modern drug-type Cannabis sativa L. BioRxiv: 028654

131. Grassa CJ, Wenger JP, Dabney C et al (2018) A complete Cannabis chromosome assembly and adaptive admixture for elevated cannabidiol (CBD) content. BioRxiv: 458083

132. Gillan R, Cole M, Linacre A, Thorpe J, Watson N (1995) Comparison of Cannabis sativa by random amplification of polymorphic DNA (RAPD) and HPLC of cannabinoids: a preliminary study. Sci Justice J Forensic Sci Soc 35: 169-77

133. Kayis SA, Hakki EE, Pinarkara E (2010) Comparison of effectiveness of ISSR and RAPD markers in genetic characterization of seized marijuana (Cannabis sativa L.) in Turkey. Afr J Agric Res 5: 2925-33

134. Miller HC, Shutler G, Abrams S et al (2003) A simple DNA extraction method for marijuana samples used in amplified fragment length polymorphism (AFLP) analysis. J Forensic Sci 48: 343-7

135. Hu Z-G, Guo H-Y, Hu X-L et al (2012) Genetic diversity research of hemp (Cannabis sativa L) cultivar based on AFLP analysis. J Plant Genet Resour 13: 555-61

136. Flachowsky H, Schumann E, Weber W, Peil A (2001) Application of AFLP for the detection of sex-specific markers in hemp. Plant Breed 120: 305-9

137. Peil A, Flachowsky H, Schumann E, Weber W (2003) Sex-linked AFLP markers indicate a pseudoautosomal region in hemp (Cannabis sativa L.). Theor Appl Genet 107: 102-9

138. Kalia RK, Rai MK, Kalia S, Singh R, Dhawan A (2011) Microsatellite markers: an overview of the recent progress in plants. Euphytica 177: 309-34

139. Gilmore S, Peakall R (2003) Isolation of microsatellite markers in Cannabis sativa L.(marijuana). Mol Ecol Res 3: 105-7

140. Alghanim H, Almirall J (2003) Development of microsatellite markers in Cannabis sativa for DNA typing and genetic relatedness analyses. Anal Bioanal Chem 376: 1225-33

141. Hsieh H-M, Hou R-J, Tsai L-C et al (2003) A highly polymorphic STR locus in Cannabis sativa. Forensic Sci Int 131: 53-8

142. Valverde L, Lischka C, Erlemann S et al (2014) Nomenclature proposal and SNPSTR haplotypes for 7 new Cannabis sativa L. STR loci. Forensic Sci Int Genet 13: 185-6

143. Gao C, Xin P, Cheng C et al (2014) Diversity analysis in Cannabis sativa based on large-scale development of expressed sequence tag-derived simple sequence repeat markers. PloS One 9

144. Shirley N, Allgeier L, LaNier T, Coyle HM (2013) Analysis of the NMI01 marker for a population database of cannabis seeds. J Forensic Sci 58

145. Dufresnes C, Jan C, Bienert F, Goudet J, Fumagalli L (2017) Broad-scale genetic diversity of Cannabis for forensic applications. PloS One 12

146. Soler S, Gramazio P, Figàs MR et al (2017) Genetic structure of Cannabis sativa var. indica cultivars based on genomic SSR (gSSR) markers: implications for breeding and germplasm management. Ind Crops Prod104: 171-8

147. Howard C, Gilmore S, Robertson J, Peakall R (2008) Developmental validation of a Cannabis sativa STR multiplex system for forensic analysis. J Forensic Sci 53: 1061-7

148. Mendoza MA, Mills DK, Lata H, Chandra S, ElSohly MA, Almirall JR (2009) Genetic individualization of Cannabis sativa by a short tandem repeat multiplex system. Anal Bioanal Chem 393: 719-26

149. Köhnemann S, Nedele J, Schwotzer D, Morzfeld J, Pfeiffer H (2012) The validation of a 15 STR multiplex PCR for Cannabis species. Int J Legal Med 126: 601-6

150. Houston R, Birck M, Hughes-Stamm S, Gangitano D (2016) Evaluation of a 13-loci STR multiplex system for Cannabis sativa genetic identification. Int J Legal Med 130: 635-47

151. Houston R, Birck M, Hughes-Stamm S, Gangitano D (2017) Developmental and internal validation of a novel 13 loci STR multiplex method for Cannabis sativa DNA profiling. Legal Med 26: 33-40

152. Allgeier L, Hemenway J, Shirley N, LaNier T, Coyle HM (2011) Field testing of collection cards for Cannabis sativa samples with a single hexanucleotide DNA marker. J Forensic Sci 56: 1245-9

153. Houston R, Mayes C, King JL, Hughes-Stamm S, Gangitano D (2018) Massively parallel sequencing of 12 autosomal STRs in Cannabis sativa. Electrophoresis 39: 2906-11

154. López P, Pereboom-de Fauw DP, Mulder PP et al (2018) Straightforward analytical method to determine opium alkaloids in poppy seeds and bakery products. Food Chem 242: 443-50

155. Cannabis, Coca, & Poppy: Nature's Addictive Plants. DEA Museum and Visitors Center Arlington, VA

156. Schmitz R (1985) Friedrich Wilhelm Sertürner and the discovery of morphine. Pharm Hist 27: 61-74

157. Tétényi P (1997) Opium poppy (Papaver somniferum): botany and horticulture. Hortic Rev 19: 373-408

158. Pert CB, Snyder SH (1973) Opiate receptor: demonstration in nervous tissue. Science 179: 1011-4

159. Pert CB, Pasternak G, Snyder SH (1973) Opiate agonists and antagonists discriminated by receptor binding in brain. Science 182: 1359-61

160. Handal KA, Schauben JL, Salamone FR (1983) Naloxone. Ann Emerg Med 12: 438-45

161. Young AM, Havens JR, Leukefeld CG (2010) Route of administration for illicit prescription opioids: a comparison of rural and urban drug users. Harm Reduct J 7: 24

162. Drug Enforcement Administration (2017) Drugs of abuse - A DEA resource guide - 2017 edition. U.S. Department of justice, Drug Enforcement Administration. pp. 40-2,5

163. Armenian P, Vo KT, Barr-Walker J, Lynch KL (2018) Fentanyl, fentanyl analogs and novel synthetic opioids: a comprehensive review. Neuropharmacol 134: 121-32

164. Volpe DA, Tobin GAM, Mellon RD et al (2011) Uniform assessment and ranking of opioid Mu receptor binding constants for selected opioid drugs. Regul Toxicol Pharmacol 59: 385-90. doi: 10.1016/j.yrtph.2010.12.007

165. Mars SG, Bourgois P, Karandinos G, Montero F, Ciccarone D (2016) The textures of heroin: User perspectives on “black tar” and powder heroin in two US cities. J Psychoact Drugs 48: 270-8

166. Ciccarone D (2009) Heroin in brown, black and white: structural factors and medical consequences in the US heroin market. Int J Drug Policy 20: 277-82

167. Trafkowski J, Madea B, Musshoff F (2006) The significance of putative urinary markers of illicit heroin use after consumption of poppy seed products. Ther Drug Monit 28: 552-8

168. Chen P, Braithwaite R, George C et al (2014) The poppy seed defense: a novel solution. Drug Test Anal 6: 194-201

169. Thevis M, Opfermann G, Schänzer W (2003) Urinary concentrations of morphine and codeine after consumption of poppy seeds. J Anal Toxicol 27: 53-6

170. Lachenmeier DW, Sproll C, Musshoff F (2010) Poppy Seed Foods and Opiate Drug Testing-Where Are We Today? Ther Drug Monit 32: 11-8

171. United States Department of Justice (2010) DrugAlert Watch: Opium Tea. National Drug Intelligence Center, U.S. Department of Justice Washington, D.C.

172. Scott GN. (2017) Poppy Seed Tea: Beware This Beverage. Medscape

173. Bailey K, Richards-Waugh L, Clay D, Gebhardt M, Mahmoud H, Kraner JC (2010) Fatality involving the ingestion of phenazepam and poppy seed tea. J Anal Toxicol 34: 527-32

174. Braye K, Harwood T, Inder R, Beasley R (2007) Poppy seed tea and opiate abuse in New Zealand. Drug Alcohol Rev 26: 215-9

175. Seyani C, Green P, Daniel L, Pegden A (2017) An interesting case of opium tea toxicity. BMJ Case Reports 2017: bcr-2016-218971

176. Nanjayya SB, Murthy P, Chand PK et al (2010) A case of poppy tea dependence in an octogenarian lady. Drug Alcohol Rev 29: 216-8

177. Lloyd-Jones D, Bonomo Y (2006) Unusual presentations for pharmacotherapy—poppy seed dependence. Drug Alcohol Rev 25: 375-6

178. (2016) Friends recall Boulder man who died after drinking poppy tea The Denver Post Denver, CO
179. (2009) Mother warns of poppy tea 'risks' after son's overdose. BBC News
180. (2012) N.S. family warns of poppy seed tea dangers CBC News
181. (2009) CU Student Dies Brewing Poppy Tea The Denver Channel
182. Powers D, Erickson S, Swortwood MJ (2018) Quantification of Morphine, Codeine, and Thebaine in Home-Brewed Poppy Seed Tea by LC-MS/MS. J Forensic Sci 63: 1229-35. doi: 10.1111/1556-4029.13664
183. Drug Enforcement Administration. (2019) Unwashed Poppy Seed
184. Duarte D (2005) Opium and opioids: a brief history. Revista brasileira de anestesiologia 55: 135-46
185. Snyder B (2014) Revisiting old friends: update on opioid pharmacology. Issues: 1.
186. Editors Hc. (2017) Heroin, Morphine and Opiates. A&E Television Networks
187. Booth M. (2013) Opium: a history. St. Martin's Griffin
188. Foster D. (1997) DEA Tries to Weed Out Opium Poppy Seeds From Gardens. Los Angeles Times
189. Smith C. (2019) What Is the Opioid Epidemic? AddictionCenter
190. Rudd RA, Aleshire N, Zibbell JE, Matthew Gladden R (2016) Increases in drug and opioid overdose deaths—United States, 2000–2014. Am J Transplant 16: 1323-7
191. Substance Abuse and Mental Health Services Administration. (2019) Key substance use and mental health indicators in the United States: Results from the 2018 National Survey on Drug Use and Health (HHS Publication No. PEP19-5068, NSDUH Series

H-54). Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration Rockville, MD

192. Centers for Disease Control and Prevention. (2018) Opioid Overdose. Centers for Disease Control and Prevention
193. Stewart CP, Stolman A. (2013) Toxicology: mechanisms and analytical methods. Elsevier
194. Barnfield C, Burns S, Byrom DL, Kemmenoe AV (1988) The routine profiling of forensic heroin samples. Forensic Sci Int 39: 107-17
195. Cone EJ, Dickerson S, Paul BD, Mitchell JM (1993) Forensic drug testing for opiates. V. Urine testing for heroin, morphine, and codeine with commercial opiate immunoassays. J Anal Toxicol 17: 156-64
196. Gustavsson E, Andersson M, Stephanson N, Beck O (2007) Validation of direct injection electrospray LC-MS/MS for confirmation of opiates in urine drug testing. J Mass Spectrom 42: 881-9
197. Smith ML, Hughes RO, Levine B, Dickerson S, Darwin WD, Cone EJ (1995) Forensic drug testing for opiates. VI. Urine testing for hydromorphone, hydrocodone, oxymorphone, and oxycodone with commercial opiate immunoassays and gas chromatography-mass spectrometry. J Anal Toxicol 19: 18-26
198. Stefanidou M, Athanaselis S, Spiliopoulou C, Dona A, Maravelias C (2010) Biomarkers of opiate use. Int J Clin Pract 64: 1712-8
199. DEA (2018) The 2016 Heroin Signature Program Report (DEA-DCW-DIR-035-18). Drug Enforcement Agency

200. DEA (2018) The 2016 Heroin Domestic Monitor Program (DEA-DCW-DIR-026-18). Drug Enforcement Agency

201. Marciano MA, Panicker SX, Liddil GD, Lindgren D, Sweder KS. (2018) Development of a Method to Extract Opium Poppy (Papaver somniferum L.) DNA from Heroin. Sci Rep 8(1): 1-15

202. Lavania U, Srivastava S (1999) Quantitative delineation of karyotype variation in Papaver as a measure of phylogenetic differentiation and origin. Curr Sci: 429-35

203. Guo L, Winzer T, Yang X et al (2018) The opium poppy genome and morphinan production. Science (New York, NY). doi: 10.1126/science.aat4096

204. Acharya HS, Sharma V (2009) Molecular Characterization of Opium Poppy. Am J Infect Dis 5: 155-60

205. Güçlü GB, Gürkök T, Koyuncu M, Arslan N, Parmaksız İ (2014) Genetic Characterization of Turkish Commercial Opium Poppy (Papaver somniferum L.) Cultivars Using ISSR and SSR Markers. J New Results Sci 7

206. Saunders JA, Pedroni MJ, Penrose LD, Fist AJ (2001) AFLP analysis of opium poppy. Crop Sci 41: 1596-601

207. Lee EJ, Hwang IK, Kim NY et al (2010) An assessment of the utility of universal and specific genetic markers for opium poppy identification. J Forensic Sci 55: 1202-8

208. Lee EJ, Jin GN, Lee KL, Han MS, Lee YH, Yang MS (2011) Exploiting expressed sequence tag databases for the development and characterization of gene-derived simple sequence repeat markers in the opium poppy (Papaver somniferum L.) for forensic applications. J Forensic Sci 56: 1131-5

209. Şelale H, Çelik İ, Gültekin V, Allmer J, Doğanlar S, Frary A (2013) Development of EST-SSR markers for diversity and breeding studies in opium poppy. Plant Breed 132: 344-51

210. Celik I, Gultekin V, Allmer J, Doganlar S, Frary A (2014) Development of genomic simple sequence repeat markers in opium poppy by next-generation sequencing. Mol Breed 34: 323-34

211. Masárová V, Mihálik D, Kraic J (2015) In Silico Retrieving of Opium Poppy (Papaver Somniferum L.) Microsatellites. Agriculture (Polnohospodárstvo) 61: 149-56.

212. Mičianová V, Ondreičková K, Muchová D et al (2017) Forensic application of EST-derived STR markers in opium poppy. Biologia 72: 587-94

213. Parthiban S, Govindaraj P, Senthilkumar S (2018) Comparison of relative efficiency of genomic SSR and EST-SSR markers in estimating genetic diversity in sugarcane. 3 Biotech 8: 144

214. Linacre A, Gusmao L, Hecht W et al (2011) ISFG: recommendations regarding the use of non-human (animal) DNA in forensic genetic investigations. Forensic Sci Int Genet 5: 501-5

215. Kati V, Le Corre V, Michel S, Jaffrelo L, Poncet C, Délye C (2012) Isolation and Characterisation of 11 Polymorphic Microsatellite Markers in Papaver rhoeas L. (Corn Poppy), a Major Annual Plant Species from Cultivated Areas. Int J Mol Sci 14(1): 470-9

216. Sansaloni CP, Petroli CD, Carling J et al (2010) A high-density Diversity Arrays Technology (DArT) microarray for genome-wide genotyping in Eucalyptus. Plant Methods 6: 16

217. Potts BM, Vaillancourt RE, Jordan G et al (2004) Exploration of the Eucalyptus globulus gene pool

218. Lanfranco D, Dungey H (2001) Insect damage in Eucalyptus: a review of plantations in Chile. Austral Ecol 26: 477-81

219. Luo J, Arnold R, Ren S et al (2013) Veneer grades, recoveries, and values from 5-year-old eucalypt clones. Ann For Sci 70: 417-28

220. Myburg AA, Potts BM, Marques CM et al. (2007) Eucalypts. Forest Trees. Springer. pp. 115-60

221. Batish DR, Singh HP, Kohli RK, Kaur S (2008) Eucalyptus essential oil as a natural pesticide. For Ecol Manag 256: 2166-74

222. Schnitzler P, Schön K, Reichling J (2001) Antiviral activity of Australian tea tree oil and eucalyptus oil against herpes simplex virus in cell culture. Die Pharmazie 56: 343-7

223. Ghalem BR, Mohamed B (2008) Antibacterial activity of leaf essential oils of Eucalyptus globulus and Eucalyptus camaldulensis. Afr J Pharm Pharmacol 2: 211-5

224. Bachir RG, Benali M (2012) Antibacterial activity of the essential oils from the leaves of Eucalyptus globulus against Escherichia coli and Staphylococcus aureus. As Pac J Trop Biomed 2: 739-42

225. Darben T, Cominos B, Lee C (1998) Topical eucalyptus oil poisoning. Australas J Dermatol 39: 265-7

226. Dhakad AK, Pandey VV, Beg S, Rawat JM, Singh A (2018) Biological, medicinal and toxicological significance of Eucalyptus leaf essential oil: a review. J Sci Food Agr 98: 833-48

227. Hendry E, Worthington T, Conway BR, Lambert P (2009) Antimicrobial efficacy of eucalyptus oil and 1, 8-cineole alone and in combination with chlorhexidine digluconate against microorganisms grown in planktonic and biofilm cultures. J Antimicrob Chemother 64: 1219-25

228. Vilela GR, de Almeida GS, D'Arce MABR et al (2009) Activity of essential oil and its major compound, 1, 8-cineole, from Eucalyptus globulus Labill., against the storage fungi Aspergillus flavus Link and Aspergillus parasiticus Speare. J Stored Prod Res 45: 108-11

229. Dey B, Mitra A (2013) Chemo-profiling of eucalyptus and study of its hypoglycemic potential. World J Diabetes 4: 170

230. Dawoud A, Shayoub M (2015) Effects of ethanolic leaf extract of Eucalyptus camaldulensis on oral glucose tolerance test in Type-2 Model diabetic rats. J Netw Commun Emerg Technol 2

231. Basak SS, Candan F (2010) Chemical composition and in vitro antioxidant and antidiabetic activities of Eucalyptus camaldulensis Dehnh. essential oil. J Iran Chem Soc 7: 216-26

232. Ikawati Z, Wahyuono S, Maeyama K (2001) Screening of several Indonesian medicinal plants for their inhibitory effect on histamine release from RBL-2H3 cells. J Ethnopharmacol 75: 249-56

233. Juergens UR, Stöber M, Vetter H (1998) Inhibition of cytokine production and arachidonic acid metabolism by eucalyptol (1.8-cineole) in human blood monocytes in vitro. Eur J Med Res 3: 508-10

234. Otsuka H, Tsukui M, Matsuoka T, Goto M, Fujimura H (1974) Studies on anti-inflammatory agents. Anti-inflammatory screening by fertile egg method (author's transl). Yakugaku zasshi: J Pharm Soc Jpn 94: 796

235. Milhau G, Valentin A, Benoit F et al (1997) In vitro antimalarial activity of eight essential oils. J Essent Oil Res 9: 329-33

236. Sugimoto K, Suzuki J, Nakagawa K et al (2005) Eucalyptus leaf extract inhibits intestinal fructose absorption, and suppresses adiposity due to dietary sucrose in rats. Br J Nutr 93: 957-63

237. Osawa T, Ide A, Su JD, Namiki M (1987) Inhibition of in vitro lipid peroxidation by ellagic acid. J Agr Food Chem 35: 808-12

238. Zhou J, Tang F, Mao G, Shao J, Wang Y, Bian R (2003) Effect of Eucalyptus globulus oil on activation of nuclear factor-kappaB in THP-1 cells. Zhejiang da xue xue bao Yi xue ban= Journal of Zhejiang University Medical Sciences 32: 315-8, 26

239. TAKASAKI M, KONOSHIMA T, FUJITANI K et al (1990) Inhibitors of Skin-Tumor Promotion. VIII.: Inhibitory Effects of Euglobals and Their Related Compounds on Epstein-Barr Virus Activation.(1). Chem Pharm Bull 38: 2737-9

240. Benyahia S, Benayache S, Benayache F et al (2004) Isolation from eucalyptus o ccidentalis and identification of a new Kaempferol derivative that induces apoptosis in human myeloid leukemia cells. J Nat Prod 67: 527-31

241. Hong C-Z, Shellock FG (1991) Effects of a topically applied counterirritant (Eucalyptamint) on cutaneous blood flow and on skin and muscle temperatures. A placebo-controlled study. Am J Phys Med Rehabil 70: 29-33

242. Ali B, Al-Wabel NA, Shams S, Ahamad A, Khan SA, Anwar F (2015) Essential oils used in aromatherapy: A systemic review. Asian Pacific Journal of Tropical Biomedicine 5: 601-11

243. Sarkar S (1994) Capillary permeability-increasing effect of Eucalyptus hybrid leaf and a Seseli indicum seed oils in rabbit. Indian J Pharmacol 26: 55

244. Tacconi L. (2012) Illegal Logging: Law Enforcement, Livelihoods and the Timber Trade. Routledge

245. Mendoza M. (2003) Losing ground to timber thieves: Illegal logging chips away at for ests, but one court puts foot down. Associated Press, May

246. Grattapaglia D, Bradshaw Jr H (1994) Nuclear DNA content of commercially important Eucalyptus species and hybrids. Can J For Res 24: 1074-8

247. Hudson CJ, Kullan AR, Freeman JS et al (2012) High synteny and colinearity among Eucalyptus genomes revealed by high-density comparative genetic mapping. Tree Genet Genom 8: 339-52

248. Oudjehih B, Abdellah B (2006) Chromosome numbers of the 59 species of Eucalyptus L'Herit.(Myrtaceae). Caryologia 59: 207-12

249. Myburg AA, Grattapaglia D, Tuskan GA et al (2014) The genome of Eucalyptus grandis. Nature 510: 356-62

250. Butler J, Vaillancourt R, Potts B et al (2017) Comparative genomics of Eucalyptus and Corymbia reveals low rates of genome structural rearrangement. BMC Genom18: 397

251. Sumathi M, Yasodha R (2014) Microsatellite resources of Eucalyptus: current status and future perspectives. Bot Stud 55: 73

252. Byrne M, Marquezgarcia M, Uren T, Smith D, Moran G (1996) Conservation and genetic diversity of microsatellite loci in the genus Eucalyptus. Aust J Bot 44: 331-41

253. Brondani RPV, Brondani C, Tarchini R, Grattapaglia D (1998) Development, characterization and mapping of microsatellite markers in Eucalyptus grandis and E. urophylla. Theor Appl Genet 97: 816-27

254. Van der Nest M, Steenkamp E, Wingfield B, Wingfield M (2000) Development of simple sequence repeat (SSR) markers in Eucalyptus from amplified inter-simple sequence repeats (ISSR). Plant Breed 119: 433-6

255. Jones ME, Stokoe RL, Cross MJ, Scott LJ, Maguire TL, Shepherd M (2001) Isolation of microsatellite loci from spotted gum (Corymbia variegata), and cross-species amplification in Corymbia and Eucalyptus. Mol Ecology Notes 1: 276-8

256. Steane D, Vaillancourt R, Russell J, Powell W, Marshall D, Potts B (2001) Development and characterisation of microsatellite loci in Eucalyptus globulus (Myrtaceae). Silvae Genet 50: 89-91

257. Glaubitz J, Emebiri L, Moran G (2001) Dinucleotide microsatellites from Eucalyptus sieberi: inheritance, diversity, and improved scoring of single-base differences. Genome 44: 1041-5

258. Brondani R, Brondani C, Grattapaglia D (2002) Towards a genus-wide reference linkage map for Eucalyptus based exclusively on highly informative microsatellite markers. Mol Genet Genom 267: 338-47

259. Thamarus K, Groom K, Murrell J, Byrne M, Moran G (2002) A genetic linkage map for Eucalyptus globulus with candidate loci for wood, fibre, and floral traits. Theor Appl Genet 104: 379-87

260. Ottewell K, Donnellan S, Moran G, Paton D (2005) Multiplexed microsatellite markers for the genetic analysis of Eucalyptus leucoxylon (Myrtaceae) and their utility for ecological and breeding studies in other Eucalyptus species. J Hered 96: 445-51

261. Brondani RP, Williams ER, Brondani C, Grattapaglia D (2006) A microsatellite-based consensus linkage map for species of Eucalyptus and a novel set of 230 microsatellite markers for the genus. BMC Plant Biol 6: 20

262. Shepherd M, Kasem S, Lee D, Henry R (2006) Construction of microsatellite linkage maps for Corymbia. Silvae Genet 55: 228-38

263. Payn KG, Dvorak WS, Janse BJ, Myburg AA (2008) Microsatellite diversity and genetic structure of the commercially important tropical tree species Eucalyptus urophylla, endemic to seven islands in eastern Indonesia. Tree Genet Genom 4: 519-30

264. Yasodha R, Sumathi R, Chezhian P, Kavitha S, Ghosh M (2008) Eucalyptus microsatellites mined in silico: survey and evaluation. J Genet 87: 21-5

265. da Silva JM, de Sousa ACB, de Souza AP et al (2009) Development and characterization of 14 microsatellite loci from an enriched genomic library of Eucalyptus camaldulensis Dehnh. Conserv Genet Resour 1: 465

266. Sexton TR, Henry RJ, McManus LJ, Bowen S, Shepherd M (2010) Capture of assay template by multiplex PCR of long amplicons for genotyping SNPs and InDels with MALDI-TOF mass spectrometry. Mol Breed 25: 471-80

267. Faria DA, Mamani EMC, Pappas MR, Pappas Jr GJ, Grattapaglia D (2010) A selected set of EST-derived microsatellites, polymorphic and transferable across 6 species of Eucalyptus. J Hered 101: 512-20

268. Zhou C, Li F, Weng Q, Yu X, Li M, Gan S (2010) Comparison between direct sequencing and pool-cloning-based sequencing of PCR products in EST-SSR marker development in Eucalyptus. Mol Plant Breed 8: 1-10

269. Faria DA, Mamani EMC, Pappas GJ, Grattapaglia D (2011) Genotyping systems for Eucalyptus based on tetra-, penta-, and hexanucleotide repeat EST microsatellites and their use for individual fingerprinting and assignment tests. Tree Genet Genom 7: 63-77

270. Acuña C, Villalba P, García M, Hopp E, Marcucci S. (2011) Functional markers development and genetic diversity analysis in Eucalyptus globulus with emphasis in wood quality candidate genes. BMC Proceedings. BioMed Central. pp. P154

271. Acuña CV, Villalba PV, García M, Pathauer P, Hopp HE, Marcucci Poltri SN (2012) Microsatellite markers in candidate genes for wood properties and its application in functional diversity assessment in Eucalyptus globulus. Electron J Biotechnol 15: 2-.

272. He X, Wang Y, Li F et al (2012) Development of 198 novel EST-derived microsatellites in Eucalyptus (Myrtaceae). Am J Bot 99: e134-e48

273. Nevill PG, Bradbury D, Jørgensen T, Krauss S, Samaraweera S, Gardner MG (2013) Microsatellite primers identified by 454 sequencing in the floodplain tree species Eucalyptus victrix (Myrtaceae). Appl Plant Sci 1: 1200402

274. Bradbury D, Smithson A, Krauss SL (2013) Development and testing of new gene-homologous EST-SSRs for Eucalyptus gomphocephala (Myrtaceae). Appl Plant Sci 1: 1300004

275. Zhou C, He X, Li F et al (2014) Development of 240 novel EST-SSRs in Eucalyptus L'Hérit. Mol Breeding 33: 221-5

276. Grattapaglia D, Mamani EM, Silva-Junior OB, Faria DA (2015) A novel genome-wide microsatellite resource for species of Eucalyptus with linkage-to-physical correspondence on the reference genome sequence. Mol Ecol Resour 15: 437-48

277. He X, Li F, Shi J, Gan S (2011) Short Note: Seven Genomic SSRs Revealed in Eucalyptus by Re-sequencing of DNA Sequences from GenBank. Silvae Genet 60: 92-4

278. Steane D, Jones R, Vaillancourt R (2005) A set of chloroplast microsatellite primers for Eucalyptus (Myrtaceae). Mol Ecol Notes 5: 538-41

279. Acuña CV, Fernandez P, Villalba PV, García MN, Hopp HE, Poltri SNM (2012) Discovery, validation, and in silico functional characterization of EST-SSR markers in Eucalyptus globulus. Tree Genet Genom 8: 289-301

280. Bundock PC, Hayden M, Vaillancourt RE (2000) Linkage maps of Eucalyptus globulus using RAPD and microsatellite markers. Silvae genet 49: 223-32

281. Arumugasundaram S, Ghosh M, Veerasamy S, Ramasamy Y (2011) Species discrimination, population structure and linkage disequilibrium in Eucalyptus camaldulensis and Eucalyptus tereticornis using SSR markers. PLoS One 6: e28252
282. Subashini V, Shanmugapriya A, Yasodha R (2014) Hybrid purity assessment in Eucalyptus F 1 hybrids using microsatellite markers. 3 Biotech 4: 367-73
283. Neves LG, Mamani EM, Alfenas AC, Kirst M, Grattapaglia D (2011) A high-density transcript linkage map with 1,845 expressed genes positioned by microarray-based Single Feature Polymorphisms (SFP) in Eucalyptus. Bmc Genom 12: 189
284. Hudson CJ, Freeman JS, Kullan AR et al (2012) A reference linkage map for Eucalyptus. BMC Genom 13: 240
285. Petroli CD, Sansaloni CP, Carling J et al (2012) Genomic characterization of DArT markers based on high-density linkage analysis and physical mapping to the Eucalyptus genome. PLoS One 7: e44684
286. Bradbury D, Krauss SL (2013) Limited impact of fragmentation and disturbance on the mating system of tuart (Eucalyptus gomphocephala, Myrtaceae): implications for seed-source quality in ecological restoration. Aust J Bot 61: 148-60
287. Bradbury D, Smithson A, Krauss SL (2013) Signatures of diversifying selection at EST-SSR loci and association with climate in natural Eucalyptus populations. Mol Ecol 22: 5112-29
288. FBI. (2020) Quality assurance standards for forensic DNA testing laboratories
289. Shaw J, Lickey EB, Beck JT et al (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am J Bot 92: 142-66

CHAPTER II

Characterization of new chloroplast markers to determine biogeographical origin and crop type of *Cannabis sativa*

This dissertation follows the style and format of *International Journal of Legal Medicine*.

Roman, MG, Gangitano, D, Houston, R (2019) Characterization of new chloroplast markers to determine biogeographical origin and crop type of *Cannabis sativa*. Int J Legal Med 133(6):1721-1732. https://doi.org/10.1007/s00414-019-02142-w

Abstract

Marijuana (*Cannabis sativa*) is the most commonly used illicit drug in the USA. Despite its schedule I classification by the federal government, 33 states and the District of Columbia have legalized its use for medicinal or recreational purposes. This state-specific legalization has created a new problem for law enforcement: preventing and tracking the diversion of legally obtained cannabis to states where it remains illegal. In addition, trafficking of the drug at the border with Mexico remains an issue for law enforcement agencies. *C. sativa* crops can be classified as marijuana (a drug containing the psychoactive chemical delta-9-tetrahydrocannabinol) or hemp (the non-drug form of the plant). Differentiation between crop types is important for forensic purposes. In addition, investigation of trafficking routes into and within the USA requires genetic association of samples from different seizures, and determining where the crop originated could provide important leads. This project seeks to exploit sequence variations in *C. sativa* chloroplast DNA (cpDNA) to allow genetic determination of biogeographic origin, discrimination between marijuana and hemp, and association between cases for *C. sativa* samples. Due to the limited discriminatory ability of common barcoding markers, the authors sought to discover more informative polymorphic regions. By comparing published whole genome cpDNA sequences, 58 polymorphisms and seven hotspot regions were identified. Hemp samples from the USA and Canada, marijuana samples from Mexico and Chile, and medical marijuana samples from Chile were evaluated using two cpDNA hotspot regions, rpl32-trnL and trnS-trnG. Principal component analysis supported some differences between the groups based on their crop type and biogeographic origin.

Keywords: Forensic plant science, Chloroplast DNA barcoding, Haplotype, Cannabis sativa

Introduction

Cannabis sativa crops can be broadly classified as marijuana (a drug containing the psychoactive chemical delta-9-tetrahydro-cannabinol, or THC) or hemp (the non-drug form of the plant, which is cultivated for its use as oil and fiber) [1]. Marijuana is the most commonly used illicit drug in the USA [2], and as a result, it is highly trafficked, with many seizures occurring at the border with Mexico. Despite the federal government's classification of marijuana as a schedule I controlled substance, 33 states and the District of Columbia have legalized marijuana for medicinal use, with ten of these states and the District of Columbia also allowing recreational use of the drug. State-specific legalization of the drug poses a new challenge for law enforcement, which must now prevent and investigate the diversion of legal marijuana to states where it remains illegal.

Identification of marijuana usually consists of confirming the presence of THC and cystolithic hairs on the leaves of the plant [3, 4]. However, these chemical and microscopic techniques can neither give information about the biogeographical origin of the sample nor be used to link seized samples. Since the 1990s, researchers have been interested in developing molecular techniques for identifying and individualizing *C. sativa* samples in order to demonstrate associations between samples. These techniques include Random Amplified Polymorphic DNA (RAPD) [5, 6], Amplified Fragment Length Polymorphism (AFLPs) [7], Inter Simple Sequence Repeats (ISSRs) [5, 8], barcoding regions in the chloroplast and mitochondrial genomes [9–16], and simple sequence repeats (SSRs), more commonly referred to as short tandem repeats (STRs) in the forensic community [17–23].

Barcoding is the use of specific regions of the genome to identify differences in the genetic sequence of individuals. It can be used for identification of species (by exploiting

inter-specific differences) or for distinguishing between populations of the same species (by exploiting intra-specific differences). Several researchers have developed universal primer sets for plant barcoding studies [24–31], but no single marker has shown universal utility in either identifying species or in distinguishing between populations [32–36].

Because the mitochondrial and chloroplast genomes are uni-parentally inherited (without recombination), most plant barcoding research has focused on these genomes. Both mitochondrial and chloroplast genomes of *C. sativa* are maternally inherited [37]. Advantages of using chloroplast DNA (cpDNA) over mitochondrial DNA (mtDNA) include its faster mutation rate and conserved size, structure, and gene content [26, 36]. Researchers have demonstrated that certain barcoding markers (*ITS1*, *ITS2*, and *trnL-trnF*) have sequences which are unique to

C. sativa plants [10, 12–15]. Though these conserved markers are useful for identifying a plant substance as *C. sativa*, they cannot be used to distinguish between samples.

Other barcoding markers show promise in distinguishing between *C. sativa* crop types and geographic origin [16]. Australian researchers discovered polymorphisms across five cpDNA (*trnL-trnF*, *rbcL-orf106*, *ccmp2*, *ccmp6*, and *trnH-trnK*) and two mtDNA (*nad* 4 exon 3 to exon 4 and *nad* 5 exon 4 to exon 5) loci which varied among hemp and marijuana samples from different geographic origins. Haplotypes consisted of alleles at each of the seven loci, and phylogenetic and parsimony analyses indicated that six major haplogroups existed, with samples differentially distributed according to both crop type (fiber, drug, and wild types) and geographic origin [16]. A subsequent study sought to build a database of organelle and autosomal *C. sativa* genotypes and to discern population structure among related and unrelated sample groups [38]. This study used the same seven

organelle markers as Gilmore et al. [16] in addition to a 13-loci autosomal STR multiplex assay [23]. Both sets of markers showed some ability to differentiate between samples from different biogeographical origins as well as crop types (marijuana versus hemp). This study also expanded the geographic range of Gilmore et al.'s study by including regions of North and South America [38]. However, these markers have not been able to definitively determine the geographic origin or crop type of samples due to limited discrimination of the organelle markers and the lack of a comprehensive worldwide *C. sativa* database.

Other chloroplast markers have also demonstrated a limited ability to determine the geographic origin of *C. sativa* samples [9, 11, 39]. Kohjyouma et al. [11] showed that the intergenic spacer region between the *trnL* and *trnF* genes (*trnL-trnF*) demonstrated variations between populations of *C. sativa*, and the authors predicted that the plant had developed local variations which would show differences among various barcoding regions. The same region showed no variation between samples seized in Brazil compared with published sequences of *C. sativa* from other countries [10]. Mello et al. [9] identified a 561-bp segment of the *rbcL* gene as a potentially informative tool for distinguishing Brazilian samples from those seized in China, the UK, or the USA. Clearly, no single marker can definitively identify the crop type and biogeographic origin, and the development of novel lineage markers in *C. sativa* would give some insight into this problem. Recently, the chloroplast genome of *C. sativa* has been sequenced and mapped [40, 41], which will greatly assist future genomic studies.

This research project seeks to discover novel informative sequence variations (single nucleotide polymorphisms, SNPs; homopolymeric STRs, hSTRs; and insertions/deletions, INDELs) in the chloroplast genome of *C. sativa*, in order to assist in

determining the biogeographical origin of marijuana samples and distinguish them from hemp. Samples tested include hemp (fiber-type) specimens from the USA and Canada, marijuana (drug-type) specimens from Mexico and Chile, and medical marijuana specimens from Chile. Differences in sequences between these groups could assist investigators in identifying licit or illicit crops as well as provide leads and allow law enforcement agencies, such as U.S. Customs and Border Protection, to associate cases involving *C. sativa*, including illegal trafficking of the drug both into and within the USA.

Materials and methods

In silico analysis of published genomes

Previously published sequences for the chloroplast genome of six *Cannabis sativa* cultivars were identified using the GenBank® database (NCBI): Carmagnola (NC_026562), Purple Kush (AGQN01337109), Cheungsam (KR184827), Yoruba Nigeria (NC_027223), *Cannabis sativa* (KY084475), and Dagestani (KR779995). The FASTA files were imported into Geneious Pro Software R7.1.9 (Biomatters, Auckland, New Zealand) and aligned using the Mauve genome alignment tool [42]. Polymorphic sites (hSTRs, INDELs, and SNPs) were identified where the sequences differed, and hotspot regions were defined as sequences of less than 1600 bp containing at least three polymorphisms. The authors chose to focus on analysis of two hotspot regions: *trnS-trnG* and *rpl32-trnL*.

Sample collection

Marijuana (THC-positive) plant material from seizures was obtained from U.S. Customs and Border Protection (CBP) (N = 422; from 21 cases). Marijuana DNA extracts from the Araucania region of southern Chile (N = 50; from 10 cases) were provided by the Policia de

Investigaciones in Chile. Additionally, DNA extracts of four medical marijuana strains (Amnesia, AK, Lemon Haze, and London Cheese) were provided by collaborators from Chile ($N = 3$ each). Hemp seeds were purchased from four companies: Manitoba Harvest (Winnipeg, MB, CA) ($N = 15$ from two bags), Badia Spices Inc. (Doral, FL, USA) ($N = 15$), Navitas™ Organics (Novato, CA, USA) (N = 15), and American Hemp Harvest (Boulder, CO, USA) ($N = 10$). Additionally, two strains (Electra and Lifter) of dried CBD hemp flowers from The Original Hemp Buds (OR, USA, or NY, USA) were purchased in Houston, TX. American Hemp Harvest states that its hemp is grown and harvested in CO, USA; The Original Hemp Buds grows hemp in Oregon and upstate New York; and the other three companies source hemp from Canada.

DNA extraction

Plant material (leaves, flowers, stems, or seeds) was homogenized in liquid nitrogen and extracted using the DNeasy® Plant Mini Kit (QIAGEN, Hilden, Germany) according to the manufacturer's protocol [43]. DNA was extracted onsite for the THC-containing samples provided by CBP (Mexican-US marijuana samples). DNA extracts were provided by the Policia de Investigaciones in Chile (Chilean marijuana and medical marijuana samples). For the hemp, DNA was extracted from randomly selected single seeds or approximately 1-cm^2 sections of dried plant material (bud, stem, or leaf), and extractions were performed at Sam Houston State University.

DNA quantification

DNA quantification was performed using a quantitative real-time PCR (qPCR) method developed and validated by Houston et al. [38] which utilizes primers specific to a sequence in *C. sativa* cpDNA, Cscp001 [16] (Integrated DNA Technologies, Coralville, IA, USA).

Sanger sequencing

Sanger sequencing was used to assess variability of each of the chosen polymorphic sites in at least ten samples from Canadian hemp, USA hemp, Chilean marijuana, USA-Mexican marijuana, and Chilean medical marijuana. Sequencing was performed using the BigDye™ Direct Cycle Sequencing Kit (Thermo Fisher Scientific, South San Francisco, CA, USA). Primer design was accomplished using Primer3 software [44], and AutoDimer software [45] was used to assess primer-primer interactions. An M13 tag was added to the primers in accordance with the BigDye™ Direct Cycle Sequencing Kit User's Manual [46]. Sequences of the primers are listed in Table 2.1. Annealing temperatures were determined as described in the "Annealing temperature determination" section, below.

PCR amplification and cycle sequencing were performed on the Veriti™ Fast Thermal Cycler (Thermo Fisher Scientific) according to the manufacturer's protocol [46] with the exception of using the optimal annealing temperature determined for each primer set (Table 2.1). Sequencing products were purified using Centri-Sep™ Spin Columns (Thermo Fisher Scientific), then reconstituted in 12 µL HiDi™ formamide (Thermo Fisher Scientific) and run on a 3500 Genetic Analyzer (Thermo Fisher Scientific) using the following conditions: oven 60 °C; prerun 18 kV, 60 s; injection 1.6 kV, 8 s; run 19.5 kV, 1020 s; capillary length 50 cm; polymer POP-7™; and dye set Z. Sequences were aligned and proofread using Geneious Pro Software R7.1.9. Allele nomenclature followed ISFG guidelines and allele sequences were reported based on the forward strand [47–49]. Sequences were submitted to GenBank® (accession numbers MK989685-MK989708).

Table 2.1 Primer sequences used for Sanger sequencing and experimentally-determined optimal annealing temperatures

Primer name	Sequence	Annealing temperature (°C)
rpl32-trnL hSTR 1 F M13	TGT AAA ACG ACG GCC AGT TAT TGG GCA GCA TTA AAA GC	61
rpl32-trnL hSTR 1 R M13	CAG GAA ACA GCT ATG ACC TTT CTA ATG TCT TTC GAA GTT	
rpl32-trnL hSTR 2 & SNP F M13	TGT AAA ACG ACG GCC AGT ATC TTT GTT GTT CTG ACT CG	61
rpl32-trnL hSTR 2 & SNP R M13	CAG GAA ACA GCT ATG ACC TGA TCA GTT CAA AAC AAA AC	
rpl32-trnL hSTR 3 F M13	TGT AAA ACG ACG GCC AGT GTT TTA TGT TTT GTT TTG AAC TG	63
rpl32-trnL hSTR 3 R M13	CAG GAA ACA GCT ATG ACC GGA TTC TTA TTT TCC CCA TCC T	
rpl32-trnL INDEL F M13	TGT AAA ACG ACG GCC AGT TCC TTT GTC TAC TCT TTT GAA	61
rpl32-trnL INDEL R M13	CAG GAA ACA GCT ATG ACC TTA ATG AGT TTC AAC GAC CT	
trnS-trnG hSTR 1 F M13	TGT AAA ACG ACG GCC AGT CCG GGC CTC TTT TAT TCC AA	63
trnS-trnG hSTR 1 R M13	CAG GAA ACA GCT ATG ACC ACG TGT TGG TGT ATT ATA AAG T	
trnS-trnG hSTR 2 F M13	TGT AAA ACG ACG GCC AGT ATA ATA CAC CAA CAC GTT TT	61
trnS-trnG hSTR 2 R M13	CAG GAA ACA GCT ATG ACC AGT CTT TTG CTA GCG GTT TT	
trnS-trnG SNPs 1-2 F M13	TGT AAA ACG ACG GCC AGT GTC CAC TCA GCC ATC TCT CC	64
trnS-trnG SNPs 1-2 R M13	CAG GAA ACA GCT ATG ACC TTG GAA TAA AAG AGG CCC GG	
trnS-trnG SNP 3 F M13	TGT AAA ACG ACG GCC AGT AAA ACC GCT AGC AAA AGA CT	66
trnS-trnG SNP 3 R M13	CAG GAA ACA GCT ATG ACC GTC GAA CAA GCA ACT CAG GTG	

Annealing temperature determination

Annealing temperature determination for all primer sets was performed by gradient PCR using the Type-it Microsatellite PCR Kit (QIAGEN) and an Eppendorf Mastercycler® Gradient (Eppendorf, Hamburg, Germany) thermal cycler. PCR reactions consisted of 6.25 μL of Type-it Microsatellite PCR Master Mix (QIAGEN), 1.25 μL of 2 μM primer mix (Integrated DNA Technologies), 1.25 μL of 5× Q solution (QIAGEN), 1.75 μL of deionized nuclease-free water, and 2 μL of DNA extract for a total reaction volume of 12.5 μL. Thermal cycling conditions consisted of an initial activation for 5 min at 95 °C, followed by 30 cycles of the following: 30 s at 94 °C, 90 s at a gradient of 60 °C ± 10 °C across 12 wells, and 30 s at 72 °C; and a final extension of 30 min at 60 °C. PCR products were separated by gel electrophoresis on a 2% agarose gel with SYBR™ Safe DNA Gel Stain (Thermo Fisher Scientific) and visualized under UV light. The highest temperature that produced a bright band was selected as the optimal annealing temperature for each primer set.

Fragment analysis development and genotyping

Fragment analysis assays were designed for easier genotyping of hSTR and INDEL markers. Primer design was accomplished using Primer3 software [44], and AutoDimer software [45] was used to assess primer-primer interactions. Forward primers were labeled with the 6-FAM™ fluorescent dye for detection via capillary electrophoresis. Optimal annealing temperature determinations were performed by gradient PCR as described in the "Annealing temperature determination" section. Due to the A/T-rich nature of the chloroplast genome and the frequency of homopolymeric regions, primer design was limited; however, multiplex assays were developed when possible.

For sample genotyping, PCR reactions targeting 80 pg of cpDNA consisted of 6.25 μL of Type-it Microsatellite PCR Master Mix (QIAGEN), 1.25 μL of primer mix (Integrated DNA Technologies) (Table 2.2), 1.25 μL of 5× Q solution (QIAGEN), 1.75 μL of deionized nuclease-free water, and 2 μL of DNA extract for a total reaction volume of 12.5 μL. Thermal cycling conditions on a Veriti™ 96-well Thermal Cycler (Thermo Fisher Scientific) consisted of an initial activation for 5 min at 95 °C, followed by 30 cycles of the following: 30 s at 95 °C, 90 s at the optimal annealing temperature (Table 2.2), and 30 s at 72 °C; and a final extension of 30 min at 60 °C.

Capillary electrophoresis was performed using a 3500 Genetic Analyzer (Thermo Fisher Scientific) with the following run conditions: oven 60 °C; prerun 15 kV, 180 s; injection 1.6 kV, 8 s; run 19.5 kV, 1330 s; capillary length 50 cm; polymer POP-7™; and dye set G5. An aliquot of 0.5 μL of amplified product was added to 9 μL HiDi™ Formamide (Thermo Fisher Scientific) and 0.5 μL LIZ 600 (Thermo Fisher Scientific). A custom panel and bins were designed for data analysis on GeneMapper ID v.5 software (Thermo Fisher Scientific). An analytical threshold of 100 RFU was used.

Following development and optimization of the fragment analysis assays, approximately every fourth sample in the larger database (N = 152) was chosen for genotyping to produce a representative data set for the *C. sativa* cpDNA database.

Table 2.2 Primer sequences and annealing temperatures for fragment analysis of hSTR and INDEL markers

	Primer name	Sequence	Concentration of primer pair in PCR (μM)	Annealing temperature (°C)
Reaction 1	*rpl32-trnL* hSTR 1 F (6-FAM) *rpl32-trnL* hSTR 1 R	/56-FAM/TA TTG GGC AGC ATT AAA AGC TTT CTA ATG TCT TTC GAA GTT	0.08	58
	rpl32-trnL hSTR 3 F (6-FAM) *rpl32-trnL* hSTR 3 R	/56-FAM/GT TTT ATG TTT TGT TTT GAA CTG GGA TTC TTA TTT TCC CCA TCC T	0.08	58
Reaction 2	*rpl32-trnL* hSTR 2 F (6-FAM) *rpl32-trnL* hSTR 2 R	/56-FAM/AA TTT ACA ACT CGA AAA CTT C TGA TCA GTT CAA AAC AAA AC	0.08	58
	rpl32-trnL INDEL F (6-FAM) *rpl32-trnL* INDEL R	/56-FAM/TC CTT TGT CTA CTC TTT TGA A TTA ATG AGT TTC AAC GAC CT	0.08	58
Reaction 3	*trnS-trnG* hSTR 1 F (6-FAM) *trnS-trnG* hSTR 1 R	/56-FAM/CC GGG CCT CTT TTA TTC CAA ACG TGT TGG TGT ATT ATA AAG T	0.04	60

Allelic ladder design

Data from the initial screening of samples via Sanger sequencing were used to design allelic ladders. Following analysis of the sequences, samples with different known alleles for each locus were used to create allelic ladders as reported in previous studies [22, 50]. Samples were amplified in singleplex, then the amplification products were combined and run on CE to ensure peak height balance. The ladders for all loci in each fragment analysis assay were then combined and balanced to create the final allelic ladder for that assay. The final, balanced ladder was then diluted 1:1000 in TE to create a stock, which was re-amplified using the multiplex assay.

SNaPshot™ assay development and genotyping

Primers were designed using Primer3 software [44] with default settings, and optimal annealing temperatures were determined as described above. To avoid primer interactions, separate SNaPshot™ assays were designed for each hotspot region. Amplification of each hotspot region was carried out using an Eppendorf MasterCycler Gradient Thermal Cycler (Eppendorf). PCR reactions consisted of 6.25 μL of Type-it Microsatellite PCR Master Mix (QIAGEN), 1.25 μL of primer mix (Integrated DNA Technologies) (Table 2.3), 1.25 μL of 5× Q solution (QIAGEN), 1.75 μL of deionized nuclease-free water, and 2 μL of DNA extract (80 pg) for a total reaction volume of 12.5 μL. Thermal cycling conditions consisted of an initial activation for 5 min at 95 °C, followed by 30 cycles of the following: 30 s at 95 °C, 90 s at the optimal annealing temperature (Table 2.3), and 30 s at 72 °C; and a final extension of 30 min at 60 °C. Products were purified to remove unincorporated deoxynucleotides and primers by adding 5 μL calf intestinal alkaline phosphatase (CIAP) (1 U/μL, Promega Corporation, Madison, WI, USA) and 2 μL

exonuclease I (10 U/μL, Invitrogen) and incubating at 37 °C for 1.5 h and 75 °C for 30 min.

Single-base extension (SBE) was achieved using the SNaPshot™ Multiplex Kit (Thermo Fisher Scientific) as per the manufacturer's instructions [51] with the exception of using half reaction volumes. Primers (Table 2.3) were designed to anneal to the 20 bp sequence immediately preceding the SNP site on either the forward or reverse DNA strand. For multiplexing, neutral sequences of different lengths were added to the SBE primers to assist in spacing of the markers during capillary electrophoresis [52]. The sequences were analyzed with AutoDimer [45] to avoid primer-primer interactions and hairpin structures. Following single-base extension on an Eppendorf Mastercycler Gradient Thermal Cycler (Eppendorf), 1 μL of CIAP was added to the products, followed by incubation at 37 °C for 1.5 h and 75 °C for 30 min.

Capillary electrophoresis was performed on a 3500 Genetic Analyzer (Thermo Fisher Scientific) using the following run conditions: oven 60 °C; prerun 15 kV, 180 s; injection 1.6 kV, 8 s; run 15 kV, 560 s; capillary length 36 cm; polymer POP-4™; and dye set E5. A custom panel and bins were created to analyze the genotypes using Genemapper ID v.5 software (Thermo Fisher Scientific). An analytical threshold of 100 RFU was used.

Following development and optimization of the SNaPshot™ assays, approximately every fourth sample from the larger database (N = 152) was chosen for genotyping to produce a representative data set for the *C. sativa* cpDNA database.

Table 2.3 Primer sequences for SNaPshot™ analysis. Neutral sequences used for spacing of markers are indicated by lowercase letters

Primer name	Sequence	SBE primer length	Concentration of primer in PCR (μM)	Annealing temperature (°C)
rpl32-trnL SNP F	ATC TTT GTT GTT CTG ACT CG	–	0.2	60
rpl32-trnL SNP R	TGA TCA GTT CAA AAC AAA AC			
rpl32-trnL SNP SBE R	AAT ATA AAT TTG TCC TTT TA	20 bp	0.2	–
trnS-trnG SNPs F	GTC CAC TCA GCC ATC TCT CC	–	0.2	66
trnS-trnG SNPs R	GTC GAA CAA GCA ACT CAG GTG			
trnS-trnG SNP1 SBE F	tga aag tct gac aaC CCC CAA TTG AAA AAA AAA A	34 bp	0.2	–
trnS-trnG SNP2 SBE R	ggt gcc acg tcg tga aag tct gac aaT TTC TAT ATT GAA AAA AAA A	46 bp	2.0	–
trnS-trnG SNP3 SBE F	TCT TTT TTG ATC CTA TTT TT	20 bp	0.2	–

Statistical analysis

Genetic Data Analysis (GDA) software [53] was used to calculate genetic distance using the neighbor-joining method with coancestry distance. In addition, the Arlequin v. 3.5 software was used to perform pairwise comparisons and test for statistically significant differences among the five populations using F_{ST} as genetic distance [54]. Principal component analysis was then performed for visualization using Past3 v.3 software [55].

Results and discussion

In silico analysis of published genomes

In comparing chloroplast genome sequences of six *C. sativa* cultivars, 58 polymorphisms were discovered. Of these, 31 were hSTRs, 23 were SNPs, and four were INDELs (Table 2.4). Seven hotspot regions with at least three polymorphisms within 1600 bp were identified: *trnK-matK-trnK*, *rps16*, *trnS-trnG*, *ycf3*, *accD-psaI*, *clpP*, and *rpl32-trnL*. The *trnK-matK-trnK* region consisted of four polymorphisms (two hSTRs, a SNP, and an INDEL) within 1130 bp; the *rps16* region contained four polymorphisms (an hSTR and three SNPs) within 800 bp; the *ycf3* region had three hSTRs in 1580 bp; the *accD-psaI* region had three polymorphisms (an hSTR and two SNPs) in 308 bp; the *clpP* region had five polymorphisms (four hSTRs and an INDEL) in 1118 bp; the *trnS-trnG* region had five polymorphisms (two hSTRs and three SNPs) in 765 bp; and the *rpl32-trnL* region had five polymorphisms (three hSTRs, a SNP, and an INDEL) in 417 bp.

The *matK* gene and its surrounding regions (*trnK/matK* and *trnK-matK-trnK*) were first explored in 1994 [56, 57] and have since been extensively studied in other plants [33, 58], but this is the first study exploring this region in the *Cannabaceae* family. A study by Shaw et al. [33] evaluating the discriminatory power of plant barcoding markers ranked the

trnK/matK region as a tier 3 (least polymorphic information content, PIC) region for plant barcoding studies. The *rps16* region is a commonly used barcoding marker [33] and was first used in phylogenetic studies by Oxelman et al. [59]. Shaw et al. [33] ranked it as a tier 2 region but noted that it is not variable enough for resolving infrageneric relationships. The *ycf3* region contains an hSTR (cscp004) previously evaluated by Gilmore et al. [16] for its utility in distinguishing crop type and geographic origin in *C. sativa* as well as two other hSTRs not previously reported. Despite its variability, this region is not in common use. The *accD-psaI* intergenic spacer has been previously reported to be moderately polymorphic in other plant species [34, 58, 60–63]. Although *clpP* has been reported to be polymorphic, it has rarely been used in phylogenetic studies [58, 64] and has never been studied in *Cannabaceae*.

The *trnS-trnG* intergenic spacer region, often coamplified with the *trnG* intron (*trnS-trnG-trnG*), was ranked as a tier 1 region by Shaw et al. [33], who noted that the region is gaining use in phylogenetic studies. The region was first explored by Hamilton in *Corythophora* [65, 66]. It has been shown to be highly polymorphic by several other studies [67-69] but has not been tested in *Cannabaceae*.

The *rpl32-trnL* region has been identified by Shaw et al. [34] as one of the most promising and variable regions in the chloroplast genome. It has also been previously reported as a highly polymorphic region by several studies [58, 70–73] but has not been previously studied in *Cannabaceae*.

Based on their size and polymorphic content, the authors chose to analyze *trnS-trnG* and *rpl32-trnL* as the two most promising hotspot regions in *C. sativa* chloroplast DNA. The two hotspot regions have a combined total of ten polymorphisms, and the

genotypes observed in the six cultivars are listed in Table 2.5. The polymorphic nature of these regions makes them promising regions of interest for discriminating between different populations of *C. sativa* to identify crop type and geographic origin. Further study is planned to analyze the usefulness of the other hotspot regions identified.

Sanger sequencing

Sanger sequencing of at least ten samples was performed to assess variability of the ten polymorphisms within the *rpl32-trnL* and *trnS-trnG* regions. This initial screening showed variable genotypes in nine of the ten polymorphisms (Table 2.6), with *trnS-trnG* hSTR2 being monomorphic for a 10-T repeat in all 16 samples tested. As a result, this locus was removed from the study. Detailed sequencing results are displayed in Fig. 2.1-2.7. The most polymorphic locus was *rpl32-trnL* hSTR3 with seven alleles, which varied in sequence as well as length (Table 2.6). However, using a size-based genotyping method, only three alleles were expected due to similar migration of sequence variation alleles. To note, it is the authors' intention to develop a massively parallel sequencing (MPS) method for future genotyping of hotspot regions, which would allow distinction of these sequence-variable alleles. All other loci were bi-allelic upon initial screening. Additionally, two new SNPs were observed: one preceding *rpl32-trnL* hSTR2 (A/C) (Fig. 2.2) and the other occurring before *trnS-trnG* hSTR1 (A/T) (Fig. 2.5). The SNP preceding *rpl32-trnL* hSTR2 was a C in all samples sequenced, compared with an A in the six reference sequences from GenBank® (NC_026562, AGQN01337109, KR184827, NC_027223, KY084475, and KR779995). These two SNPs were not genotyped in this study.

Table 2.4 Polymorphisms in *C. sativa* chloroplast genome. Full chloroplast genome sequences for six cultivars were aligned, and 58 polymorphisms were detected

Start location in consensus sequence (bp)	Gene name	Polymorphism type	Carmagnola NC_026562	Cheungsam KR184827	Cannabis sativa KY084475	Dagastani KR779995	Yoruba Nigeria NC_027223	Purple Kush AGQN01337 109
124	*trnH-psbA*	SNP (G/T)	T	G	T	G	G	T
2988	*trnK*	Indel (GAATAC)	present	absent	present	present	absent	present
3267	*trnK*	SNP (C/T)	C	T	C	C	C	C
3818	*trnK*	Homopolymer (A)	11	10	11	11	11	11
4118	*trnK/ matK*	Homopolymer (T)	10	11	9	10	10	10
5207	*trnK-rps16*	Homopolymer (A)	9	9	10	9	9	10
5314	*rps16*	SNP (A/G)	A	G	A	A	A	A
5528	*rps16*	SNP (A/C)	A	C	A	A	A	A
5529	*rps16*	Homopolymer (C)	11	12 (13 with SNP)	11	11	12	11
6114	*rps16*	SNP (G/A)	A	G	G	G	G	A
7096	*rps16 -trnQ*	Homopolymer (A)	10	9	10	10	10	10
8370	*psbI-trnS*	Homopolymer (T)	8	7	8	7	7	8

(continued)

Start location in consensus sequence (bp)	Gene name	Polymorphism type	Carmagnola NC_026562	Cheungsam KR184827	Cannabis sativa KY084475	Dagastani KR779995	Yoruba Nigeria NC_027223	Purple Kush AGQN01337109
8607	*trnS-trnG*	SNP (T/C)	C	T	T	T	T	C
8696	*trnS-trnG*	SNP (A/T)	T	A	T	A	A	T
9030	*trnS-trnG*	Homopolymer (T)	12	11	TTA(9T)	12	12	12
9032	*trnS-trnG*	SNP (T/A)	T	T	A	T	T	T
9116	*trnS-trnG*	Homopolymer (T)	11	11	10	11	11	11
9372	*trnS-trnG*	SNP (T/A)	T	T	T	A	A	T
10551	*trnG-trnR*	Homopolymer	(4A)G(6A)	6A	(5A)G(6A)	(4A)G(6A)	A	(5A)G(6A)
14678	*atpH-atpI*	Homopolymer (T)	11	13	15	11	11	14
16500	*rps2-rpoC2*	Homopolymer (T)	14	14	14	13	13	14
16550	*rps2-rpoC2*	Homopolymer (A)	7	8	7	8	8	7
22571	*rpoC1*	SNP (C/T)	C	T	C	C	C	C
22892	*rpoC1*	Homopolymer (T)	8	8	9	8	8	8

(continued)

Start location in consensus sequence (bp)	Gene name	Polymorphism type	Carmagnola NC_026562	Cheungsam KR184827	Cannabis sativa KY084475	Dagastani KR779995	Yoruba Nigeria NC_027223	Purple Kush AGQN01337109
36499	*trnS-lhbA/psbZ*	Homopolymer (T)	10	9	10	9	9	10
39630	*psaB*	SNP (T/G)	G	T	G	T	T	G
43480	*ycf3*	Homopolymer (T)	11	11	12	11	12	11
44033	*ycf3*	Homopolymer (T)	11	11	12	11	11	12
45060	*ycf3*	Homopolymer (A)	10	9	10	10	10	10
49760	*ndhJ-ndhK*	Homopolymer (T)	10	11	10	10	10	10
57175	*rbcL-accD*	Homopolymer	(11A)(4T)	(11A)(5T)	(11A)(4T)	(11A)(4T)	(12A)(4T)	(11A)(4T)
58861	*accD-psaI*	SNP (A/G)	G	A	G	A	A	G
59009	*accD-psaI*	SNP (T/G)	G	T	G	T	T	G
59169	*accD-psaI*	Homopolymer (A)	10	10	11	10	10	11
63645	*petA-psbJ*	SNP (C/A)	C	C	C	A	A	C
64999	*psbL-psbF*	SNP (T/G)	G	T	G	T	T	G

(continued)

Start location in consensus sequence (bp)	Gene name	Polymorphism type	Carmagnola NC_026562	Cheungsam KR184827	Cannabis sativa KY084475	Dagastani KR779995	Yoruba Nigeria NC_027223	Purple Kush AGQN01337109
69406	*rpl20-rps12*	Homopolymer (T)	13	13	14	13	13	14
70942	*clpP*	Homopolymer (A)	10	10	10	10	11	10
71012	*clpP*	Homopolymer	4T	4T	4T	4T	TATTT	4T
71693	*clpP*	Homopolymer (T)	15	14	15	15	14	15
72047	*clpP*	Homopolymer (T)	12	13	13	12	13	11
72060	*clpP*	Indel (TTCAATTTA)	absent	absent	present	absent	absent	present
78222	*petD-rpoA*	Homopolymer (T)	10	9	9	10	9	10
79722	*rps11*	SNP (T/G)	G	T	G	T	T	G
80148	*rpl36-rps8*	SNP (T/C)	T	T	T	C	C	T
82462	*rpl16-rps3*	SNP (C/T)	C	T	C	C	C	C
82562	*rpl16*	Homopolymer (T)	11	10	11	11	11	11
110335	*ndhF*	SNP (T/G)	G	T	T	T	T	G

(continued)

Start location in consensus sequence (bp)	Gene name	Polymorphism type	Carmagnola NC_026562	Cheungsam KR184827	Cannabis sativa KY084475	Dagastani KR779995	Yoruba Nigeria NC_027223	Purple Kush AGQN01337109
112594	*ndhF-rpl32*	INDEL ((11A)GAATTGA)	present (11A)GAATTGA	absent (with AA insert)	present (16A)TTAA	present (11A)GAATTGA	absent	present (11A)GAATTG, missing AA at end
112889	*rpl32-trnL*	Homopolymer (A)	11	12	12	11	12	11
113020	*rpl32-trnL*	Homopolymer (A)	11	12	11	11	11	10
113104	*rpl32-trnL*	SNP (A/C)	C	A	C	A	A	C
113209	*rpl32-trnL*	Homopolymer	6A	6T	7T	6A	(7T)A	5T
113306	*rpl32-trnL*	Indel (AATAAA)	absent	present	absent	absent	present	absent
115034	*ndhD*	Homopolymer (A)	8	11	8	8	11	8
116326	*ndhD*	SNP (T/A)	A	T	T	T	T	A
123192	*rps15-ycf1*	SNP (C/T)	T	C	T	C	C	T
130028	*trnR-5S rRNA*	SNP (C/G)	G	C	C	G	C	C

Table 2.5 Chloroplast genotypes of the six GenBank *C. sativa* cultivars

Accession	Strain	Type	Geographic origin	*rpl32-trnL*					*trnS-trnG*				
				hSTR 1	hSTR 2	hSTR 3	INDEL	SNP	hSTR 1	hSTR 2	SNP 1	SNP 2	SNP 3
NC_026562.1	Carmagnola	Hemp	Italy	11A	11A	6A	Absent	C	12T	11T	C	T	T
KY084475.1	C. Sativa	Hemp	China	12A	11A	7T	Absent	C	11T	11T	T	A	T
AGQN01337109.1	Purple Kush	Medical Marijuana	California	11A	10A	5T	Absent	C	12T	11T	T	A	A
KR184827.1	Cheungsam	Hemp	Korean	12A	12A	6T	Present AATAAA	A	12T	11T	T	A	A
NC_027223.1	Yoruba Nigeria	Hemp	Nigeria	12A	11A	T(7A)	Present AATAAA	A	TTA(9T)	10T	T	T	T
KR779995.1	Dagestani	Hemp	Russia	11A	11A	6A	Absent	A	12T	11T	C	T	T

Table 2.6 Sanger sequencing results showing variability of alleles

Polymorphism	Motif	Canadian hemp (N=4-5)	USA hemp (N=1)	USA-Mexico marijuana (N=3-6)	Chile marijuana (N=2)	Chile medical marijuana (N=4)
rpl32-rnL hSTR1	A Homopolymer	12A	12A	11A	11A	Not sequenced
rpl32-trnL hSTR2	A homopolymer	11A	12A	11A	11A	Not sequenced
rpl32-trnL hSTR3	A or T homopolymer	8T, 8T+A, 7T+A	6A, 3T+3A, 2T+4A	6A, T+5A	6A	6A
rpl32-trnL SNP	A/C SNP	A	A	C	C	Not sequenced
rpl32-trnL INDEL	AATAAA	Absent	Present	Present	Present	Not sequenced
trnS-trnG hSTR1	Variable	TTA(12T)	TTA(12T)	TTA(12T)	TTA(12T), TTATTA(10T)	TTA(12T)
trnS-trnG hSTR2	T homopolymer	10T	10T	10T	10T	10T
trnS-trnG SNP1	T/C SNP	T	C	C	C	C
trnS-trnG SNP2	A/T SNP	A	T	T	T	T
trnS-trnG SNP3	A/T SNP	A	T	T	T	T

This is a note for Fig. 2.1 through 2.7. In the consensus sequences, the forward and reverse primer binding locations are underlined and the single base extension (SBE) primer binding sites are highlighted.** The location of hSTR repeats are indicated in the consensus sequences as **[REPEAT]**, indels are indicated as **[INDEL]**, and SNPs are indicated by their **nucleotide ambiguity code.** The GenBank accession numbers are also referenced in the table. N refers to the total number of samples with the indicated haplotype.

TATTGGGCAGCATTAAAAGCTTTTTCTTTAGGTAAATCTCTTTCTACAAGAAATT
CAAAAAGTTTCTTTTATACGCCAAATAAATAATACAAAATTGGCATAATCTTTGT
TGTTCTGACTCG**[REPEAT]**TAAGCATAATTTTTAATTTACAACTCGAAAACTTC
GAAAGACATTAGAAA

Allele	[REPEAT]	N	Genbank accesion number
	A		
10	10	3	MK989685
11	11	7	MK989686
12	12	5	MK989687

Fig. 2.1 Consensus sequence of *rpl32-trnL* hSTR1 locus, haplotypes found and allele nomenclature proposal

ATCTTTGTTGTTCTGACTCG**[hSTR1]**TAAGCATAATTTTTAATTTACAACTCGAA
AACTTCGAAAGACATTAGAAAAAAATTCTAAGAATTATAATAATACTTC**M**CTTT
ATTTATATATTTATTCCTTTCCTTTTAAAATAGAAT**[REPEAT]**CGACAAATTCTAT
TAGATAGAAATCTATAAAAAAATATAAAAACTAAAAAAAAAATATATGTTCAAA
GATTT**M**TAAAAGGACAAATTTATATTCTATTTATATTCTTTAATGTTTTATGTTTT
GTTTTGAACTGATCA

hSTR2 allele	Preceding SNP	[REPEAT]	SNP allele		N	Genbank accesion number
	M (forward)	A	M (forward)	K (reverse)		
11	C	11	A	T	4	MK989688
11	C	11	C	G	5	MK989689
12	C	12	A	T	1	MK989690

Fig. 2.2 Consensus sequence of *rpl32-trnL* hSTR2 and *rpl32-trnL* SNP loci, haplotypes found and allele nomenclature proposal. The sequenced amplicon also included *rpl32-trnL* hSTR1, indicated as [hSTR1] in the sequence. The preceding SNP was a C in all samples sequenced compared to an A in the reference sequences (NC_026562, AGQN01337109, KR184827, NC_027223, KY084475, and KR779995)

GTTTTATGTTTTGTTTTGAACTGATCAACAACAATATAATATTAAATTCAATTT
GTTTCCTTTTTTT**[REPEAT]**AAAAAAAACTAAGAATTCCTTTGTCTACTCTTTTG
AATATGCTTTCTCATTTTTAGGATGGGGAAAATAAGAATCC

Allele	[REPEAT]		N	Genbank accesion number
	T	A		
6	0	6	17	MK989691
6	1	5	2	MK989692
6	3	3	1	MK989693
6	2	4	1	MK989694
8	8	0	1	MK989695
8	7	1	2	MK989696
9	8	1	2	MK989697
14	6	8	1	MK989698

Fig. 2.3 Consensus sequence of *rpl32-trnL* hSTR3 locus, haplotypes found and allele nomenclature proposal

TCCTTTGTCTACTCTTTTGAATATGCTTTCTCATTTTTAGGATGGGGAAAATAA
GAATCCCCATCAATTCGAT**[INDEL]**AATAAAAATGATTTCTCTTTTATTTATAT
TATTTTTCTATTATTTTATACTATTTAAATTTTAAATATTTCTTAGACTAACAT
AGAAATTAGAGTATAGAAGAGCATATATATATAATTTGTAGTCAAAACTAAT
AGGTCGTTGAAACTCATTAA

Allele	[INDEL]	N	Genbank accesion number
	AATAAA		
Present	Present	8	MK989699
Absent	Absent	4	MK989700

Fig. 2.4 Consensus sequence of *rpl32-trnL* INDEL locus, haplotypes found and allele nomenclature proposal

CCGGGCCTCTTTTATTCCAAAAAATCAAATTT**W**AATTATTTAATTTATTACAT
TTGAAAACACAAATGCTTATTACTTATTATTAATATTAAATAATATTAAAACA
ATCATAAGATCATAAGAAAGACTGTTTTGATTCCTTTGATATATAATCACAAT
GTCATTTCTTA**[REPEAT]**AAACACTAAAATTGCTTTGATTTACTTTATAATAC
ACCAACACGT

Allele	Preceding SNP	[REPEAT]		N	Genbank accesion number
	W	TTA	T		
15	A	1	12	25	MK989701
16	T	2	10	1	MK989702

Fig. 2.5 Consensus sequence of *trnS-trnG* hSTR1 locus, haplotypes found and allele nomenclature proposal

GTCCACTCAGCCATCTCTCCCCCAATTGAAAAAAAAAA**Y**ATATATATAATAAGA
AGGGTTTTTTCAAGCCTTATTTTGGCTTATGGAACCTTATGGAACATAGTAATTA
TTCTTATTATATATATAT**W**TTTTTTTTCAATATAGAAATATAACTATAAACTATAT
AAAAAAAACAATAT**[INDEL]**AAAAATAAGAATTAAGAAATAAAATACAAAAAA
ATACTATATCTCTTTGATTTTTTCCAAAGAAACCTTATTCTTTCCGCGGCTTGGC
CTGGTCAATACCTAGCTGGGCCGGGCCTCTTTTATTCCAA

SNP1		SNP2		Post-INDEL	N	Genbank accesion number
Y (forward)	R (reverse)	W (forward)	W (reverse)	CAATAT		
C	G	T	A	–	6	MK989703
T	A	A	T	–	4	MK989704
T	A	A[1]	T[1]	–	7	MK989705
T	A	A[1]	T[1]	CAATAT	2	MK989706

Fig. 2.6 Consensus sequence of *trnS-trnG* SNP1 and *trnS-trnG* SNP2 loci, haplotypes found and allele nomenclature proposal

[1] Sequences which were haplotype A with one fewer T in SBE primer binding site for SNP2 – these 9 samples were sequenced due to failure of single base extension

AAAACCGCTAGCAAAAGACTTGGTTTTTAGTTAGTTAAATGAGTCTGCGTTTC
CTAATCTCAATCTCATTAAGTCCTCTTGGTACCGCTCTAATTTTCATGATTCTTTT
TTGATCCTATTTTTWAATTACACCTGAGTTGCTTGTTCGAC

SNP3		N	Genbank accesion number
W (forward)	W (reverse)		
T	A	6	MK989707
A	T	4	MK989708

Fig. 2.7 Consensus sequence of *trnS-trnG* SNP3 locus, haplotypes found and allele nomenclature proposal

Fragment analysis development and genotyping

Due to the A/T-rich nature of the chloroplast genome and the frequency of homopolymeric regions, primer design was limited. As a result, three fragment analysis assays were developed to avoid primer interactions. The *rpl32-trnL* markers hSTR1 and hSTR3 formed one multiplex assay; *rpl32-trnL* hSTR2 and INDEL formed another; and *trnS-trnG* hSTR1 was amplified in singleplex. The primer sequences, concentrations, and optimal annealing temperatures are listed in Table 2.2. Allelic ladders were created for each of the three assays (Fig. 2.8). The *rpl32-trnL* INDEL and hSTR2 assay ladder contained two alleles per locus; *rpl32-trnL* hSTR1 and hSTR3 contained two alleles for hSTR1 and four alleles for hSTR3; and *trnS-trnG* hSTR1 contained two alleles. Following sample genotyping, an off-ladder allele was observed for *rpl32-trnL* hSTR1 in six samples. It was sequenced and found to be a 10-A repeat (10 allele). Two new alleles were observed for *rpl32-trnL* hSTR3, but the motifs were unable to be clarified by Sanger sequencing due to poor sequencing quality; by their size, they were predicted to be 11 and 15 alleles, which occurred in only one and two samples, respectively. It is expected that using an MPS method would allow for determination of the repeat motif as well as provide distinction between isoalleles since alleles of different sequences

but similar fragment lengths were observed in the initial sequencing screen. Example electropherograms are shown in Fig. 2.9. Despite the high stutter levels characteristic of hSTRs, all sample profiles were relatively balanced and easy to interpret.

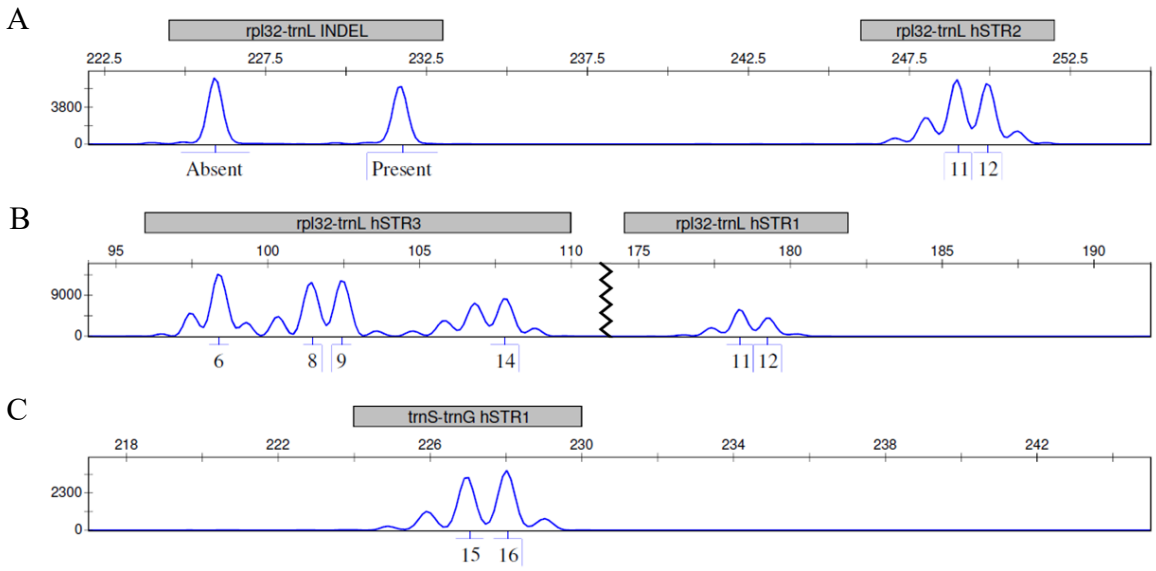


Fig. 2.8 Allelic ladders for *rpl32-trnL* INDEL & hSTR2 (A), *rpl32-trnL* hSTR1 & hSTR3 (B), and *trnS-trnG* hSTR1 (C)

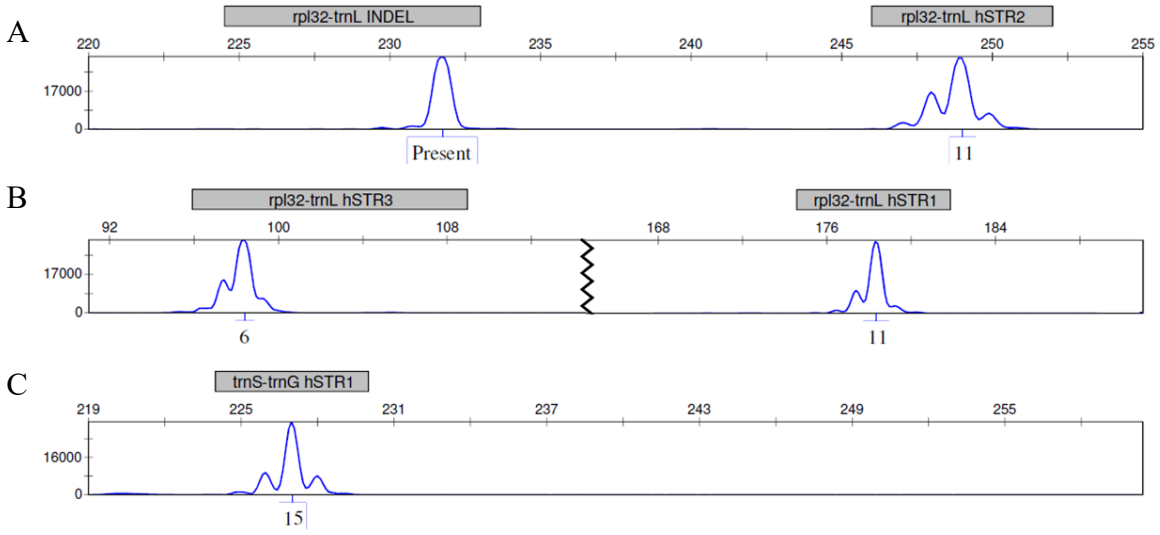


Fig. 2.9 Example electropherograms of a marijuana sample for *rpl32-trnL* INDEL & hSTR2 (A), *rpl32-trnL* hSTR1 & hSTR3 (B), and *trnS-trnG* hSTR1 (C)

SNaPshot™ assay development and genotyping

Two SNaPshot™ assays were developed, one for each hotspot region. The *rpl32-trnL* region contained only one SNP, and trnS-trnG contained three SNPs, which were analyzed in multiplex. During genotyping, it was observed that trnS-trnG SNP2 dropped out in several samples (six Chilean marijuana, two Mexican marijuana, and one USA hemp). The region was sequenced to determine why genotyping failed. In all nine samples, the SNP genotype was A, and there was a mutation in the SBE primer-binding region. In the majority of samples, nine Ts follow the SNP site, and the SBE primer contains nine As. However, in the nine samples where SNP genotyping failed, only eight Ts are present, which affects the binding of the reverse SBE primer. The sample genotypes were recorded as A; however, an MPS method of genotyping would easily distinguish between samples with and without this mutation, creating an additional haplotype. Another interesting observation during sequencing of the region containing *trnS-trnG* SNP2 was an unexpected INDEL seen in two samples (both from Chilean marijuana case 7) with a CAATAT motif. The INDEL occurred outside of the SBE primer binding site and did not affect genotyping of the SNP; therefore, it is unknown whether other samples contain this INDEL marker. An example SNP profile is shown in Fig. 2.10; when determining the genotype, it is important to consider that the *rpl32-trnL* SNP and *trnS-trnG* SNP2 were amplified using reverse SBE primers.

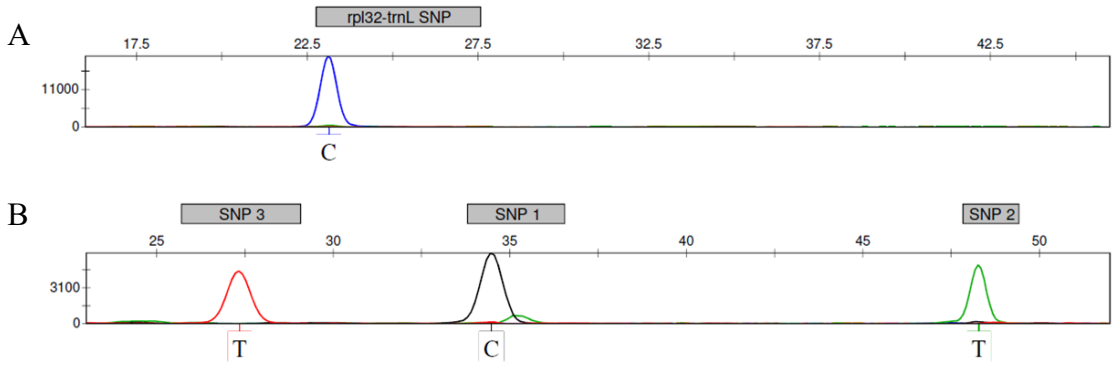


Fig. 2.10 Example SNaPshot™ profile of a marijuana sample for *rpl32-trnL* SNP (A) and *trnS-trnG* SNPs 1-3 (B). Alleles are named for the base on the forward DNA strand

Haplotype analysis

Eight different haplotypes (Table 2.7) were observed in the 152 samples tested from USA or Canada hemp, USA-Mexican or Chilean marijuana, and Chilean medical marijuana. Haplotype 1 was the most common (81.6% of samples), and it occurred in all sample groups except in Canadian hemp. As expected, haplotype analysis was not as discriminatory as previously used autosomal STR multiplexes [23] due to the inheritance pattern and lack of recombination in the chloroplast genome. Another factor for this extreme haplotype sharing was that 71.6% percent of the samples in our database were marijuana samples from Mexico.

Canadian hemp consisted of haplotypes 5 and 6, which were unique to that group. USA hemp consisted of haplotypes 1, 7, and 8, with 7 and 8 being unique to that group; haplotype 8 was observed in only one sample. The USA-Mexican marijuana group consisted of haplotypes 1, 2, and 3 with haplotype 2 being unique to the group and seen in six out of 109 samples tested. Haplotype 3 was observed in only two out of 109 USA-Mexican marijuana samples. Chilean marijuana samples consisted of haplotypes 1, 3, and 4. Previous autosomal typing revealed that all Chilean marijuana cases except for case 8 were clonally propagated (all samples within each of the nine cases shared an autosomal genotype) [38]. As expected,

samples from the same case shared a haplotype, and interestingly, some cases could be distinguished from others by their haplotype. For example, haplotype 4 was unique to the Chilean marijuana group and occurred only in the two samples from case 7. The samples from cases 5 and 9 shared haplotype 3. All four strains of medical marijuana from Chile shared haplotype 1. Fig. 2.11 shows the haplogroups mapped by geographic origin, and a complete breakdown of the haplotypes for each case or sample bag is available in Table 2.8. Before using these data to determine the geographic origin or crop type of an unknown sample, it is important to expand the database to include samples from all around the world. Indeed, the discriminatory power of this haplotype analysis will not be truly known until a more diverse sample set can be genotyped.

Table 2.7 Haplotypes observed. Eight distinct haplotypes were observed in this study. For simplicity, the haplotypes are referred to as 1-8

Haplotype	*rpl32-trnL*					*trnS-trnG*			
	hSTR 1	hSTR 2	hSTR 3	INDEL	SNP	hSTR 1	SNP 1	SNP 2	SNP 3
1	11	11	6	Present	C	15	C	T	T
2	10	11	6	Present	C	15	C	T	T
3	11	12	14	Present	A	16	T	A	T
4	11	12	15	Present	A	16	T	A	T
5	12	11	8	Absent	A	15	T	A	A
6	12	11	9	Absent	A	15	T	A	A
7	12	12	6	Present	A	15	T	A	T
8	11	12	11	Present	A	16	T	A	T

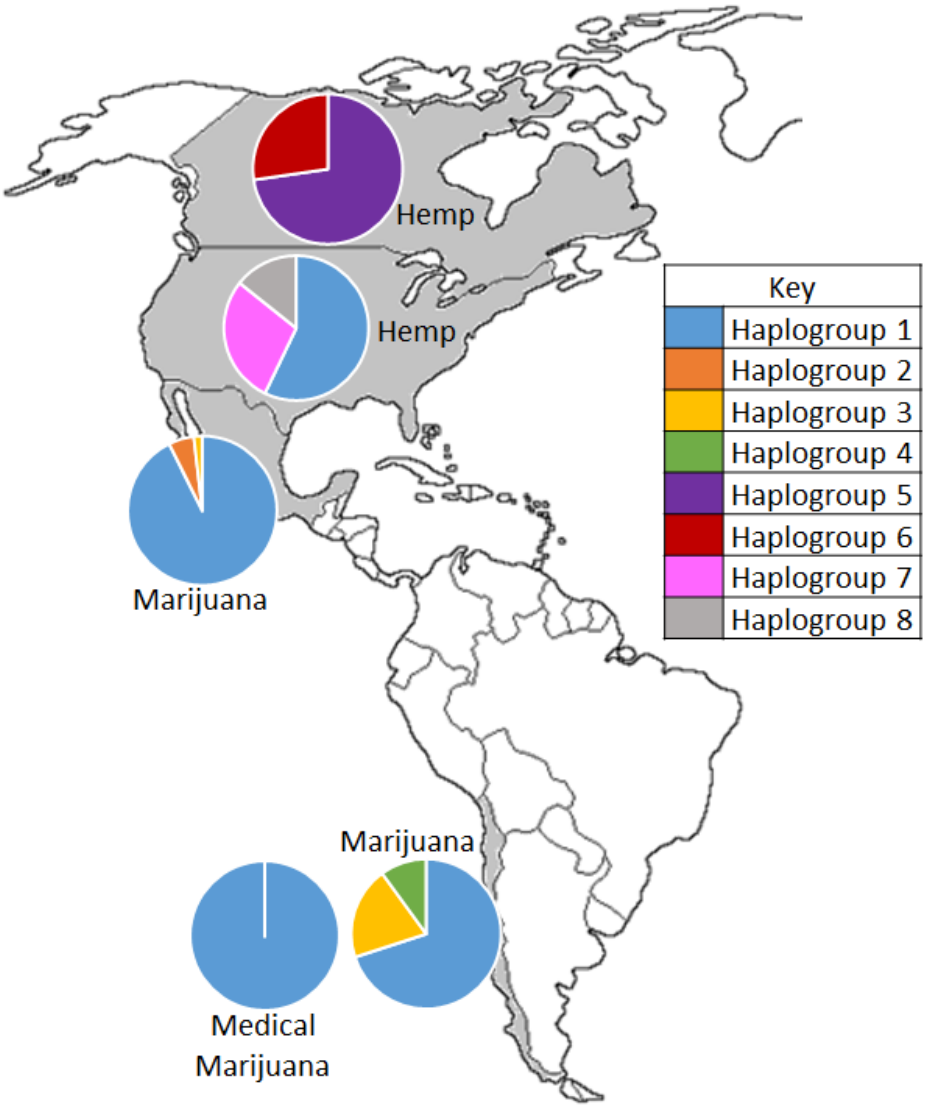


Fig. 2.11 Map of haplogroups. The haplotype proportions observed in the five sample groups (N=152) are displayed by crop type and geographic origin

Table 2.8 Breakdown of haplotypes observed in each case/bag for samples from Mexico, Chile, Canada, and USA

Population	N	Country of origin	Haplotypes
Drug-type (marijuana)			
CBP Case 1	4	Mexico	1
CBP Case 2	3	Mexico	1(N=2), 2(N=1)
CBP Case 3	4	Mexico	1
CBP Case 4	16	Mexico	1(N=15), 3(N=1)
CBP Case 5	2	Mexico	1
CBP Case 6	2	Mexico	1
CBP Case 7	5	Mexico	1
CBP Case 8	3	Mexico	1
CBP Case 9	3	Mexico	1
CBP Case 10	3	Mexico	1
CBP Case 11	7	Mexico	1
CBP Case 12	5	Mexico	1
CBP Case 13	3	Mexico	1
CBP Case 14	4	Mexico	1(N=3), 2(N=1)
CBP Case 15	3	Mexico	1(N=2), 2(N=1)
CBP Case 16	2	Mexico	1(N=1), 3(N=1)
CBP Case 17	9	Mexico	1
CBP Case 18	7	Mexico	1
CBP Case 19	7	Mexico	1
CBP Case 20	8	Mexico	1(N=6), 2(N=2)
CBP Case 21	9	Mexico	1(N=8), 2(N=1)
Chile Case 1	2	Chile	1
Chile Case 2	2	Chile	1
Chile Case 3	2	Chile	1
Chile Case 4	2	Chile	1
Chile Case 5	2	Chile	3
Chile Case 6	2	Chile	1
Chile Case 7	2	Chile	4
Chile Case 8	3	Chile	1
Chile Case 9	2	Chile	3
Chile Case 10	2	Chile	1
Medical Amnesia	1	Chile	1
Medical AK	1	Chile	1
Medical Lemon Haze	1	Chile	1
Medical London Cheese	1	Chile	1

(continued)

Population	N	Country of origin	Haplotypes
Fiber-type (hemp)			
Manitoba Harvest Bag 1	4	Canada	5(N=2), 6(N=2)
Manitoba Harvest Bag 2	1	Canada	5
Badia Spices	3	Canada	5(N=2), 6(N=1)
Navitas Organics	3	Canada	5
American Hemp Harvest	5	USA	1(N=2), 7(N=2), 8(N=1)
Original Hemp Buds Electra	1	USA	1
Original Hemp Buds Lifter	1	USA	1

Statistical analysis

Phylogenetic analysis (neighbor-joining method) and pairwise comparison of the five populations using F_{ST} as genetic distance revealed the genetic association of two sets of populations (Table 2.9, Table 2.10). The distance matrix comparing the five sample groups is shown in Table 2.9. Using the neighbor-joining method with coancestry as genetic distance, it was determined that genetic similarities exist between the following populations: USA-Mexican marijuana and Chilean medical marijuana, and Chilean medical marijuana, Chilean marijuana, and USA hemp. No statistically significant differences were detected for any of these pairs of populations ($p > 0.05$) (Table 2.10). PCA analysis (Fig. 2.12) revealed a clear distinction between Canadian hemp compared with the other four groups. However, the pairs of USA-Mexican marijuana and Chilean marijuana and the group of Chilean marijuana, Chilean medical marijuana, and USA hemp samples are shown to be closely related as predicted from the genetic distance calculations. To note, only the Canadian hemp samples clustered close to each other. The other populations had haplotypes that were found in multiple populations. Genotyping population samples and more markers in other hotspot regions and the higher discriminatory power of MPS may help to resolve these sample groups in future studies.

Table 2.9 Distance matrix from GDA software estimated using the Neighbor-joining method with coancestry distance

Population	USA-Mexico marijuana	Chile marijuana	Canada hemp	USA hemp
Chile marijuana	0.428119			
Canada hemp	3.03761	1.36497		
USA hemp	0.93009	0.00232648	1.32309	
Chile medical marijuana	0[a]	0.0458501	3.04893	0.173353

[a] Negative distance values were transformed to zero

Table 2.10 Population-to-population comparison among five cannabis populations using pairwise genetic-distance analysis based on F_{ST}.

Population	USA-Mexico marijuana	Chile marijuana	Canada hemp	USA hemp
Chile marijuana	0.35963 (0.00000[a])			
Canada hemp	0.95903 (0.00000 [a])	0.76052 (0.00000 [a])		
USA hemp	0.63793 (0.00901[a])	0.02678 (0.20721)	0.73369 (0.00000 [a])	
Chile medical marijuana	-0.13017 (0.99099)	0.04353 (0.53153)	0.95259 (0.00000 [a])	0.15916 (0.27027)

Probability values of F_{ST} are displayed in parentheses

[a] Statistically significant differences at 0.01 levels

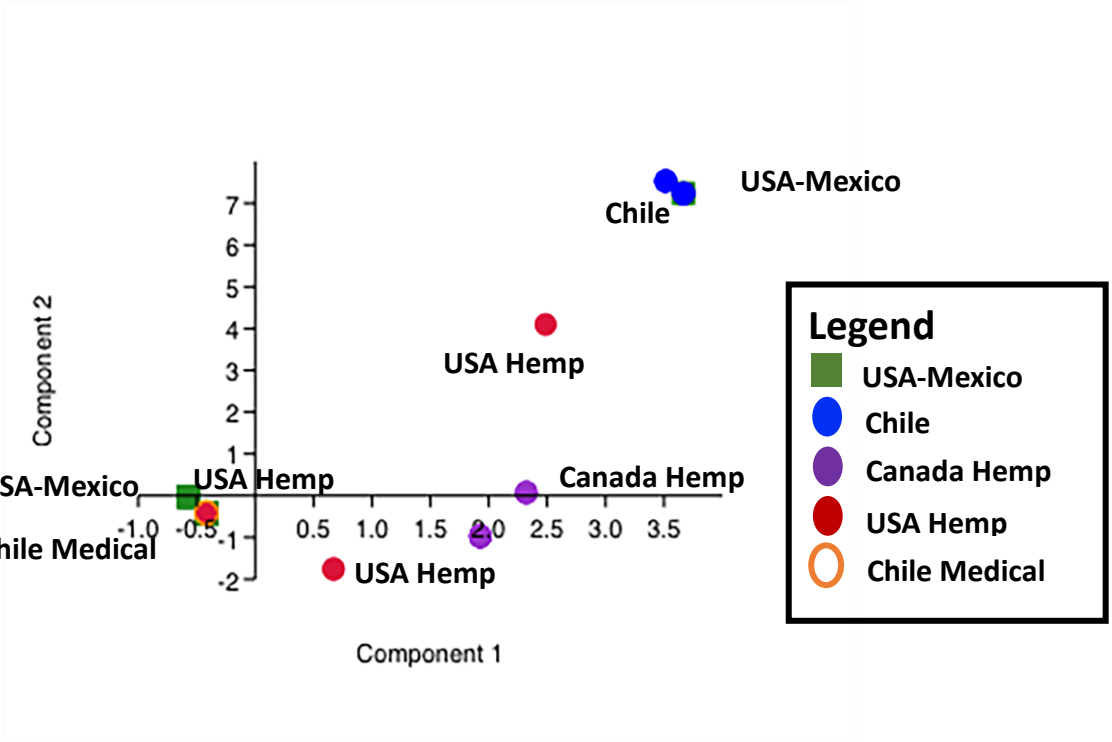


Fig. 2.12 PCA plot. Principal component analysis was performed using Past3 software on haplotype data (N=152). The results show some differentiation between the sample groups with USA hemp, Chile Medical, and USA-Mexico marijuana being the most similar. To note, labels have been added for points that have been hidden

Conclusions

This project proposed seven promising new polymorphic hotspot regions to be used in discriminating between crop type (marijuana versus hemp) and biogeographic origin of *C. sativa* samples. Assays were developed to genotype the *trnS-trnG* and *rpl32-trnL* regions, which had four and five polymorphic loci, respectively, and these assays were used to genotype samples from five different groups (USA hemp, Canadian hemp, USA-Mexican marijuana, Chilean marijuana, and Chilean medical marijuana). Despite over 80% of samples sharing the same haplotype, characteristic differences were observed among the five sample groups, indicating that these markers may have limited use in determining geographic origin and crop type. Increasing the number of markers and expanding the

sample database is imperative to more accurately assign origin and crop type of unknown samples.

Principal component analysis showed clear distinction between Canadian hemp compared with the other four groups. However, the USA-Mexican marijuana and Chilean medical marijuana samples and the Chilean marijuana, Chilean medical marijuana, and USA hemp samples were closely related. Using additional markers, including the other hotspot regions reported in this paper and those previously reported by Gilmore et al. [16], may help to further separate these groups. The authors also contributed to the development of a *C. sativa* chloroplast DNA database, which will need to be expanded to include more chloroplast markers as well as genotypes from samples obtained around the world.

Additionally, four unexpected polymorphisms were identified which were not observed when aligning the published chloroplast genomic sequences from Genbank®. One was a homopolymeric repeat immediately following SNP2 in the *trnS-trnG* region, which was found in a minority of samples to contain eight Ts rather than the nine Ts seen in other samples (Fig. 2.6). Also in the *trnS-trnG* region was a SNP (A/T) occurring before hSTR1 (Fig. 2.5). A new INDEL marker in the *rpl32-trnL* region was observed by chance after sequencing two Chilean marijuana samples (Fig. 2.4), and it may have been present in other samples. Additionally, a SNP (A/C) was observed preceding the *rpl32-trnL hSTR2* locus (Fig. 2.2). These four additional polymorphisms would be easily genotyped if the hotspot regions were incorporated into a targeted MPS assay, improving the discriminatory ability of these markers. In addition, known isoalleles (sequence variations of similar lengths) exist for the *rpl32-trnL* locus hSTR3, and these would be easily distinguished in an MPS assay.

References

1. Werf H, Mathussen E, Haverkort A (1996) The potential of hemp (Cannabis sativa L.) for sustainable fibre production: a crop physiological appraisal. Ann Appl Biol 129:109–123
2. Administration SAaMHS (2018) Key substance use and mental health indicators in the United States: results from the 2017 National Survey on Drug Use and Health (HHS Publication No. SMA 18-5068, NSDUH Series H-53). Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration Rockville, MD
3. Mikes F, Hofmann A, Waser PG (1971) Identification of (−)-Δ9-6a, 10a-trans-tetrahydro-cannabinol and two of its metabolites in rats by use of combination gas chromatography-mass spectrometry and mass fragmentography. Biochem Pharmacol 20:2469–2476
4. Mitosinka GT, Thornton JI, Hayes TL (1972) The examination of cystolithic hairs of Cannabis and other plants by means of the scanning electron microscope. J Forensic Sci Soc 12:521–529
5. Kayis SA, Hakki EE, Pinarkara E (2010) Comparison of effectiveness of ISSR and RAPD markers in genetic characterization of seized marijuana (Cannabis sativa L.) in Turkey. Afr J Agric Res 5:2925–2933
6. Pinarkara E, Kayis SA, Hakki EE, Sag A (2009) RAPD analysis of seized marijuana (Cannabis sativa L.) in Turkey. Electron J Biotechnol 12:5–6
7. Datwyler SL, Weiblen GD (2006) Genetic variation in hemp and marijuana (Cannabis sativa L.) according to amplified fragment length polymorphisms. J Forensic Sci

51:371–375

8. Kojoma M, Iida O, Makino Y, Sekita S, Satake M (2002) DNA fingerprinting of Cannabis sativa using inter-simple sequence repeat (ISSR) amplification. Planta Med 68:60–63
9. Mello I, Ribeiro A, Dias Vet al (2016) A segment of rbcL gene as a potential tool for forensic discrimination of Cannabis sativa seized at Rio de Janeiro, Brazil. Int J Legal Med 130:353–356
10. Dias V, Ribeiro A, Mello I et al (2015) Genetic identification of Cannabis sativa using chloroplast trnL-F gene. Forensic Sci Int Genet 14:201–202
11. Kohjyouma M, Lee I-J, Iida O et al (2000) Intraspecific variation in Cannabis sativa L. based on intergenic spacer region of chloroplast DNA. Biol Pharm Bull 23:727–730
12. Gigliano GS, Caputo P, Cozzolino S (1997) Ribosomal DNA analysis as a tool for the identification of Cannabis sativa L. specimens of forensic interest. Sci Justice 37:171–174
13. Gigliano GS (1998) Identification of Cannabis sativa L. (Cannabaceae) using restriction profiles of the Internal Transcribed Spacer II (ITS2). Sci Justice 38:225–230
14. Gigliano SS (1999) Preliminary data on the usefulness of internal transcribed spacer I (ITS1) sequence in Cannabis sativa L. identification. J Forensic Sci 44:475–477
15. Linacre A, Thorpe J (1998) Detection and identification of cannabis by DNA. Forensic Sci Int 91:71–76
16. Gilmore S, Peakall R, Robertson J (2007) Organelle DNA haplotypes reflect crop-use characteristics and geographic origins of Cannabis sativa. Forensic Sci Int 172:179–

190

17. Gilmore S, Peakall R (2003) Isolation of microsatellite markers in Cannabis sativa L. (marijuana). Mol Ecol Resour 3:105–107
18. Gilmore S, Peakall R, Robertson J (2003) Short tandem repeat (STR) DNA markers are hypervariable and informative in Cannabis sativa: implications for forensic investigations. Forensic Sci Int 131:65–74
19. Alghanim H, Almirall J (2003) Development of microsatellite markers in Cannabis sativa for DNA typing and genetic relatedness analyses. Anal Bioanal Chem 376:1225–1233
20. Köhnemann S, Nedele J, Schwotzer D, Morzfeld J, Pfeiffer H (2012) The validation of a 15 STR multiplex PCR for Cannabis species. Int J Legal Med 126:601–606
21. Hsieh H-M, Hou R-J, Tsai L-C, Wei CS, Liu SW, Huang LH, Kuo YC, Linacre A, Lee JCI (2003) A highly polymorphic STR locus in Cannabis sativa. Forensic Sci Int 131:53–58
22. Houston R, Birck M, Hughes-Stamm S, Gangitano D (2016) Evaluation of a 13-loci STR multiplex system for Cannabis sativa genetic identification. Int J Legal Med 130:635–647
23. Houston R, Birck M, Hughes-Stamm S, Gangitano D (2017) Developmental and internal validation of a novel 13 loci STR multiplex method for Cannabis sativa DNA profiling. Legal Med 26: 33–40
24. Dumolin-Lapegue S, Pemonge M, Petit R (1997) An enlarged set of consensus primers for the study of organelle DNA in plants. Mol Ecol 6:393–397
25. Weising K, Gardner RC (1999) A set of conserved PCR primers for the analysis of

simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. Genome 42:9–19

26. Haider N (2011) Chloroplast-specific universal primers and their uses in plant studies. Biol Plant 55:225–236
27. Demesure B, Sodzi N, Petit R (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Mol Ecol 4:129–134
28. Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol 17:1105–1109
29. Santos C, Pereira F (2017) Design and evaluation of PCR primers for amplification of four chloroplast DNA regions in plants. Conserv Genet Resour 9:9–12
30. Heinze B (2007) A database of PCR primers for the chloroplast genomes of higher plants. Plant Methods 3:4
31. Ferri G, Corradini B, Ferrari F, Santunione AL, Palazzoli F, Alu M (2015) Forensic botany II, DNA barcode for land plants: which markers after the international agreement? Forensic Sci Int Genet 15:131–136
32. Hollingsworth PM, Li D-Z, van der Bank M, Twyford AD (2016) Telling plant species apart with DNA: from barcodes to genomes. Phil Trans R Soc B 371:20150338
33. Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am J Bot 92:142–166
34. Shaw J, Lickey EB, Schilling EE, Small RL (2007) Comparison of whole chloroplast

genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. Am J Bot 94:275–288

35. Shaw J, Shafer HL, Leonard OR, Kovach MJ, Schorr M, Morris AB (2014) Chloroplast DNA sequence utility for the lowest phylogenetic and phylogeographic inferences in angiosperms: the tortoise and the hare IV. Am J Bot 101:1987–2004
36. Schaal B, Hayworth D, Olsen KM, Rauscher J, Smith W (1998) Phylogeographic studies in plants: problems and prospects. Mol Ecol 7:465–474
37. Zhang Q, Liu Y (2003) Examination of the cytoplasmic DNA in male reproductive cells to determine the potential for cytoplasmic inheritance in 295 angiosperm species. Plant Cell Physiol 44:941– 951
38. Houston R, Birck M, LaRue B, Hughes-Stamm S, Gangitano D (2018) Nuclear, chloroplast, and mitochondrial data of a US cannabis DNA database. Int J Legal Med 1–13
39. Shajitha P, Dhanesh N, Ebin P et al (2016) A combined chloroplast atpB–rbcL and trnL-F phylogeny unveils the ancestry of balsams (Impatiens spp.) in the Western Ghats of India. 3 Biotech 6:258
40. Vergara D, White KH, Keepers KG, Kane NC (2016) The complete chloroplast genomes of Cannabis sativa and Humulus lupulus. Mitochondrial DNA A 27:3793–3794
41. Oh H, Seo B, Lee S et al (2016) Two complete chloroplast genome sequences of Cannabis sativa varieties. Mitochondrial DNA A 27: 2835–2837
42. Darling AE, Mau B, Perna NT (2010) Progressivemauve: multiple genome alignment with gene gain, loss and rearrangement. PLOS One 5:e11147

43. QIAGEN (2018) DNeasy® plant handbook, Hilden, Germany

44. Koressaar T, Remm M (2007) Enhancements and modifications of primer design program Primer3. Bioinformatics 23:1289–1291

45. Vallone PM, Butler JM (2004) AutoDimer: a screening tool for primer-dimer and hairpin structures. Biotechniques 37:226–231

46. BigDye® direct cycle sequencing kit. (2011) Thermo Fisher Scientific, South San Francisco, CA

47. Mayr W (1995) DNA Recommendations-1994 report concerning further recommendations of the DNA commission of the ISFH regarding PCR-based polymorphisms in STR (short tandem repeat) systems. Vox Sang 69:70–71

48. Gill P, Brinkmann B, d'Aloja E, Andersen J, Bar W, Carracedo A, Dupuy B, Eriksen B, Jangblad M, Johnsson V, Kloosterman AD, Lincoln P, Morling N, Rand S, Sabatier M, Scheithauer R, Schneider P, Vide MC (1997) Considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature. Forensic Sci Int 87:185–192

49. Parson W, Ballard D, Budowle B, Butler JM, Gettings KB, Gill P, Gusmão L, Hares DR, Irwin JA, King JL, Knijff P, Morling N, Prinz M, Schneider PM, Neste CV, Willuweit S, Phillips C (2016) Massively parallel sequencing of forensic STRs: considerations of the DNA commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements. Forensic Sci Int Genet 22:54–63

50. Griffiths R, Barber M, Johnson P et al (1998) New reference allelic ladders to improve allelic designation in a multiplex STR system. Int J Legal Med 111:267–272

51. ABI PRISM® SNaPshot™Multiplex Kit Protocol 4323357b. (2010) Thermo Fisher Scientific, South San Francisco, CA

52. Sanchez JJ, Børsting C, Hallenberg C, Buchard A, Hernandez A, Morling N (2003) Multiplex PCR and minisequencing of SNPs—a model with 35 Y chromosome SNPs. Forensic Sci Int 137:74–84

53. Lewis P, Zaykin D. (2001) Genetic data analysis: computer program for the analysis of allelic data. Version 1.0 (d16c). https://phylogeny.uconn.edu/software/. Accessed 31 July 2019

54. Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564–567

55. Hammer Ø, Harper DA, Ryan PD (2001) PAST: paleontological statistics software package for education and data analysis. Palaeontol Electron 4:9

56. Johnson LA, Soltis DE (1994) matK DNA sequences and phylogenetic reconstruction in Saxifragaceae s. str. Syst Bot 143-56

57. Steele KP, Vilgalys R (1994) Phylogenetic analyses of Polemoniaceae using nucleotide sequences of the plastid gene matK. Syst Bot 19:126–142

58. Dong W, Liu J, Yu J, Wang L, Zhou S (2012) Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. PLoS One 7:e35071

59. Oxelman B, Lidén M, Berglund D (1997) Chloroplastrps16 intron phylogeny of the tribeSileneae (Caryophyllaceae). Plant Syst Evol 206:393–410

60. Takayama K, OHI-TOMA T, Kudoh H, Kato H (2005) Origin and diversification of

Hibiscus glaber, species endemic to the oceanic Bonin Islands, revealed by chloroplast DNA polymorphism. Mol Ecol 14:1059–1071

61. Smith SD, Cowan RS, Gregg KB, Chase MW, Maxted N, Fay MF (2004) Genetic discontinuities among populations of Cleistes (Orchidaceae, Vanilloideae) in North America. Bot J Linn Soc 145:87–95
62. Small RL, Ryburn JA, Cronn RC, Seelanan T, Wendel JF (1998) The tortoise and the hare: choosing between noncoding plastome and nuclear Adh sequences for phylogeny reconstruction in a recently diverged plant group. Am J Bot 85:1301–1315
63. Hu C, Zheng X, Teng Y (2011) Characterization and phylogenetic utility of non-coding chloroplast regions trnL-trnF and accD-psaI in Pyrus. Acta Horticulturae Sinica 38:2261–2272
64. Erixon P, Oxelman B (2008) Whole-gene positive selection, elevated synonymous substitution rates, duplication, and indel evolution of the chloroplast clpP1 gene. PLoS One 3:e1386
65. Hamilton M (1999) Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. Mol Ecol 8: 521–523
66. Hamilton MB (1999) Tropical tree gene flow and seed dispersal. Nature 401:129–130
67. Olson ME (2002) Intergeneric relationships within the Caricaceae-Moringaceae clade (Brassicales) and potential morphological synapomorphies of the clade and its families. Int J Plant Sci 163:51–65
68. Xu D, Abe J, Sakai M, Kanazawa A, Shimamoto Y (2000) Sequence variation of non-coding regions of chloroplast DNA of soybean and related wild species and its

implications for the evolution of different chloroplast haplotypes. Theor Appl Genet 101: 724–732

69. Byrne M, Hankinson M (2012) Testing the variability of chloroplast sequences for plant phylogeography. Aust J Bot 60:569–574

70. Yang J-B, Yang H-Q, Li D-Z, Wong K-M, Yang Y-M (2010) Phylogeny of Bambusa and its allies (Poaceae: Bambusoideae) inferred from nuclear GBSSI gene and plastid psbA-trnH, rpl32-trnL and rps16 intron DNA sequences. Taxon 59:1102–1110

71. Dunbar-Co S, Wieczorek AM, Morden CW (2008) Molecular phylogeny and adaptive radiation of the endemic Hawaiian Plantago species (Plantaginaceae). Am J Bot 95:1177–1188

72. Timme RE, Kuehl JV, Boore JL, Jansen RK (2007) A comparative analysis of the Lactuca and Helianthus (Asteraceae) plastid genomes: identification of divergent regions and categorization of shared repeats. Am J Bot 94:302–312

73. Sakai M, Kanazawa A, Fujii A, Thseng F, Abe J, Shimamoto Y (2003) Phylogenetic relationships of the chloroplast genomes in the genus Glycine inferred from four intergenic spacer sequences. Plant Syst Evol 239:29–54

CHAPTER III

Investigation of chloroplast regions *rps16* and *clpP* for determination of *Cannabis sativa* crop type and biogeographical origin

---

This dissertation follows the style and format of *International Journal of Legal Medicine*.

Roman, MG, Houston, R (2020) Investigation of chloroplast regions *rps16* and *clpP* for determination of *Cannabis sativa* crop type and biogeographical origin. Submitted to Legal Med

Abstract

Cannabis sativa can be classified as either hemp (a legal crop containing less than 0.3% delta-9-tetrahydrocannabinol, THC) or marijuana (an illegal drug containing more than 0.3% THC). Despite its legalization in 33 states for medicinal or recreational use, marijuana remains the most commonly used illicit drug in the USA, and it is heavily trafficked into and within the country. Discriminating between marijuana and hemp is critical to the legal process. Genetic analysis provides a means of analyzing samples unsuitable for chemical analysis, and in addition to discriminating between crop types, DNA may be able to determine the biogeographical origin of samples. In addition, the sharing of rare haplotypes between different seizures may be useful for linking cases and providing investigative leads to law enforcement.

This study evaluates the potential of two highly polymorphic regions of the chloroplast genome of *C. sativa*, *rps16* and *clpP*, to be used for determination of crop type and biogeographical origin. Custom fragment analysis and SNaPshot™ assays were developed to genotype nine polymorphic loci in hemp samples from the USA and Canada, marijuana samples from USA-Mexico and Chile, and medical marijuana samples from Chile.

Haplotype analysis revealed eight haplotypes. Only Canadian hemp could be completely differentiated from the other sample groups by haplotype. Phylogenetic analysis and principal component analysis suggested a closer relationship among USA-Mexico marijuana, Chilean marijuana and medical marijuana, and USA hemp. Genotyping additional polymorphisms in future studies is expected to reveal further differences between these sample groups.

Keywords: *Cannabis sativa*, Marijuana, Chloroplast DNA, DNA Barcoding, Forensic Plant Science, Haplotype

Introduction

Cannabis sativa is cultivated worldwide as a source of industrial fiber and seed oil and as a drug with recreational, religious, and medicinal uses [1]. The psychoactive compound delta-9-tetrahydrocannabinol (THC) distinguishes the drug marijuana from the non-drug form of the plant, known as hemp. In the United States, the 2018 Farm Bill broadly legalized hemp, defined as *C. sativa* that contains no more than 0.3% THC. On the other hand, marijuana is categorized as schedule I by the Controlled Substances Act (CSA), and possession of the drug is a criminal offense in most states. Marijuana has been legalized for medicinal or recreational use in 33 states and the District of Columbia. Despite this, marijuana remains the most commonly used illicit drug in the United States [2], and it is heavily trafficked, with many seizures occurring at the border with Mexico. Additionally, state-specific legalization of the drug requires law enforcement to investigate the diversion of legal marijuana to states where it remains illegal.

Before the 2018 Farm Bill, the goal of forensic scientists was simply to identify a plant as *C. sativa*, but with the legalization of hemp, a new challenge has unfolded: distinguishing between marijuana and hemp. Traditional methods of identifying *C. sativa*, including chemical methods to detect the presence (rather than quantity) of THC and microscopical analysis of plant material [3, 4], are no longer enough evidence to prove a plant is marijuana. Additionally, these methods cannot provide information about the source of a sample or be used to link growers and distributors. DNA analysis has the potential to discriminate between marijuana and hemp as well as show genetic associations among plants with common growers and distributors, thus elucidating criminal networks and trafficking routes. Another advantage to using sensitive genetic methods is that they

can be used to analyze material that is unsuitable for chemical or microscopical analysis, including seeds or roots, trace residues, juvenile plants, and highly fragmented materials.

DNA barcoding is a technique that involves sequencing specific sections of the genome (usually chloroplast or mitochondria) to identify differences between individuals of different species (using interspecific sequence differences) or different populations (using intraspecific sequence differences). This allows identification of species by comparison to databases such as the Barcode of Life Database (BOLD) or determination of biogeographical origin, which requires databases containing the sequences of individuals of a species from different geographic regions. In animals, the mitochondrial cytochrome c oxidase 1 gene (CO1) is universally used to identify animal species [5]. However, many studies have shown that no single marker can distinguish between all species of plants [6-11]. Plant mitochondrial genomes have a low rate of substitution and are prone to heteroplasmy and intramolecular recombination, making chloroplast regions more optimal for barcoding studies [6, 12]. The most commonly used barcoding markers in plants are *rbcL*, *matK*, *trnH-psbA*, and *ITS2* [13]. However, several studies have investigated a variety of regions in the nuclear, mitochondrial, and chloroplast genomes and found that other markers are more variable and more informative for both species identification and determination of biogeographical origin [7, 8, 10].

Previous studies in marijuana have shown that the *ITS1*, *ITS2*, and *trnL-trnF* regions have sequences which are unique to *C. sativa*, making them useful for species identification [14-18]. Additionally, several markers have been tested for their ability to distinguish between hemp and marijuana samples from different geographic origins [19-22]. Kohjyouma et al. [19] studied interspecific variation in the intergenic spacer region

trnL-trnF and concluded that *C. sativa* had developed local mutations in this and other regions of the genome. Mello et al. [20] used a portion of the *rbcL* gene and found differences between samples seized in Brazil compared to database sequences of samples from China, The United Kingdom, and the United States. Gilmore et al. [21] identified five chloroplast DNA (cpDNA) and two mitochondrial DNA (mtDNA) polymorphisms which varied among *C. sativa* plants from different biogeographical origins and crop types (fiber, drug, and wild types). A subsequent study by Houston et al. [22] tested the same seven organelle markers, as well as nuclear DNA (nDNA) markers in samples from North and South America. Though the organelle markers showed promise in both studies for distinguishing between different crop types and populations of samples, they could not definitively identify either crop type or biogeographical origin. Exploration of additional barcode regions could help to strengthen these classifications. Seven "hotspots," or highly polymorphic regions, in the chloroplast genome of *C. sativa* were identified by Roman et al. [23]: *trnK-matK-trnK*, *rps16*, *trnS-trnG*, *ycf3*, *accD-psaI*, *clpP*, and *rpl32-trnL*. The study also analyzed ten polymorphisms within the *rpl32-trnL* and *trnS-trnG* regions and showed that they differed between marijuana and hemp and between samples from the United States, Canada, Chile, and Mexico.

The purpose of this study was to evaluate the hotspot regions *rps16* and *clpP*, previously reported by Roman et al. [23], for their variability and to determine whether they are useful for determining the biogeographical origin and crop type of *C. sativa* samples. The *rps16* region contains three single nucleotide polymorphisms (SNPs) and a homopolymeric short tandem repeat (hSTR), and the *clpP* region contains four hSTRs and an insertion-deletion (INDEL) [23]. Fragment analysis assays were developed to quickly

genotype the size-based polymorphisms (hSTRs and INDEL), and SNaPshot™ assays were used to sequence the SNPs. Marijuana seized at the USA-Mexico border, Chilean marijuana, Chilean medical marijuana, hemp grown in the USA, and hemp grown in Canada were genotyped, and a haplotype map was developed to visualize differences between the sample groups. Identifying regional differences in these hotspot DNA sequences could assist law enforcement in determining whether seized samples represent licit or illicit *C. sativa*. It could also provide investigative leads to agencies investigating marijuana trafficking into and within the United States by providing evidence for associations between cases and suggesting entry points into the country.

Materials and methods

Sample collection

Samples (N=166) from five groups were examined in this study: marijuana from USA-Mexico (N=108), marijuana from Chile (N=21), medical marijuana from Chile (N=4), hemp from Canada (N=11), and hemp from the USA (N=22). Approximately every fourth sample from a larger collection was tested. USA-Mexico marijuana was obtained from the United States Customs and Border Protection (CBP) (N=422 samples from 21 cases). DNA extracts of marijuana from the Araucania region of southern Chile (N=50 samples from 10 cases) were obtained from the Policia de Investigaciones. Chilean medical marijuana extracts (N=12 samples; 3 each of the strains Amnesia, AK, Lemon Haze, and London Cheese) were provided by collaborators in Chile. Hemp seeds grown in Canada were purchased from Manitoba Harvest (Winnipeg, MB, CA; N=15 samples from two bags), Badia Spices Inc. (Doral, FL, USA; N=15), and Navitas™ Organics (Novato, CA, USA; N=15). USA hemp seeds were purchased from American Hemp Harvest (Boulder,

CO, USA; N=10). USA Hemp CBD flowers were purchased from The Original Hemp Buds (OR or NY, USA) (strains: Electra (2 bags), Lifter (2 bags), Sour Space Candy, Special Sauce, and Cascade; N=1 from each bag) and CBD Hemp Direct (Las Vegas, NV, USA) (strains: Hemp World Haze, Sunset Rd Sherbert #2, Paradise OG, Casino Cookies #2, Durban Potion #2, Tangie, Juicy Fruit #2, and Trophy Wife; N=1 each). USA Hemp CBG flowers were purchased from CBD Hemp Direct (strains: Desert Snow and Jazzy CBG; N=1 each).

DNA extraction and quantification

DNA extraction was performed using the DNeasy® Plant Mini kit (QIAGEN, Hilden, Germany) according to the manufacturer's protocol [24]. Plant material (leaves, flowers, stems, or seeds) was homogenized in liquid nitrogen using Kimble-Chase Kontes™ Pellet Pestles™ (Kimble-Chase, Rockwood, TN). DNA extraction from USA-Mexico marijuana samples was performed onsite at CBP. DNA extracts for the Chilean marijuana and medical marijuana samples were provided by collaborators in Chile. Hemp DNA extractions were performed at Sam Houston State University.

Quantification of cpDNA was performed on a StepOne™ Real-Time PCR System (Thermo Fisher Scientific, South San Francisco, CA, USA) using a previously reported real-time PCR assay [22].

Sanger sequencing

Sanger sequencing was used to assess allelic variability at all loci previously identified in the hotspots, *rps16* and *clpP*. A small number of samples (N=3-10) were sequenced using the BigDye™ Direct Cycle Sequencing Kit (Thermo Fisher Scientific) according to the manufacturer's instructions [25]. Primers were designed using Primer3

software [26], and annealing temperatures were determined using gradient PCR as previously described [23]. Geneious Pro Software R7.1.9 was used to align and analyze the sequences, and sequences were submitted to GenBank® (accession numbers MT053877-MT053885).

Fragment analysis development and genotyping

Genotyping of the size-based polymorphisms (five hSTRs and one INDEL) was performed in multiplex using custom fragment analysis assays. Primers were designed using Primer3 software [26] (Table 3.1), and AutoDimer software [27] was used to predict primer interactions. Optimal annealing temperatures for each primer set were determined by gradient PCR as previously described [23]. As a note, due to the proximity of *clpP* hSTR1 and *clpP* hSTR2 and limitations in primer design due to the A/T-rich nature of the chloroplast genome, one primer set was used to amplify both polymorphisms. Additionally, one primer set was used to amplify both *clpP* INDEL and *clpP* hSTR4 since the INDEL occurs immediately following the hSTR.

The PCR master mix contained 6.25 µL of Type-it Microsatellite PCR Master Mix (QIAGEN), 1.25 µL of primer mix (Integrated DNA Technologies, Coralville, IA, USA) (Table 3.1), and 1.25 µL of 5x Q-solution (QIAGEN). A 3.75 µL aliquot of sample DNA (20 pg/µL) was added to each reaction, and thermal cycling was carried out on a Veriti™ 96-well Thermal Cycler (Thermo Fisher Scientific). Amplification parameters were as follows: 5 min activation at 95 °C; 30 cycles of 30 s at 95 °C, 90 s at 63 °C, and 30 s at 72°C; and 30 min at 60°C.

Capillary electrophoresis (CE) was carried out on a 3500 Genetic Analyzer (Thermo Fisher Scientific). Mixtures of 9 µL HiDi™ Formamide (Thermo Fisher

Scientific), 0.5 µL GeneScan™ 600 LIZ™ (Thermo Fisher Scientific), and 0.5 µL amplified DNA were denatured at 95 °C for 5 min. CE run parameters were as follows: oven 60°C; pre-run 15 kV, 180s; injection 1.6kV, 8s; run 19.5kV, 1330 s; capillary length 50 cm; polymer POP-7™; and dye set G5.

Data analysis was performed using GeneMapper ID v.5 software (Thermo Fisher Scientific). A custom panel and bins were designed, and the analytical threshold was set to 100 relative fluorescence units (RFU).

Allelic ladder design and allele nomenclature

An allelic ladder was developed as previously described [23] with sequenced alleles. Allele nomenclature followed ISFG guidelines where appropriate [28-30]. Since *clpP* hSTR1 and hSTR2 were amplified together, the allele was designated as the sum of the two repeats (e.g., an 11 allele for hSTR1 with a 5 allele for hSTR2 was designated as 16). The *clpP* INDEL and *clpP* hSTR4 were also coamplified, and alleles were designated by the number of repeats in hSTR4, followed by "A" for absence of the INDEL or "P" for presence (e.g., an 11 repeat for hSTR4 and a present INDEL was designated 11P). Where accurate sequence data could not be obtained, alleles were designated by their base pair (bp) size.

SNaPshot™ assay development and genotyping

Primer3 software [26] was used to design primers for amplification of the *rps16* SNPs. Amplification reactions were prepared with 6.25 µL of Type-it® Microsatellite PCR Master Mix (QIAGEN), 1.25 µL of 2 µM primer mix (Integrated DNA Technologies) (Table 3.2), 1.25 µL of 5x Q-solution (QIAGEN), and 3.75 µL of DNA sample (20 pg/µL). Thermal cycling was carried out on an Eppendorf Mastercycler® Gradient Thermal Cycler

(Eppendorf, Hamburg, Germany) as follows: 5 min activation at 95 °C; 30 cycles of 30 s at 95 °C, 90 s at 63 °C, and 60 s at 72°C; and a 30 min final extension at 60°C. Cleanup of the PCR products was performed by adding 5 µL calf intestinal alkaline phosphatase (CIAP) (1 U/µL, Promega Corporation, Madison, WI, USA) and 2 µL exonuclease I (10 U/µL, Invitrogen), followed by a 90 min incubation at 37 °C and a 30 min enzyme inactivation at 75 °C.

Single base extension (SBE) was carried out using the SNaPshot™ Multiplex Kit (Thermo Fisher Scientific) according to the manufacturer's instructions [31] using half reaction volumes. SBE primers were designed to complement the 20 bp sequence adjacent each SNP site (Table 3.2). Neutral sequences were added to the primers to allow size separation of each SBE product [32], and AutoDimer [27] was used to predict primer interactions. To avoid formation of hairpin and dimer structures, a reverse SBE primer was developed for SNP1; however, allele nomenclature for all SNPs was based on the forward strand. Cycling was carried out on an Eppendorf Mastercycler® Gradient Thermal Cycler (Eppendorf). Cleanup of the SBE products was performed by adding 1 µL CIAP (1 U/µL, Promega Corporation) and incubating at 37°C for 90 min and 75 °C for 30 min.

CE was performed on a 3500 Genetic Analyzer (Thermo Fisher Scientific) with the following parameters: oven 60°C; pre-run 15kV, 180 s; injection 1.6 kV, 8 s; run 15 kV, 560s; capillary length 36 cm; polymer POP-4™; and dye set E5. GeneMapper ID v.5 (Thermo Fisher Scientific) was used to analyze the data with an analytical threshold of 100 RFU. A custom panel and bin set was developed for allele calling.

Table 3.1 Primer sequences and annealing temperatures for fragment analysis assays

Primer name	5’ Fluorescent tag	Sequence	Approximate amplicon size (bp)	Conc. of primer pair in PCR (µM)	Annealing temperature (°C)
rps16 hSTR F	6FAM	TCATAAAACCCCACTTTCCGA	155	0.04	63
rps16 hSTR R	—	CTTGAGCCGTACGAGGAGAA			
clpP hSTR1&2 F	VIC	ACTATGATGGTTCCGTTGCTT	231	0.04	63
clpP hSTR1&2 R	—	GAGACCCATTTCAGCGTCAC			
clpP hSTR3 F	VIC	TCGATAAAGTCGGTTGATTGG	154	0.04	63
clpP hSTR3 R	—	ATTCGACAGGGCCTGCTAT			
clpP hSTR4&INDEL F	6FAM	CCCGATTTGGATTTGCCTAT	248	0.03	63
clpP hSTR4&INDEL R	—	CCAGGCTCCGTTTAGAAACAA			

Table 3.2 Primer sequences for SNaPshot™ analysis. Neutral sequences used for spacing of markers are indicated by lowercase letters

Primer name	Sequence	SBE primer length (bp)	Concentration of primer in PCR (µM)	Annealing temperature (°C)
rps16 SNPs F	GGTTGGTGATTAAGGCGAAG	—	0.2	63
rps16 SNPs R	TCGATGGAGAAATCGAAAGAG			
rps16 SNP1 SBE R	TTTCAAAAAAGGCAGGGGTT	20	0.2	—
rps16 SNP2 SBE F	gtgaaagtctgacaaGATAGATGTAGATAAAAAAT	35	0.2	—
rps16 SNP3 SBE F	aaactaggtgccacgtcgtgaaagtctgacaaTTTGAATTTGAAACTTG CTT	52	0.2	—

Data Analysis

Haplotypes were assigned based on the multi-locus genotypes as shown in Table 3.3. A haplotype map was developed showing the frequency of each haplotype in samples from different biogeographical origins. Genetic distance was calculated using the neighbor-joining method with coancestry genetic distance with Genetic Data Analysis (GDA) software [33]. Pairwise comparisons and tests for significant differences among the five populations were performed using Arlequin v. 3.5 software [34] with F_{ST} as genetic distance. Finally, principal component analysis (PCA) was performed using Past3 v.3 software [35].

Table 3.3 Breakdown of haplotypes

Haplotype	*rps16*		*clpP*		
	SNP3	hSTR	hSTRs 1 & 2	hSTR3	hSTR4 & INDEL
1	A	150	14	150	11P
2	A	150	14	151	11P
3	G	153	16	147	10A
4	G	152	16	147	10A
5	G	150	16	150	12A
6	A	150	14	151	12P
7	G	150	15	150	12A
8	G	152	15	148	11A

Results

Sanger sequencing

Sequencing of the *rps16* SNPs revealed two alleles for SNP3 (A and G) (Fig. 3.1), but only A alleles were observed for SNPs 1 and 2. It was decided that all three SNPs would be included in the SNaPshot™ assay, allowing all 166 samples to be screened for rare alleles in SNPs 1 and 2. The locus *rps16* hSTR (an hSTR with a C motif) resulted in

poor Sanger sequencing that was difficult to align; to avoid incorrect nomenclature, the two alleles were named by their bp sizes (150 and 153) rather than the number of repeats.

Sequencing of the *clpP* polymorphisms revealed three alleles for hSTR1 (10, 11, and 12 A repeats), two alleles for hSTR2 (a 4 allele with the motif TTTT and a 5 allele with the motif TATTT), three alleles for hSTR4 (believed to be 10, 11, and 12 T repeats), and two alleles for the INDEL (present, with an insert of the 9 bp sequence TTCAATTTA, and absent) (Fig. 3.1-3.3). The *clpP* hSTR3 locus (an hSTR with a T motif) resulted in a low-quality sequence, so its alleles were designated as 147 and 150 for their bp size.

This is a note for Fig. 3.1 through 3.3. In the consensus sequences, the forward and reverse primer binding locations are underlined and the single base extension (SBE) primer binding sites are highlighted.** The location of hSTR repeats are indicated in the consensus sequences as **[REPEAT]**, indels are indicated as **[INDEL]**, and SNPs are indicated by their **nucleotide ambiguity code.** The GenBank accession numbers are also referenced in the table. N refers to the total number of samples with the indicated haplotype.

AGCAACATACCATTTTGAGGGATTTCTTTCTATCAAAGAATCATACGAATGGTT
GATTCCTGTGTAATACACTTTTGATTTTATTGAAAGAGTTTTACCTATTCACCAA
AAAATTTACTTTTGAATTTGAAACTTGCTTRAATTGGACTCTTTCGATTTCTCC
ATCGAAAATTTACTTACAAAGTTGTCCCAATTTATTAATTGATACTAACCTTAGA
TTCTTGCCTCCG

SNP3		N	Genbank accesion number
R (forward)	Y (reverse)		
A	T	4	MT053884
G	C	5	MT053885

Fig. 3.1 Consensus sequence of *rps16* SNP 3 locus, haplotypes found and allele nomenclature proposal

TGATCCAATAACCACCCTTCCTTTTTGAGTAGTTAAAAAAATACTATGATGGT
TCCGTTGCTTTATATATATATTTCGTCTGTTAGTTATTAGTTAGCAATCCCAAA
GTTTTTTTATTTG**[hSTR1]**TATATATATATAAATATATATATCAAAAATATATA
TCTAAGTAAAAAACGAAATTCTA**[hSTR2]**CGTATTCATTCATAATATAAATAT
ATATTGTTAAAAGTTTTTCGGTGTGAAAAAAGTTTGTGACGCTGAAATGGGT

Allele	[hSTR1]	[hSTR2]	N	Genbank accesion number
	A	Var.		
14	10	TTTT	2	MT053877
15	11	TTTT	2	MT053878
16	12	TTTT	1	MT053879
16	11	TATTT	2	MT053880

Fig. 3.2 Consensus sequence of *clpP* hSTR 1 and 2 loci, haplotypes found and allele nomenclature proposal

CCCGATTTGGATTTGCCTATATAGGACAAATGGACAAATACTATGTCTTTTTG
CTACGACTTC**[hSTR4][INDEL]**TCAATTTATTTCATATCTCCTACCAAATATTCT
ATTTGAAATCACGTCTATTCATATTAGAAATTAGAAATCGAATATAATAGAA
AATATGATATAAAATATGATCATCTAAGTAGAAATCCTAGATATATTACCAA
TTGTTTTGTTTCTAAACGGAGCCTGG

Allele	[hSTR4]	[INDEL]	N	Genbank accesion number
	T	TTCAATTTA		
10A	10	Absent	2	MT053881
11A	11	Absent	1	MT053882
12A	12	Absent	2	TBDMT053883
11P	11	Present	2	Insufficient quality
12P	12	Present	1	Insufficient quality

Fig. 3.3 Consensus sequence of *clpP* hSTR 4 and INDEL loci, haplotypes found and allele nomenclature proposal

Fragment analysis and allelic ladder

An example electropherogram for the fragment analysis assay is shown in Fig. 3.4. All 166 samples tested resulted in full profiles (Table 3.4). The allelic ladder (Fig. 3.5)

contained two alleles (150 and 153) for *rps16* hSTR, two alleles (14 and 16) for *clpP* hSTRs 1 and 2, two alleles (147 and 150) for *clpP* hSTR3, and three alleles (10A, 12A, and 11P) for *clpP* hSTR4 and INDEL.

Sample genotyping revealed several off-ladder alleles. Three samples had a 152 bp allele for *rps16* hSTR. For *clpP* hSTRs 1 and 2, two samples had an off-ladder allele that was between the bins for 14 and 16; sequencing revealed a 15 allele with 11 repeats of A for hSTR1 and 4 repeats of T for hSTR2 in both samples. One sample had a 148 bp allele for the *clpP* hSTR3 locus, and 67 samples had a 151 bp allele. For *clpP* hSTR4 and INDEL, an off-ladder allele between the bins for 10A and 12A was observed in one sample; sequencing revealed an 11A allele with 11 repeats of T at hSTR4 and an absent INDEL.

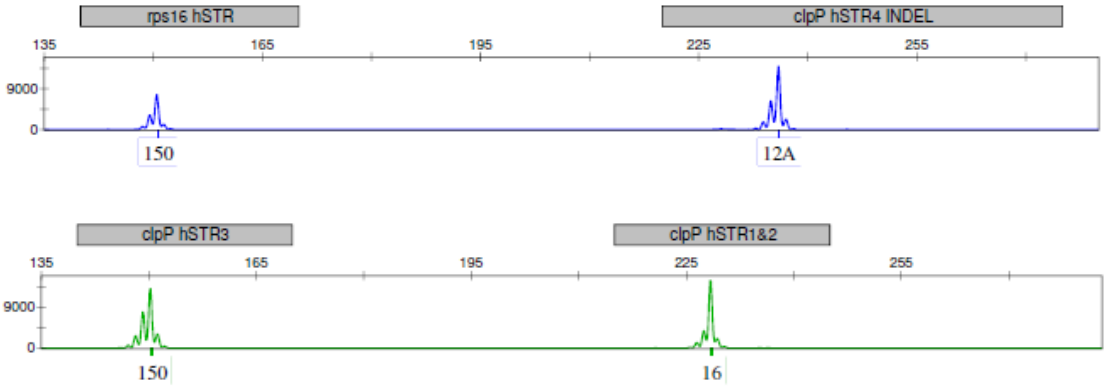


Fig. 3.4 Example fragment analysis electropherogram of a Canadian hemp sample

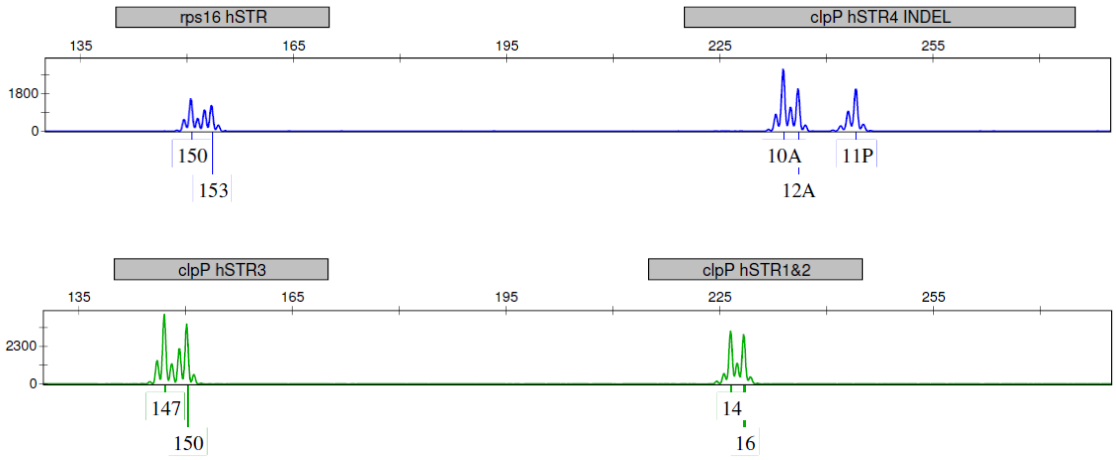


Fig. 3.5 Allelic ladder for the fragment analysis assay

Table 3.4 Breakdown of haplotypes observed in each case/bag for samples from Mexico, Chile, Canada, and USA

Population	N	Country of Origin	Haplotypes
Drug-type (marijuana)			
CBP Case 1	4	Mexico	1(N=1), 2(N=3)
CBP Case 2	3	Mexico	1(N=2), 2(N=1)
CBP Case 3	4	Mexico	1(N=3), 2(N=1)
CBP Case 4	16	Mexico	1(N=11), 2(N=4), 3(N=1)
CBP Case 5	2	Mexico	1
CBP Case 6	2	Mexico	1
CBP Case 7	5	Mexico	1(N=4), 2(N=1)
CBP Case 8	3	Mexico	1(N=2), 2(N=1)
CBP Case 9	3	Mexico	1
CBP Case 10	3	Mexico	1(N=1), 2(N=2)
CBP Case 11	7	Mexico	1(N=5), 2(N=2)
CBP Case 12	5	Mexico	1(N=3), 2(N=2)
CBP Case 13	3	Mexico	2
CBP Case 14	4	Mexico	1(N=1), 2(N=3)
CBP Case 15	3	Mexico	1(N=1), 2(N=2)
CBP Case 16	2	Mexico	1(N=1), 3(N=1)
CBP Case 17	9	Mexico	1(N=8), 2(N=1)
CBP Case 18	7	Mexico	1(N=4), 2(N=3)
CBP Case 19	7	Mexico	1(N=3), 2(N=4)
CBP Case 20	7	Mexico	1(N=2), 2(N=5)
CBP Case 21	9	Mexico	1(N=5), 2(N=4)
Chile Case 1	2	Chile	2
Chile Case 2	2	Chile	2
Chile Case 3	2	Chile	1
Chile Case 4	2	Chile	2

(continued)

Population	N	Country of Origin	Haplotypes
Chile Case 5	2	Chile	3
Chile Case 6	2	Chile	2
Chile Case 7	2	Chile	4
Chile Case 8	3	Chile	1
Chile Case 9	2	Chile	3
Chile Case 10	2	Chile	1
Medical Amnesia	1	Chile	1
Medical AK	1	Chile	1
Medical Lemon Haze	1	Chile	1
Medical London Cheese	1	Chile	2
Fiber-type (Hemp)			
Manitoba Harvest Bag 1	4	Canada	5
Manitoba Harvest Bag 2	1	Canada	5
Badia Spices	3	Canada	5
Navitas Organics	3	Canada	5
American Hemp Harvest	5	USA	2(N=2), 6(N=1), 7(N=1), 8(N=1)
Original Hemp Buds Electra Bag 1	1	USA	2
Original Hemp Buds Electra Bag 2	1	USA	1
Original Hemp Buds Lifter Bag 1	1	USA	2
Original Hemp Buds Lifter Bag 2	1	USA	2
Original Hemp Buds Sour Space Candy	1	USA	2
Original Hemp Buds Special Sauce	1	USA	2
Original Hemp Buds Cascade	1	USA	2
CBD Hemp Direct Hemp World Haze	1	USA	2
CBD Hemp Direct Sunset Rd Sherbert #2	1	USA	2
CBD Hemp Direct Paradise OG	1	USA	2
CBD Hemp Direct Casino Cookies #2	1	USA	2
CBD Hemp Direct Durban Potion #2	1	USA	1
CBD Hemp Direct Tangie	1	USA	2
CBD Hemp Direct Juicy Fruit #2	1	USA	2
CBD Hemp Direct Trophy Wife	1	USA	1
CBD Hemp Direct Desert Snow	1	USA	2
CBD Hemp Direct Jazzy CBG	1	USA	2

SNaPshot™ assay

All 166 samples yielded full profiles for the SNaPshot™ 3-plex assay. An example electropherogram is shown in Fig. 3.6. SNPs 1 and 2 in the *rps16* region were monomorphic for A alleles. SNP3, however, did show polymorphism, as expected, with 21 samples having a G allele.

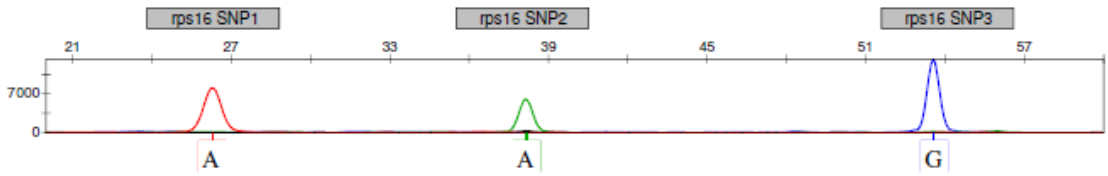


Fig. 3.6 Example SNaPshot™ electropherogram of a Canadian hemp sample

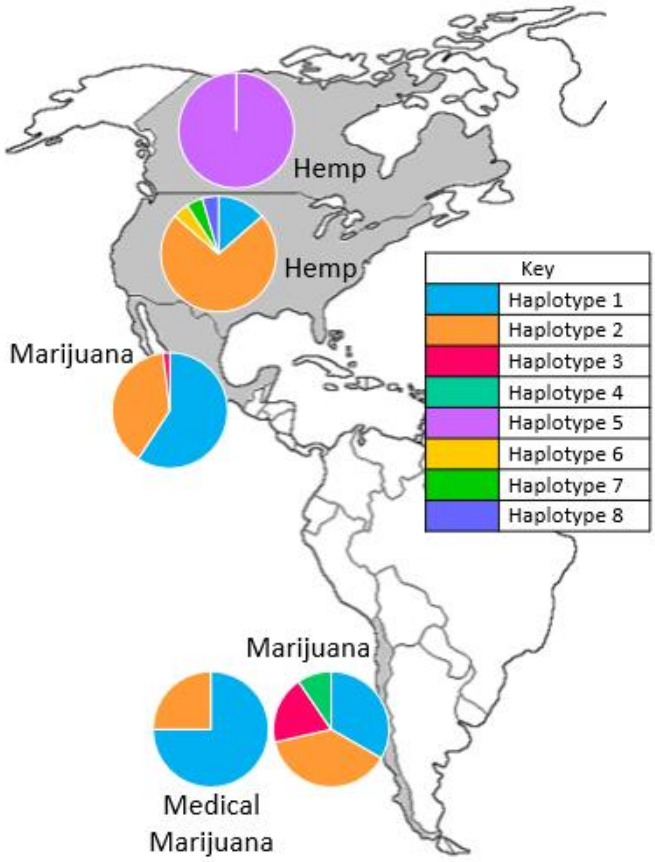


Fig. 3.7 Haplotype map showing proportions of each haplotype observed in the sample groups (N=166 individuals), displayed by crop type and geographic origin

Haplotype analysis

Eight haplotypes were observed (Table 3.3), and their distribution between sample groups are shown in the haplotype map (Fig. 3.7). Haplotypes 1 and 2 were the most common, occurring in 46.39% and 40.36% of samples, respectively. Both were observed

in all sample groups except Canadian hemp. Haplotype 3 (3.61% of samples) occurred in two USA-Mexico marijuana samples from different cases and in all samples from Chile marijuana cases 5 and 9. Haplotype 4 was only observed in two samples, both from Chile marijuana case 7. Haplotype 5 was unique to Canadian hemp, with all Canadian hemp samples tested resulting in the same haplotype. Haplotypes 6, 7, and 8 were each unique to single USA hemp samples from USA Hemp Harvest.

Statistical analysis

Phylogenetic analysis and pairwise comparisons (Tables 3.5 and 3.6) showed a relatedness between Chilean marijuana and USA-Mexico marijuana, Chilean marijuana and medical marijuana, and USA hemp and Chilean medical marijuana ($p>0.05$). Only Canadian hemp was completely differentiated from the other four populations ($p<0.01$). PCA (Fig. 3.8) graphically displays the variation among sample groups. All Canadian hemp samples group together (represented by a single point). The other sample groups are represented by multiple dots (representing samples with different haplotypes), with some samples from different groups being indistinguishable from each other (showing no haplotype variation). In particular, USA-Mexico and Chilean marijuana share three points (representing haplotypes 1, 2, and 3), with two of those points (haplotypes 1 and 2) also being shared by USA hemp and Chilean medical marijuana.

Table 3.5 Distance matrix from GDA software estimated using the neighbor-joining method with coancestry distance

Population	USA-Mexico marijuana	Chile marijuana	Canada hemp	USA hemp
Chile marijuana	0.288787			
Canada hemp	1.81171	0.827868		
USA hemp	0.204683	0.135601	1.66983	
Chile medical marijuana	0	0.0524657	3.30949	0.164869

Table 3.6 Population-to-population comparison among five cannabis populations using pairwise genetic distance analysis based on F_{ST}

Population	USA-Mexico marijuana	Chile marijuana	Canadian hemp	USA hemp
Chile marijuana	0.26470 (0.00000[a])			
Canada hemp	0.86452 (0.00000[a])	0.57033 (0.00000[a])		
USA hemp	0.17679 (0.00901[a])	0.13673(0.00000[a])	0.83404 (0.00000[a])	
Chile medical marijuana	-0.10432 (0.72973)	0.04486 (0.44144)	0.97225 (0.00000[a])	0.13810 (0.09009)

Probability values of F_{ST} are displayed in parentheses

[a] Statistically significant differences at 0.01 level

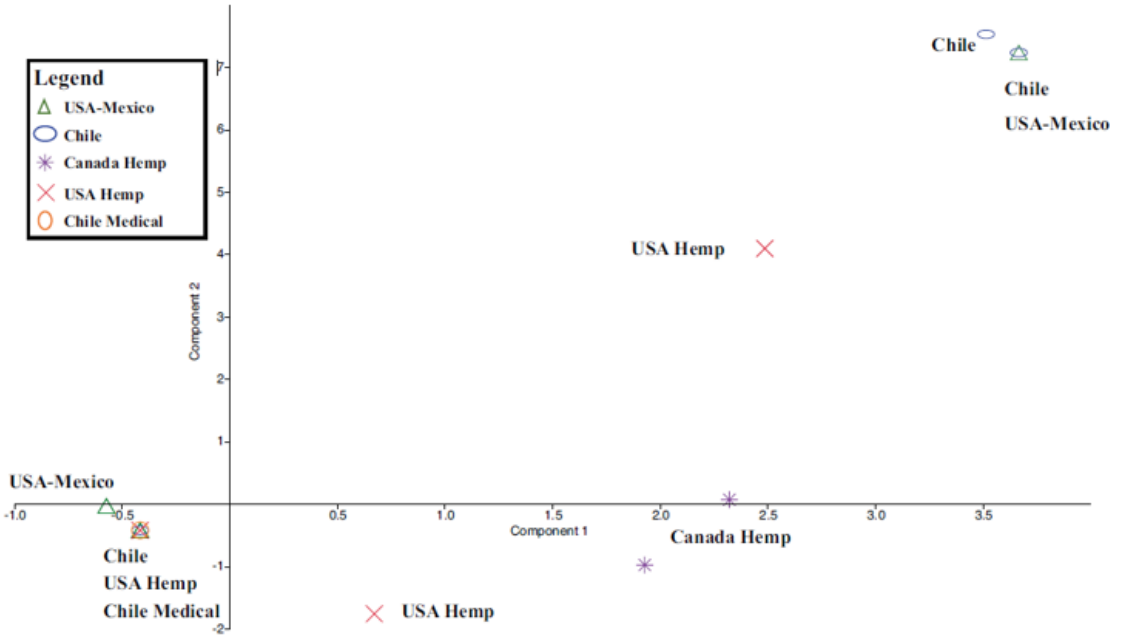


Fig. 3.8 PCA plot created from haplotype data (N=166) using Past3 software

Discussion

Sanger sequencing

Due to the long homopolymeric stretch in *clpP* hSTR4, Sanger sequencing resulted in a low quality sequence which was difficult to align, especially when the INDEL was present. Two alleles with the present INDEL were designated 11P and 12P (Fig. 3.3) based on the sequences obtained and the size of the alleles; however, the quality of the sequences was not sufficient for submission to GenBank®.

It is important to note that *clpP* hSTRs 1 and 2 were sequenced separately but amplified together in the fragment analysis assay. Though three distinct combination genotypes for *clpP* hSTRs 1 and 2 were observed during sequencing (hSTR1/2: 10/4, 11/5, and 12/4), only two size-based alleles were expected (14 and 16, the sum of the individual alleles) due to similar migration of the 11/5 and 12/4 alleles. It is the authors' intention to

develop an MPS method for simultaneous, high-throughput sequencing of the seven hotspot regions identified in Roman et al. [23]. An MPS method would be capable of discriminating between sequence-variable alleles of the same size (such as the 11/5 and 12/4 variations of the 16 allele for *clpP* hSTRs 1 and 2), as well as elucidating the sequences of *rps16* hSTR and *clpP* hSTR3 alleles.

Haplotype analysis

Due to a lack of variability, *rps16* SNPs 1 and 2 were not useful for determining crop type or biogeographical origin in this sample set. Therefore, these SNPs were not included in the haplotypes used for statistical analysis. However, Roman et al. [23] found variability in these sites when comparing GenBank® database sequences of *C. sativa* from other countries. Genotyping of a more extensive sample set from other countries around the world is likely to reveal diversity at these sites.

As expected, extensive haplotype sharing was observed. Due to single parent inheritance of the chloroplast genome and lack of recombination, it was predicted that samples from the same population would share haplotypes. Samples from different populations were expected to show some regional variations which could be used to characterize unknown samples by their origin or crop type.

Samples which originated from the same case numbers among the Chilean marijuana samples shared haplotypes without exception. Interestingly, samples from case 7 shared haplotype 4, which was unique to this case alone and was not observed in any of the other sample groups. This could indicate that this Chilean marijuana seizure was unrelated to any of the other cases, perhaps even originating in a different country. Cases

5 and 9, however, shared the relatively rare haplotype 3, indicating that the two cases may be linked.

This single haplotype sharing among cases was not universal for the USA-Mexico marijuana. Samples from cases 5, 6, and 9 shared haplotype 1, and samples from case 13 shared haplotype 2. However, samples from the other 17 cases fell into multiple haplogroups. This may indicate that these large seizures of marijuana from the USA-Mexico border contain crops from multiple growers. Cases 4 and 16 each had one sample with haplotype 3, which was relatively uncommon; it is possible these two samples have a common origin.

Among different crop types, haplotypes 1 and 2 were shared by both hemp (from the USA only) and marijuana. Four haplotypes (haplotypes 5-8) were only observed in hemp, and two haplotypes (haplotypes 3 and 4) were only observed in marijuana. While this is promising, a much more diverse set of samples must be tested. Additionally, analysis of more regions in the cpDNA may help to further distinguish hemp and marijuana samples.

Previously, many of these same samples were assessed for variability at the *rpl32-trnL* and *trnS-trnG* hotspot regions regions [23]. This previous study also resulted in 8 haplotypes; however, when the data from all four hotspots are combined, 10 haplogroups are created. Additionally, many of these samples (excluding USA hemp and Chilean medical marijuana) were genotyped with the Gilmore et al. [21] markers, as reported by Houston et al. [22]. Combining data from all of these studies results in additional haplogroups, providing better differentiation between samples of different crop types and from different biogeographical origins.

Though the data indicate that polymorphisms in *rps16* and *clpP* vary to a limited extent in samples from different biogeographical origins and in different crop types, determination of crop type or biogeographical origin for unknown samples is not possible without an extensive database containing samples from around the world. The true discriminatory power of these hotspots will not be known until a more diverse sample set is analyzed.

Statistical analysis

A similar relatedness among the sample groups was observed previously using polymorphisms in the *rpl32-trnL* and *trnS*-trnG hotspot loci, with Chilean marijuana and USA-Mexico marijuana grouping together and Chilean marijuana, Chilean medical marijuana, and USA hemp grouping together [23]. In both studies, Canadian hemp was the only group completely separated genetically from the other populations. However, the five sample groups may become more distinct from each other through the genotyping of more samples and more polymorphisms in other hotspot regions. In future studies, the authors plan to genotype additional hotspot regions (*trnK-matK-trnK*, *ycf3*, and *accD-psaI*) in these populations and develop an MPS assay capable of simultaneously sequencing all seven hotspot regions in a high-throughput manner.

Conclusions

Fragment analysis and SNaPshot™ assays were developed to genotype nine polymorphic loci in the *rps16* and *clpP* hotspot regions in *C. sativa*. Variability of these markers was assessed in hemp and marijuana samples from North and South America (USA and Canadian hemp, USA-Mexico and Chilean marijuana, and Chilean medical

marijuana). It was expected that haplotype sharing would be observed between samples with common crop types and biogeographical origins.

Two of the loci, *rps16* SNPs 1 and 2, were monomorphic, making them unable to discriminate between samples from different populations. Variation was observed at the other seven loci, and samples were assigned haplotypes based on their genotypes at each locus. Haplogroups 1 and 2 accounted for approximately 46% and 40% of the sample database, respectively, and were made up of both marijuana and USA hemp samples. Canadian hemp was the only population which could be completely separated from the others by haplotype, with all samples belonging to haplogroup 5. Chilean marijuana samples originating from the same case number shared haplotypes, but 81% of USA-Mexico marijuana cases were made up of samples from two or three different haplogroups, possibly indicating that these large seizures contained marijuana from different origins or growers. Phylogenetic analysis, PCA, and pairwise comparisons revealed a clear distinction between Canadian hemp and the other groups as well as high genetic similarities between USA-Mexico marijuana, Chile marijuana, and Chile medical marijuana samples at these loci, limiting their use in determining biogeographical origin and crop type.

Overall, variability in the *rps16* and *clpP* hotspot regions has the potential to assist in differentiating samples from different biogeographical origins as well as to discriminate between marijuana and hemp. However, it is crucial to assess variability at more loci, including the previously reported organelle markers from Gilmore et al. [21] and Roman et al. [23]. Additionally, the true discriminatory power of these markers will not be known until a more extensive database of samples from around the world can be collected and analyzed. The authors intend to develop a massively parallel sequencing assay

incorporating these and other markers in order to more easily and rapidly genotype additional samples.

References

1. Werf H, Mathussen E, Haverkort A (1996) The potential of hemp (Cannabis sativa L.) for sustainable fibre production: a crop physiological appraisal. Ann Appl Biol 129: 109-23
2. Substance Abuse and Mental Health Services Administration (2018) Key substance use and mental health indicators in the United States: Results from the 2017 National Survey on Drug Use and Health (HHS Publication No. SMA 18-5068, NSDUH Series H-53). Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration Rockville, MD
3. Mikes F, Hofmann A, Waser PG (1971) Identification of (−)-Δ9-6a, 10a-trans-tetrahydro-cannabinol and two of its metabolites in rats by use of combination gas chromatography-mass spectrometry and mass fragmentography. Biochem Pharmacol 20: 2469-76
4. Mitosinka GT, Thornton JI, Hayes TL (1972) The examination of cystolithic hairs of Cannabis and other plants by means of the scanning electron microscope. J Forensic Sci Soc 12: 521-9
5. Hebert PD, Ratnasingham S, De Waard JR (2003) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc R Soc B Biol Sci 270: S96-S9
6. Schaal B, Hayworth D, Olsen KM, Rauscher J, Smith W (1998) Phylogeographic studies in plants: problems and prospects. Mol Ecol 7: 465-74

7. Shaw J, Lickey EB, Beck JT et al (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am J Bot 92: 142-66
8. Shaw J, Lickey EB, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. Am J Bot 94: 275-88
9. Hollingsworth PM, Graham SW, Little DP (2011) Choosing and using a plant DNA barcode. PLOS One 6: e19254
10. Shaw J, Shafer HL, Leonard OR, Kovach MJ, Schorr M, Morris AB (2014) Chloroplast DNA sequence utility for the lowest phylogenetic and phylogeographic inferences in angiosperms: the tortoise and the hare IV. Am J Bot 101: 1987-2004
11. Hollingsworth PM, Li D-Z, van der Bank M, Twyford AD (2016) Telling plant species apart with DNA: from barcodes to genomes. Phil Trans R Soc B 371: 20150338
12. Haider N (2011) Chloroplast-specific universal primers and their uses in plant studies. Biologia Plantarum 55: 225-36
13. Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S (2015) Plant DNA barcoding: from gene to genome. Biol Rev 90: 157-66.
14. Dias V, Ribeiro A, Mello I et al (2015) Genetic identification of Cannabis sativa using chloroplast trnL-F gene. Forensic Sci Int Genet 14: 201-2
15. Gigliano GS, Caputo P, Cozzolino S (1997) Ribosomal DNA analysis as a tool for the identification of Cannabis sativa L. specimens of forensic interest. Sci Justice 37: 171-4

16. Gigliano GS (1998) Identification of Cannabis sativa L.(Cannabaceae) using restriction profiles of the Internal Transcribed Spacer II (ITS2). Sci Justice 38: 225-30
17. Gigliano SS (1999) Preliminary data on the usefulness of internal transcribed spacer I (ITS1) sequence in Cannabis sativa L. identification. J Forensic Sci 44: 475-7
18. Linacre A, Thorpe J (1998) Detection and identification of cannabis by DNA. Forensic Science International 91: 71-6.
19. Kohjyouma M, LEE I-J, Iida O et al (2000) Intraspecific variation in Cannabis sativa L. based on intergenic spacer region of chloroplast DNA. Biol Pharm Bull 23: 727-30
20. Mello I, Ribeiro A, Dias V et al (2016) A segment of rbcL gene as a potential tool for forensic discrimination of Cannabis sativa seized at Rio de Janeiro, Brazil. Int J Legal Med 130: 353-6
21. Gilmore S, Peakall R, Robertson J (2007) Organelle DNA haplotypes reflect crop-use characteristics and geographic origins of Cannabis sativa. Forensic Sci Int 172: 179-90
22. Houston R, Birck M, LaRue B, Hughes-Stamm S, Gangitano D (2018) Nuclear, chloroplast, and mitochondrial data of a US cannabis DNA database. Int J Legal Med: 1-13
23. Roman MG, Gangitano D, Houston R (2019) Characterization of new chloroplast markers to determine biogeographical origin and crop type of Cannabis sativa. Int J Legal Med 133: 1721-32
24. QIAGEN (2018) DNeasy® Plant Handbook. Hilden, Germany

25. Thermo Fisher Scientific (2011) BigDye® Direct Cycle Sequencing Kit. South San Francisco, CA

26. Koressaar T, Remm M (2007) Enhancements and modifications of primer design program Primer3. Bioinformatics 23: 1289-91

27. Vallone PM, Butler JM (2004) AutoDimer: a screening tool for primer-dimer and hairpin structures. Biotechniques 37: 226-31

28. Mayr W (1995) DNA Recommendations-1994 Report Concerning Further Recommendations of the DNA Commission of the ISFH Regarding PCR-Based Polymorphisms in STR (Short Tandem Repeat) Systems. Vox Sanguinis 69: 70-1

29. Gill P, Brinkmann B, d'Aloja E et al (1997) Considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature. Forensic Sci Int 87: 185-92

30. Parson W, Ballard D, Budowle B et al (2016) Massively parallel sequencing of forensic STRs: considerations of the DNA commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements. Forensic Sci Int Genet 22: 54-63

31. Thermo Fisher Scientific (2010) ABI PRISM® SNaPshot™ Multiplex Kit. South San Francisco, CA

32. Sanchez JJ, Børsting C, Hallenberg C, Buchard A, Hernandez A, Morling N (2003) Multiplex PCR and minisequencing of SNPs—a model with 35 Y chromosome SNPs. Forensic Sci Int 137: 74-84

33. Lewis P (2001) Genetic data analysis: computer program for the analysis of allelic data. Version 1.0 (d16c). http://lewiseebuconnedu/lewishome/software.html

34. Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10: 564-7

35. Hammer Ø, Harper DA, Ryan PD (2001) PAST: paleontological statistics software package for education and data analysis. Palaeontologia Electronica 4: 9

CHAPTER IV

Massively parallel sequencing of *Cannabis sativa* chloroplast hotspots for determination of biogeographical origin and crop type

---

This dissertation follows the style and format of *International Journal of Legal Medicine*.

Roman, MG, Gutierrez, R, Houston, R (2020) Massively parallel sequencing of *Cannabis sativa* chloroplast hotspots for determination of biogeographical origin and crop type. In preparation

Abstract

Massively parallel sequencing (MPS) offers distinct advantages over capillary electrophoresis (CE), including more comprehensive coverage of target loci, analysis of hundreds of markers simultaneously, and high throughput capabilities. This study reports on the development of a MiSeq FGx® assay targeting seven "hotspot" regions in the *Cannabis sativa* chloroplast genome that are highly polymorphic and informative for determining biogeographical origin and distinguishing between marijuana and hemp. Sequencing results were compared to previous studies that used CE-based genotyping methods. A total of 49 polymorphisms were observed, 16 of which have not been previously reported. Additionally, sequence data revealed isoalleles at one locus, which were able to differentiate two samples that had the same haplotype using CE-based methods. This study reports preliminary results from sequencing 14 hemp and marijuana samples from different countries using the developed MPS assay. Future studies should genotype a more comprehensive sample set from around the world to build a haplotype database, which could be used to determine crop type and biogeographical origin of unknown *C. sativa* samples and provide investigative leads for law enforcement agencies investigating marijuana trafficking.

Keywords: Massively parallel sequencing, *Cannabis sativa*, Haplotype, MiSeq FGx, Chloroplast genome, Biogeographical origin

Introduction

Massively parallel sequencing (MPS), also called next generation sequencing (NGS), is a high throughput technique capable of collecting DNA sequence data from multiple targets and multiple samples in parallel. It offers several distinct advantages over traditional DNA typing using capillary electrophoresis (CE), including providing more comprehensive coverage of target markers (sequence data in addition to length) and the ability to analyze hundreds or thousands of targets at a time [1, 2], compared to only about 25 loci by CE for a five-dye short tandem repeat (STR) kit [3]. Sequence data may elucidate isoalleles, alleles which have the same length and appear identical on CE but actually have different sequences, leading to more discriminatory results. Isoalleles may differ in their repeat structure or contain variants in the flanking regions, and the International Society of Forensic Genetics (ISFG) has reported guidelines to standardize the nomenclature for these sequence variable alleles [4]. Recently, costs and run times associated with MPS have dropped substantially, making targeted MPS assays a cost-effective approach for characterizing samples for genetic individualization or identification [1].

The two main MPS platforms used in forensic science are Thermo Fisher Scientific's Ion Torrent™ Ion S5™ (semi-conductor sequencing) and Verogen's MiSeq FGx® (reversible dye terminator sequencing) [5]. Semi-conductor sequencing platforms work by detecting a pH change caused by release of a proton when a pyrophosphate is cleaved during extension [6]. Reversible dye terminator sequencing uses dideoxynucleotide triphosphates (ddNTPs) with a reversible blocking group. The instrument records the fluorescence of a single base once it is incorporated, then the blocking group is removed and the next nucleotide is incorporated and recorded [7]. This

approach provides higher accuracy for homopolymeric stretches compared to semi-conductor sequencing [8-10].

In forensics, MPS assays have been used for human identification purposes, including sequencing autosomal STRs and single nucleotide polymorphisms (SNPs) [11-15], mitochondrial DNA analysis [16], phenotype prediction [17], and other purposes [2, 18]. While much of the forensic research on MPS has focused on human DNA, its use for forensic plant science has recently been investigated [19]. Houston et al. reported an MPS panel for the Ion S5™ consisting of twelve autosomal STRs in *Cannabis sativa* (marijuana). Results showed concordance with CE methods, and isoalleles were found at eight of the loci, providing a higher discriminatory power compared to CE [19]. MPS has been also been used for DNA barcoding studies in animals and has shown better recovery, reduced costs, and faster processing times compared to traditional Sanger sequencing [20]. Studies involving the use of MPS for DNA barcoding in plants have been limited but do show the advantage of simultaneous analysis of multiple barcodes for enhanced phylogenetic resolution [21, 22]. A recent study reported chloroplast DNA barcoding markers in *C. sativa* that were informative for biogeographical origin and crop type prediction [23]. Since these regions represent the most highly polymorphic regions of the *C. sativa* chloroplast genome, they are referred to as “hotspots.” The polymorphisms were genotyped using CE-based methods, and Sanger sequencing revealed isoalleles at several loci with different repeat sequences or variations in the flanking regions that were not detected by CE [23]. This study seeks to expand upon the previous study by incorporating the “hotspot” barcoding regions into an MPS assay to provide more discriminatory results and a high throughput method for building a database of *C. sativa* chloroplast haplotypes.

Full chloroplast genome sequences have been reported for several marijuana and hemp cultivars [24-26]. The full genome is 153,871 bp (Carmagnola and Dagestani cultivars) and contains 83 genes [25]. In comparison, the human mitochondrial genome is 16,569 bp [27] (about a tenth of the size), and typically mitochondrial DNA analyses only involve sequencing a portion of the genome, usually the hypervariable regions (HV1 and HV2) [28, 29]. Due to the large size of the *C. sativa* chloroplast genome, sequencing targeted regions (barcoding markers) gives better coverage and increased throughput capabilities compared to whole genome sequencing. The chloroplast genome is AT-rich (63%) and contains numerous homopolymeric stretches [25]. The MiSeq FGx® platform was chosen for sequencing in this study because has been shown to have a higher fidelity when sequencing homopolymeric stretches of DNA [8-10], and many of the polymorphisms identified in the previous study [23] were homopolymeric STRs (hSTRs).

This study seeks to design an MPS panel for the MiSeq FGx® consisting of seven highly polymorphic "hotspot" regions in the *C. sativa* chloroplast (*trnK-matK-trnK*, *rps16*, *trnS-trnG*, *ycf3*, *accD-psaI*, *clpP*, and *rpl32-trnL*) to provide additional sequence data, discover isoalleles, and provide a high throughput method for creating a haplotype DNA database for hemp and marijuana samples. This assay could provide important investigative leads for law enforcement agencies investigating marijuana trafficking into and within the United States.

Materials and methods

DNA samples

Hemp samples from Canada were purchased online from Badia Spices Inc. (Doral, FL, USA; N=1) and Navitas Organics (Novato, CA, USA; N=1). USA hemp samples were

purchased from American Hemp Harvest (Boulder, CO, USA; N=4) and The Original Hemp Buds (OR or NY, USA; N=1; strain: Electra). THC-positive marijuana samples from the USA-Mexico border were obtained from U.S. Customs and Border Protection (N=5 from different seizures). Chile marijuana samples were received from the Policia de Investigaciones in southern Chile (N=2 from separate cases). A Chilean medical marijuana sample was provided by collaborators in Chile (N=1; strain: London Cheese). Chloroplast DNA was quantified using a real-time PCR method reported by Houston et al. [30].

Target enrichment

Seven polymorphic "hotspot" regions in the chloroplast of *C. sativa* (*trnK-matK-trnK*, *rps16*, *trnS-trnG*, *ycf3*, *accD-psaI*, *clpP*, and *rpl32-trnL*) were amplified in single PCR [23]. Primers were designed using Primer3 [31] and checked for specificity using the Primer-BLAST tool (NCBI). The optimal annealing temperature for each primer set was determined by gradient PCR as previously described [23]. PCR was carried out on a Veriti™ 96-well Thermal Cycler (Thermo Fisher Scientific, Waltham, MA) using the TaKaRa LA PCR™ Kit Ver.2.1 (TaKaRa Bio Inc., Kusatsu, Shiga, Japan). Reactions consisted of 0.25 µL TaKaRa LA Taq polymerase, 2.5 µL 10X LA PCR Buffer II (Mg^{2+} free), 2.5 µL 25 µM $MgCl_2$, 4 µL dNTP mix, 2.5 µL 2 µM primer mix (Table 4.1), 4 µL template DNA (20 pg/ µL), and 9.25 µL water. Cycling conditions consisted of a 2 min initial denaturation at 94 °C; followed by 30 cycles of 98 °C for 10 s, the optimal annealing temperature (Table 4.1) for 1 min, and 68 °C for 2 min; and a 10 min final extension at 72 °C. A negative template control (NTC) was included.

Table 4.1 Primer sequences and optimal annealing temperatures for single PCR reactions

Region	Forward primer	Reverse primer	Product size (bp)	Ta (°C)
trnK-matK-trnK	ACGAGCCAAAGTTTTAACACAGG	TCGGCTTTTAAGTGCGGCTA	2111	69
rps16	AGAAAAGGGTGTAGACGAACG	TCGTTTCTCGGAGGCAAGAAT	1398	66
trnS-trnG	TCTAATGATCCGGGGCGTAA	TGCATTCAAAACGACCTGC	1668	66
ycf3	ACGGCTCAGCAGTCAAGTTC	TTCGAAATTCATGAAAGGCCCC	2095	68
accD-psaI	GGCTGTTCAAACAGGTACAGG	TGCCGGAAATACTAAGCCCA	1424	68
clpP	TAAATTCCCCTGTCGGTGCC	ATGCCTATTGGTGTTCCAAAAGTA	1984	66
rpl32-trnL	GGAAAAACCCACATACGGCG	TAACACTCGGCGCGGTTATT	1964	69

Following amplification, samples were quantified using the Qubit™ dsDNA HS Assay Kit™ (Invitrogen, Carlsbad, CA) on a Qubit™ 2.0 fluorometer (Invitrogen). The seven PCR targets were then diluted and pooled to a final concentration of 1 ng/μL. A 25 μL aliquot of the mixed PCR products was moved to a new tube and incubated with 2 μL exonuclease I (10 U/μL, Invitrogen) at 35 °C for 72 min and 75 °C for 15 min to remove excess primers.

Library preparation and sequencing

The Nextera XT DNA Library Preparation Kit (Illumina, San Diego, CA) was used for tagmentation and indexing of libraries. A 1 ng input of amplified, cleaned DNA was used (1 μL), and libraries were prepared according to the manufacturer's instructions with the exception that 15 cycles of PCR were used for indexing instead of 12 cycles to ensure adequate library quantity [32]. Sample libraries were checked on the 2100 Bioanalyzer

(Agilent, Santa Clara, CA) using the High Sensitivity DNA kit (Agilent). Following manual normalization, libraries were pooled in equal amounts and denatured according to the manufacturer's instructions, resulting in a 15 pM library [33]. A denatured PhiX control (20 pM) was spiked in at 5% volume. Sequencing on the MiSeq FGx® (Verogen, San Diego, CA) was performed using the MiSeq FGx® Reagent Micro Kit (Verogen).

Data analysis

Data analysis was performed using the MiSeq Reporter and Real Time Analysis software installed on the instrument. Sequences were compared to a *C. sativa* chloroplast reference genome (Yoruba Nigeria cultivar, GenBank accession NC_027223.1). Variants to the reference genome were reported in a variant call file, and .bam files were viewed in Integrative Genomics Viewer (*IGV*) 2.8.0 [34].

Results and discussion

Sequencing metrics

Sequencing quality was high, with 81.2% of base calls having a quality score of 30 or higher, which indicates at least 99.9% accuracy at each base. The error rate in the PhiX control was below 3%. The yield was 1.48 Mb, cluster density was 1045 ± 26 K/mm^2, and 99.33% ± 0.26% of clusters passed the filter. Phasing/prephasing rate was 0.139/0.036. Coverage varied within each of the amplicons. The *trnK-matK-trnK* and *rps16* amplicons consistently had the lowest coverage, and *ycf3* consistently had the highest. The clusters passing filter (PF) and clusters aligned to the reference genome for each sample are shown in Table 4.2.

Table 4.2 Sequencing coverage for each sample

Sample name	Cluster PF	Cluster align
H2-4	354478	340754/338903
H3-3	421534	397646/396664
H5-4	499828	468880/466048
NT H5-1	324939	318659/316935
NT H5-2	329130	316440/314649
NT H5-4	162214	145299/144912
H8-1	430363	410957/409288
10-A1	247712	234158/233555
12-A7	295382	285361/283439
16-B1	259612	251007/249592
21-A16	312004	300208/298747
35	286421	275212/272900
41	345513	335350/333284
MedMJ10	301630	282215/280864
NTC	18667	144/191

Due to the homopolymeric nature of many of the repeat units (hSTRs), forward and reverse stutter was observed. However, based on the sequence coverage, confident allele calls were made. Some reads aligned to portions of the genome outside of the hotspot regions and were not interpreted, and additionally, several reads misaligned within hotspot regions. These misalignments had low read depth and did not affect interpretation of sequencing results.

Sequence data

There were 33 known polymorphisms within the "hotspot" amplicons [23], and 16 more were discovered, bringing the total to 49 polymorphisms. Newly discovered polymorphisms are indicated by asterisks in Tables 4.3-4.9. The *trnS-trnG* and *rpl32-trnL* hotspots had the most polymorphisms with 12 and 11, respectively, and *ycf3* had the least with only three. Analysis of a higher number of polymorphic loci is expected to show increased differences between samples from different populations. The genotypes at each

locus for 14 samples are displayed in Tables 4.3-4.9. Samples H8-1 (USA hemp) and 12-A7 (USA-Mexico marijuana) were the only two that produced the same haplotype. These two samples were analyzed previously with CE-based methods and also shown to have the same haplotype. Samples 10-A1 (USA-Mexico marijuana) and MedMJ 10 (Chile medical marijuana) were also shown to have the same haplotype in previous studies. However, using MPS, they were distinguished by their sequences at *rpl32-trnL* hSTR3; 10-A1 has a 6 bp allele with the sequence TAAAAA, and MedMJ10 has a 6 bp allele with the sequence AAAAAA. Since they are the same size, these two isoalleles could not be distinguished in previous CE-based studies [23].

Table 4.3 Genotypes for polymorphisms in the *trnK-matK-trnK* region (2,233-4,337 bp)

Sample name	trnK-matK INDEL (GAATAC)	trnK/matK SNP (C/T)	*SNP* (T/C)*	*SNP* (A/C)*	trnK-matK STR1 (A)	trnK-matK STR2 (T)
Start location (bp)	2984	3258	3561	3752	3809	4109
Yoruba Nigeria (NC_027223.1)	Absent	C	T	A	11	10
H2-4	Absent	C	T	A	11	10
H3-3	Absent	C	T	A	11	10
H5-4	Present	C	T	A	11	10
NT H5-1	Absent	C	T	A	11	10
NT H5-2	Present	C	T	A	11	10
NT H5-4	Present	C	C	C	15	9
H8-1	Present	C	T	A	11	10
10-A1	Present	C	T	A	11	10
12-A7	Present	C	T	A	11	10
16-B1	Present	C	C	C	13	9
21-A16	Present	C	T	A	11	10
35	Present	C	C	C	15	9
41	Present	C	C	C	14	9
MedMJ10	Present	C	T	A	11	10

* indicates new polymorphisms discovered in this study
Italics indicate polymorphisms not analyzed previously by CE

Table 4.4 Genotypes for polymorphisms in the *rps16* region (4,803-6,201 bp)

Sample name	*INDEL* (AAAGTA)*	*trnK-rps16 hSTR (A)*	rps16 SNP1 (A/G)	rps16 SNP2 (A/C)	rps16 hSTR (C)	rps16 SNP3 (G/A)
Start location (bp)	4911	5197	5303	5517	5518	6103
Yoruba Nigeria (NC_027223.1)	Absent	9	A	A	12	G
H2-4	Absent	9	A	A	11	G
H3-3	Absent	9	A	A	11	G
H5-4	Absent	10	A	A	11	A
NT H5-1	Absent	9	A	A	11	G
NT H5-2	Absent	10	A	A	11	A
NT H5-4	Present	11	A	A	13	G
H8-1	Absent	10	A	A	11	A
10-A1	Absent	10	A	A	11	A
12-A7	Absent	10	A	A	11	A
16-B1	Present	11	A	A	14	G
21-A16	Absent	10	A	A	11	A
35	Present	11	A	A	13	G
41	Present	11	A	A	14	G
MedMJ10	Absent	10	A	A	11	A

* indicates new polymorphisms discovered in this study

Italics indicate polymorphisms not analyzed previously by CE

Table 4.5 Genotypes for polymorphisms in the *trnS-trnG* region (8,300-9,967 bp). Discrepancies between the genotypes obtained by MPS and CE are indicated in parentheses

Sample name	*psbI-trnS hSTR (T)*	*hSTR* (A)*	trnS-trnG SNP1 (T/C)	trnS-trnG SNP2 (A/T)	*hSTR* (T)*	*INDEL* (CAATAT)*	*SNP* (A/T)*	trnS-trnG hSTR1	trnS-trnG hSTR2 (T)	*hSTR* (A)*	trnS-trnG SNP3 (T/A)	*SNP* (A/C)*
Start location (bp)	8359	8585	8595	8684	8685	8731	8877	9018	9104	9149	9360	9463
Yoruba Nigeria (NC_027223.1)	7	10	T	A	9	Absent	A	15	11	12	A	A
H2-4	7	10	T	A	9	Absent	A	15	11	12	A	A
H3-3	7	10	T	A	9	Absent	A	15	11	12	A	A
H5-4	8	10	C	T	9	Absent	A	15	11	12	T	A
NT H5-1	7	10	T	A	9	Absent	A	15	11	12	T	A
NT H5-2	8	10	C	T	9	Absent	A	15	11	12	T	A
NT H5-4	8	12	T	A	8	Absent	A	16	12 (CE: 11)	11	T	C
H8-1	8	10	C	T	9	Absent	A	15	11	12	T	A
10-A1	8	10	C	T	9	Absent	A	15	11	12	T	A
12-A7	8	10	C	T	9	Absent	A	15	11	12	T	A
16-B1	8	10	T	A	8	Absent	T	16	12 (CE: 11)	11	T	C
21-A16	8	10	C	T	9	Absent	A	15	11	12	T	A
35	8	10	T	A	8	Present	A	16	12 (CE: 11)	11	T	C
41	8	10	T	A	8	Absent	T	16	12 (CE: 11)	11	T	C
MedMJ10	8	10	C	T	9	Absent	A	15	11	12	T	A

* indicates new polymorphisms discovered in this study

Italics indicate polymorphisms not analyzed previously by CE

Table 4.6 Genotypes for polymorphisms in the *ycf3* region (43,383-45,478 bp)

Sample name	ycf3 hSTR1 (T)	ycf3 hSTR2 (T)	ycf hSTR3 (A)
Start location (bp)	43454	44007	45034
Yoruba Nigeria (NC_027223.1)	11	11	10
H2-4	11	11	10
H3-3	11	11	10
H5-4	10	12	10
NT H5-1	11	11	10
NT H5-2	9	12	10
NT H5-4	9	10	10
H8-1	10	12	10
10-A1	10	12	10
12-A7	10	12	10
16-B1	9	10	10
21-A16	10	12	10
35	9	10	10
41	9	10	10
MedMJ10	10	12	10

Table 4.7 Genotypes for polymorphisms in the *accD-psaI* region (58,173-59,596 bp)

Sample name	accD-psal SNP1 (A/G)	accD-psal SNP3 (A/C)	SNP* (G/A)	SNP* (T/C)	accD-psal SNP2 (T/G)	accD-psal STR (A)
Start location (bp)	58833	58851	58921	58924	58981	59141
Yoruba Nigeria (NC_027223.1)	A	A	G	T	T	10
H2-4	A	A	G	T	T	10
H3-3	A	A	G	T	T	10
H5-4	G	A	G	T	G	11
NT H5-1	A	A	G	T	T	10
NT H5-2	G	A	G	T	G	11
NT H5-4	G	C	G	T	T	10
H8-1	G	A	G	T	G	11
10-A1	G	A	G	T	G	11
12-A7	G	A	G	T	G	11
16-B1	G	C	A	C	T	10
21-A16	G	A	G	T	G	11
35	G	C	A	C	T	10
41	G	C	A	C	T	10
MedMJ10	G	A	G	T	G	11

* indicates new polymorphisms discovered in this study

Italics indicate polymorphisms not analyzed previously by CE

Table 4.8 Genotypes for polymorphisms in the *clpP* region (70,502-72,486 bp)

Sample name	clpP hSTR1 (A)	clpP hSTR2	clpP hSTR3 (T)	clpP hSTR4 (T)	clpP INDEL (TTCAATTTA)
Start location (bp)	70912	70981	71663	72016	72028
Yoruba Nigeria (NC_027223.1)	11	TATTT	14	13	Absent
H2-4	11	TATTT	14	12	Absent
H3-3	11	TATTT	14	12	Absent
H5-4	10	TTTT	15	12	Present
NT H5-1	11	TTTT	14	12	Absent
NT H5-2	10	TTTT	15	11	Present
NT H5-4	11	TTTT	12	11	Absent
H8-1	10	TTTT	14	11	Present
10-A1	10	TTTT	15	11	Present
12-A7	10	TTTT	14	11	Present
16-B1	12	TTTT	11	10	Absent
21-A16	10	TTTT	15	11	Present
35	12	TTTT	11	10	Absent
41	12	TTTT	11	10	Absent
MedMJ10	10	TTTT	15	11	Present

Table 4.9 Genotypes for polymorphisms in the *rpl32-trnL* region (112,153-114,100 bp). Discrepancies between the genotypes obtained by MPS and CE are indicated in parentheses

Sample name	hSTR* (T)	hSTR* (A)	ndhF-rpl32 INDEL (Variable)	rpl32-trnL hSTR1 (A)	SNP* (A/T)	rpl32-trnL hSTR2	hSTR* (A)	rpl32-trnL SNP	rpl32-trnL hSTR3	rpl32-trnL INDEL	SNP* (A/G)
Start location (bp)	112294	112551	112562	112830	112900	112961	113017	113044	113149	113246	113459
Yoruba Nigeria (NC_027223.1)	8	12	Absent	12	A	11	10	A	8 (7T+A)	Present	A
H2-4	8	12	Absent	12	A	11	10	A	9(8T+A)	Absent	A
H3-3	8	12	Absent	12	A	11	10	A	8 (7T+A)	Absent	A
H5-4	8	13	GAATTG+10A	11	A	11	10	C	6 (6A)	Present	A
NT H5-1	8	11	Absent	12	A	12	10	A	6 (2T+4A)	Present	A
NT H5-2	8	13	GAATTG+11A	11	A	11	10	C	6 (6A)	Present	A
NT H5-4	8	11	GAATTG+10A	11	T	11 (CE:12)	11	A	11 (7T+4A)	Present	A
H8-1	8	13	GAATTG+12A	11	A	11	10	C	6 (6A)	Present	A
10-A1	8	13	GAATTG+11A	11	A	11	10	C	6 (T+5A)	Present	A
12-A7	8	13	GAATTG+12A	11	A	11	10	C	6 (6A)	Present	A
16-B1	7	11	GAATTG+10A	11	T	11 (CE:12)	11	A	14 (6T+8A)	Present	G
21-A16	8	13	GAATTG+11A	10	A	11	10	C	6 (6A)	Present	A
35	7	10	GAATTG+10A	11	T	11 (CE:12)	11	A	15 (11T+4A)	Present	G
41	7	10	GAATTG+10A	11	T	11 (CE:12)	11	A	14 (6T+8A)	Present	G
MedMJ10	8	13	GAATTG+11A	11	A	11	10	C	6 (6A)	Present	A

* indicates new polymorphisms discovered in this study

Italics indicate polymorphisms not analyzed previously by CE

Concordance

Previously, 30 polymorphisms within the seven hotspot regions were analyzed in our laboratory using CE-based methods (unpublished data and [23]). The genotypes obtained by MPS are fully concordant with the CE-based genotypes with exceptions at the *trnS-trnG* hSTR2 and *rpl32-trnL* hSTR2 loci. In four samples (NT H5-4, 16-B1, 35, and 41), the sequence genotypes at both loci appeared to be off by 1 bp from the CE-based genotypes (indicated in Tables 4.5 and 4.9). The sequence data showed that the CE fragment assays for both of these loci amplified regions containing the locus of interest as well as an additional hSTR locus that was unknown at the time. Variation at this new hSTR locus explains the discrepancy between the CE and sequencing genotypes for all four samples at both loci.

Previously, *rps16* hSTR and *clpP hSTR3* alleles were reported by their bp size due to unclear results using Sanger sequencing (unpublished data). However, as expected, the MPS method was able to elucidate the sequences of all alleles, and the fragment size and sequence data were determined to be concordant. Additionally, the 11 and 15 alleles at the *rpl32-trnL* hSTR3 locus were unable to be confirmed by Sanger sequencing in the previous study [23], but MPS was able to provide sequence data for these alleles.

Conclusions

The MPS assay developed in this study provided an effective method for genotyping seven chloroplast regions previously shown to be informative for determining crop type and biogeographical origin of *C. sativa*. It provided multiple benefits over previous CE-based assays, including simultaneous analysis of all 7 regions in multiple samples, higher confidence for haplotype calls, and better discrimination through

sequencing more polymorphisms and identifying isoalleles. A preliminary set of 14 samples was sequenced, and a total of 49 polymorphisms were observed, 16 of which have not been previously published. The sequence data were concordant with CE genotypes from previous studies. The high throughput ability of MPS will allow for the creation of a worldwide haplotype database of *C. sativa* samples. A comprehensive database is necessary for determining the biogeographical origin and crop type of unknown samples.

References

1. Moorthie S, Mattocks CJ, Wright CF (2011) Review of massively parallel DNA sequencing technologies. HUGO J 5: 1-12
2. Bruijns B, Tiggelaar R, Gardeniers H (2018) Massively parallel sequencing techniques for forensics: a review. Electrophoresis 39: 2642-54
3. Lazaruk K, Walsh PS, Oaks F et al (1998) Genotyping of forensic short tandem repeat (STR) systems based on sizing precision in a capillary electrophoresis instrument. Electrophoresis 19: 86-93
4. Parson W, Ballard D, Budowle B et al (2016) Massively parallel sequencing of forensic STRs: considerations of the DNA commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements. Forensic Sci Int Genet 22: 54-63
5. Quail MA, Smith M, Coupland P et al (2012) A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC Genom 13: 341
6. Rothberg JM, Hinz W, Rearick TM et al (2011) An integrated semiconductor device enabling non-optical genome sequencing. Nature 475: 348-52

7. Bentley DR, Balasubramanian S, Swerdlow HP et al (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456: 53-9
8. Duke J, Lind C, Mackiewicz K et al (2015) Towards allele-level human leucocyte antigens genotyping–assessing two next-generation sequencing platforms: Ion Torrent Personal Genome Machine and Illumina MiSeq. Int J Immunogenet 42: 346-58
9. McElhoe JA, Holland MM, Makova KD et al (2014) Development and assessment of an optimized next-generation DNA sequencing approach for the mtgenome using the Illumina MiSeq. Forensic Sci Int Genet 13: 20-9
10. Loman NJ, Misra RV, Dallman TJ et al (2012) Performance comparison of benchtop high-throughput sequencing platforms. Nature Biotechnol 30: 434
11. Wang Z, Zhou D, Wang H et al (2017) Massively parallel sequencing of 32 forensic markers using the Precision ID GlobalFiler™ NGS STR Panel and the Ion PGM™ System. Forensic Sci Int Genet 31: 126-34
12. Kim EH, Lee HY, Yang IS, Jung S-E, Yang WI, Shin K-J (2016) Massively parallel sequencing of 17 commonly used forensic autosomal STRs and amelogenin with small amplicons. Forensic Sci Int Genet 22: 1-7
13. Guo F, Yu J, Zhang L, Li J (2017) Massively parallel sequencing of forensic STRs and SNPs using the Illumina® ForenSeq™ DNA signature prep kit on the MiSeq FGx™ forensic genomics system. Forensic Sci Int Genet 31: 135-48
14. Seo SB, King JL, Warshauer DH, Davis CP, Ge J, Budowle B (2013) Single nucleotide polymorphism typing with massively parallel sequencing for human identification. Int J Legal Med 127: 1079-86

15. Eduardoff M, Santos C, De La Puente M et al (2015) Inter-laboratory evaluation of SNP-based forensic identification by massively parallel sequencing using the Ion PGM™. Forensic Sci Int Genet 17: 110-21

16. Davis C, Peters D, Warshauer D, King J, Budowle B (2015) Sequencing the hypervariable regions of human mitochondrial DNA using massively parallel sequencing: enhanced data acquisition for DNA samples encountered in forensic testing. Legal Med 17: 123-7

17. Mehta B, Daniel R, Phillips C, Doyle S, Elvidge G, McNevin D (2016) Massively parallel sequencing of customised forensically informative SNP panels on the MiSeq. Electrophoresis 37: 2832-40

18. Budowle B, Schmedes SE, Wendt FR (2017) Increasing the reach of forensic genetics with massively parallel sequencing. Forensic Sci Med Pathol 13: 342-9

19. Houston R, Mayes C, King JL, Hughes-Stamm S, Gangitano D (2018) Massively parallel sequencing of 12 autosomal STRs in Cannabis sativa. Electrophoresis 39: 2906-11

20. Shokralla S, Porter TM, Gibson JF et al (2015) Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform. Sci Rep 5: 9687

21. Sucher NJ, Hennell JR, Carles MC. (2012) DNA fingerprinting, DNA barcoding, and next generation sequencing technology in plants. Plant DNA Fingerprinting and Barcoding. Springer. pp. 13-22

22. Parks M, Cronn R, Liston A (2009) Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. Bmc Biol 7: 84

23. Roman MG, Gangitano D, Houston R (2019) Characterization of new chloroplast markers to determine biogeographical origin and crop type of Cannabis sativa. Int J Legal Med 133: 1721-32

24. Oh H, Seo B, Lee S et al (2016) Two complete chloroplast genome sequences of Cannabis sativa varieties. Mitochondrial DNA Part A 27: 2835-7

25. Vergara D, White KH, Keepers KG, Kane NC (2016) The complete chloroplast genomes of Cannabis sativa and Humulus lupulus. Mitochondrial DNA Part A 27: 3793-4

26. Matielo CBDO, Lemos RPM, Sarzi DS et al (2020) Whole Plastome Sequences of Two Drug-Type Cannabis: Insights Into the Use of Plastid in Forensic Analyses. J Forensic Sci 65: 259-65

27. Anderson S, Bankier AT, Barrell BG et al (1981) Sequence and organization of the human mitochondrial genome. Nature 290: 457-65

28. Miller KW, Budowle B (2001) A compendium of human mitochondrial DNA control region: development of an international standard forensic database. Croat Med J 42: 315-27

29. Ingman M, Gyllensten U (2006) mtDB: Human Mitochondrial Genome Database, a resource for population genetics and medical sciences. Nucleic Acids Res 34: D749-D51

30. Houston R, Birck M, LaRue B, Hughes-Stamm S, Gangitano D (2018) Nuclear, chloroplast, and mitochondrial data of a US cannabis DNA database. Int J Legal Med: 1-13
31. Koressaar T, Remm M (2007) Enhancements and modifications of primer design program Primer3. Bioinformatics 23: 1289-91
32. Illumina (2018) Nextera XT DNA Library Prep Kit Reference Guide. San Diego, CA
33. Illumina (2019) MiSeq System Denature and Dilute Libraries Guide. San Diego, CA
34. Robinson JT, Thorvaldsdóttir H, Winckler W et al (2011) Integrative genomics viewer. Nature Biotechnol 29: 24-6

CHAPTER V

Evaluation of tetrahydrocannabinolic acid (THCA) synthase polymorphisms for distinguishing between marijuana and hemp

---

This dissertation follows the style and format of *International Journal of Legal Medicine*.

Roman, MG, Cheng, YC, Kerrigan, S, Houston, R (2020) Evaluation of tetrahydrocannabinolic acid (THCA) synthase polymorphisms for distinguishing between marijuana and hemp. In preparation

Abstract

The Controlled Substances Act (CSA) classifies marijuana (*Cannabis sativa*) as a Schedule I illicit drug. However, the recent Agriculture Improvement Act of 2018 (Farm Bill) removed hemp (*C. sativa* with less than 0.3% of the psychoactive chemical tetrahydrocannabinol, THC) from the definition of marijuana in the CSA, making it a legal crop. As a result, many hemp products are now available, including strains of hemp buds for smoking that are high in other cannabinoids (usually cannabidiol, CBD, or cannabigerol, CBG). The genetic inheritance of chemical phenotype (chemotype) has been widely studied, with the gene for tetrahydrocannabinolic acid (THCA) synthase at the forefront. Previous studies have speculated that two forms of the gene exist, one that produces an active enzyme (present in marijuana) and one that produces an inactive enzyme (present in hemp). A DNA analysis method is desirable for determining crop type in sample types inconducive to chemical analysis, such as immature crops, trace residues, small leaf fragments, seeds, and root material. This study optimized and evaluated a previously reported single nucleotide polymorphism (SNP) assay for determining *C. sativa* crop type (marijuana or hemp). Additionally, the presence or absence of fifteen cannabinoids, including THC and THCA, was reported in fifteen legal hemp flower samples. The SNP assay correctly identified crop type in the majority of samples. However, several hemp seeds, which generally have low levels of cannabinoids, were classified as marijuana, and two strains of legal CBG hemp flowers were classified as marijuana, indicating that factors other than genetic variation of the THCA synthase gene (including gene expression at the time of harvest) should be considered when determining crop type.

Keywords: *Cannabis sativa*, THCA synthase, Chemotype, Single nucleotide polymorphisms, Cannabinoids

Introduction

Cannabis sativa is cultivated world-wide for its use as fiber and oil (hemp) or as a psychoactive drug (marijuana) [1]. It is a flowering plant that contains over 120 cannabinoids, which are present in the plant at various concentrations [2]. In the United States, crops with less than 0.3% delta-9-tetrahydrocannabinol (THC) on a dry weight basis are considered hemp. THC is a psychoactive drug, and marijuana is listed in Schedule I of the federal Controlled Substances Act. It is the most commonly used recreational drug in the U.S. [3], and it is analyzed routinely in forensic laboratories. Traditionally, to identify a substance as marijuana, microscopic identification of cystolithic hairs [4] or instrumental techniques such as gas chromatography-mass spectrometry (GC-MS) are used [5]. Microscopic techniques are presumptive in nature, and instrumental analyses to identify THC are typically qualitative, not quantitative. Since these approaches do not *quantify* THC, they cannot be used to distinguish between marijuana and hemp. Recently, the Agriculture Improvement Act of 2018, referred to as the Farm Bill, legalized the possession of hemp and necessitated that forensic laboratories find ways to distinguish the legal crop (hemp) from the illegal crop (marijuana). Under this law, hemp is defined as any cannabis material with less than 0.3% THC by dry weight [6]. The U.S. Department of Agriculture (USDA) testing guidelines for hemp material dictate that "total THC" concentration must be measured, which means that testing methods must measure THC concentrations post-decarboxylation [6]. Decarboxylation involves the conversion of tetrahydrocannabinolic acid (THCA) to THC, which can be achieved after heating, and is therefore capable of producing THC *in-situ* to the user upon smoking [7].

As a result of the 2018 Farm Bill, strains of cannabis with cannabinoids other than THC have been marketed for their medicinal and healing properties and are widely available to the public through smoke shops and online vendors. Cannabidiol (CBD) is traditionally the most prevalent cannabinoid in hemp and is often the main cannabinoid marketed in hemp flower products at high concentrations (>10% by weight). Another cannabinoid, cannabigerol (CBG), which is a precursor to both THC and CBD, is also increasing in popularity in legal cannabis products. In order to remain compliant with the Farm Bill, these products should be sold with an accompanying certificate of analysis (COA) that reports the THC concentration. Although the USDA guidelines recommend that laboratories performing these analyses are accredited to ISO/IEC 17025 (General Requirements for the Competence of Testing and Calibration Laboratories), this is not required currently. COAs provided by the supplier are also used to report the concentrations of other cannabinoids perceived to be "desirable" to the user.

THC, CBD, and CBG are found in very low quantities in fresh cannabis. Instead, these cannabinoids are present as cannabinolic acids, principally THCA, cannabidiolic acid (CBDA), and cannabigerolic acid (CBGA) [8, 9]. The conversion of these cannabinolic acids to the neutral cannabinoids (THC, CBD, and CBG) occurs via non-enzymatic decarboxylation during heating (e.g., when the plant is smoked), drying, or storage [10]. There are several enzymes involved in the biosynthesis of these cannabinoids; however, the final step is the most studied in regard to chemical phenotype. The precursor CBGA is converted into either THCA (catalyzed by THCA synthase) or CBDA (catalyzed by CBDA synthase) (Fig. 5.1) [8].

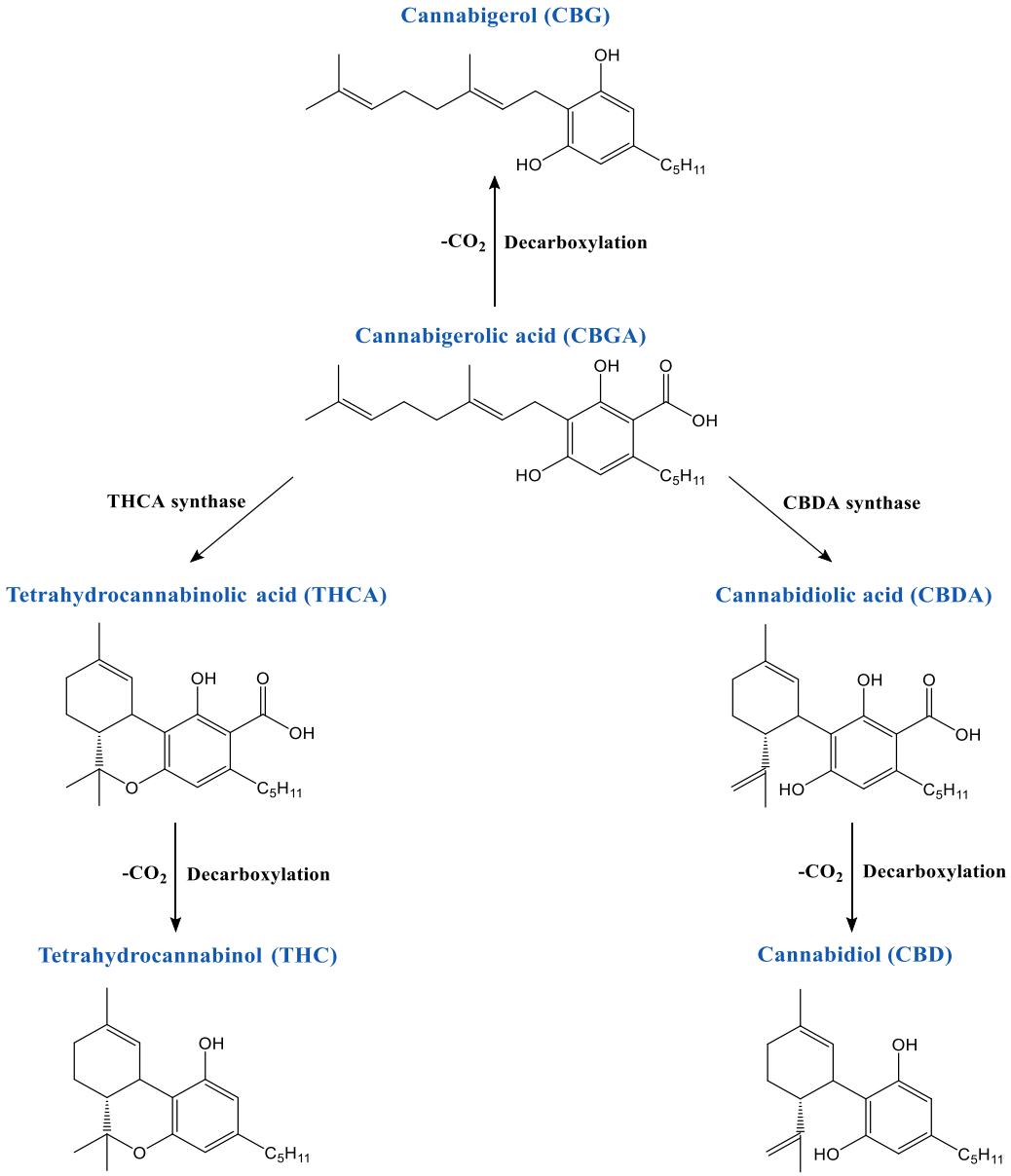


Fig. 5.1 CBGA conversion to THCA, CBDA, and their neutral forms (THC and CBD) through the cannabinoid synthase and decarboxylation pathways

The use of polymorphisms in the THCA and CBDA synthase genes to predict *C. sativa* chemical phenotype (also known as chemotype) has been extensively studied [11-22]. Two popular chemotype inheritance models exist. The first proposes that chemotype is influenced by a single genetic locus, locus B, with two codominant alleles, B_D and B_T [11]. The allele B_D codes for CBDA synthase, and the allele B_T codes for THCA synthase.

Genotypes at the B locus correspond to five chemical chemotypes: chemotype I ('drug type' with a low CBD/THC ratio; B_T/B_T), chemotype II ('intermediate;' B_D/B_T), chemotype III ('fiber type' with a high CBD/THC ratio; B_D/B_D), chemotype IV (high in CBG due to the presence of allele $B_{0,}$ which produces neither THCA nor CBDA, resulting in an accumulation of the precursor CBGA), and chemotype V (lacking any detectable cannabinoids, perhaps due to a gene upstream of the B locus preventing formation of the precursor CBG) [11, 12].

A second genetic model postulates the existence of separate, tightly linked loci for the THCA and CBDA synthase genes [13-21, 23]. In this model, expression or allelic variation at these loci affects enzymatic efficiency and determines chemotype. Due to tight linkage of the two loci, inheritance follows the same pattern as would be expected from single locus inheritance [23].

The gene encoding THCA synthase was first reported by Sirikantamarus et al. [9]. Later, Kojoma et al. [13] reported two variants of the THCA synthase gene: one active, found only in drug-type plants, and one inactive, present in hemp and some drug-type crops. The gene sequences for these isoforms varied by 62 single nucleotide polymorphisms (SNPs), resulting in 37 amino acid substitutions. Rotherham and Harbison [14] designed a SNaPshot™ assay using four of the SNPs to differentiate between drug and non-drug samples, reporting 100% accuracy in their preliminary study. Several other research groups developed assays based on the idea of active versus inactive (or less active) THCA synthase genes [13, 15-18].

However, two recent studies have suggested that no such isoform of the THCA synthase gene exists. Weiblen et al. [19] demonstrated that active THCA synthase was

expressed in both marijuana and hemp and that isoforms of CBDA synthase were responsible for chemotype; i.e., a less active CBDA synthase enzyme, resulting in less competition for the substrate, was responsible for high THC content in chemotype I plants. The authors proposed that the THCA synthase-like sequence detected by Kojoma et al. [13] in hemp was likely another cannabinoid synthase gene. This result was supported in a recent paper by Laverty et al. [24], which identified the "inactive" gene amplified by Kojoma et al. [13] as cannabichromene acid (CBCA) synthase [24].

Studies by Vergara et al. [25] and McKernan et al. [26] suggest that copy number variation of THCA and CBDA synthase genes also contributes to chemotype. Additionally, it is likely that other genes contribute to chemotype (for example, the gene for aromatic prenyltransferase, which catalyzes the production of CBGA, the substrate of both THCA and CBDA synthases) [19, 24, 27], and transcriptional variation occurs in different strains and at different stages of plant development [28].

Current knowledge indicates that inheritance of chemotype is much more complicated than originally expected. Interestingly, assays assuming a simpler model with active versus inactive isoforms of THCA synthase were surprisingly accurate, with researchers reporting complete success in identifying hemp versus marijuana [13-15, 17]. The purpose of this work was to re-evaluate the SNaPshot™ assay reported by Rotherham and Harbison [14] using the current knowledge. The method was applied to hemp samples from the USA and Canada and marijuana samples from USA-Mexico and Chile. In order to achieve the most accurate results, the assay was first optimized to ensure sufficient amplification. If proven to be accurate, this SNaPshot™ assay could provide a simple, rapid method of distinguishing between marijuana and hemp for all botanical samples (e.g., trace

samples and juvenile plants). In addition to genetically predicting the crop type (marijuana or hemp), this study also measures cannabinoid content for purchased legal hemp flowers in order to ensure accuracy of the classification. Lastly, this study provides a preliminary evaluation of the COA reported cannabinoid content for 15 strains of hemp flowers available in the marketplace.

Materials and methods

DNA analysis

Sample collection

Hemp from the USA (N=23) and Canada (N=6), marijuana from the USA-Mexico border (N=11) and Chile (N=3), and medical marijuana from Chile (N=4) were tested. One package of hemp seeds produced in the USA was purchased from American Hemp Harvest (Boulder, CO, USA; N=6). Hemp seeds harvested in Canada was purchased from Manitoba Harvest (Winnipeg, MB, CA; N=4 from two packages), Badia Spices Inc. (Doral, FL, USA; N=1), and Navitas™ Organics (Novato, CA, USA; N=1). Hemp CBD flowers were purchased from The Original Hemp Buds (OR or NY, USA); five strains included Elektra (2 packages), Lifter (2 packages), Sour Space Candy, Special Sauce, and Cascade (N=1 from each package). Additional hemp CBD flowers were purchased from CBD Hemp Direct (Las Vegas, NV, USA); eight strains included Hemp World Haze, Sunset Rd Sherbert #2, Paradise OG, Casino Cookies #2, Durban Potion #2, Tangie, Juicy Fruit #2, and Trophy Wife; N=1 each). CBG hemp flowers were also purchased from CBD Hemp Direct; two strains included Desert Snow and Jazzy CBG; N=1 each. USA-Mexico marijuana was provided by U.S. Customs & Border Protection Houston Laboratory in Houston, TX (Southwest Regional Science Center) (N=11 from 10 cases). Chile marijuana

DNA extracts were donated by The Policia de Investigaciones in the Araucania region of southern Chile (N=3 from 3 cases). Medical marijuana DNA extracts from Chile were obtained from collaborators in Chile (N=4 strains: Amnesia, AK, Lemon Haze, and London Cheese).

DNA extraction and quantification

The DNeasy® Plant Mini kit (QIAGEN, Hilden, Germany) was used for DNA extraction in accordance with the manufacturer's instructions [29]. Extraction from USA-Mexico marijuana was performed onsite at CBP, and extraction from hemp was performed at Sam Houston State University. DNA extracts for the Chilean samples were provided by collaborators.

Quantification of nuclear DNA was achieved on a StepOne™ Real-Time PCR System (Thermo Fisher Scientific, South San Francisco, CA, USA) using the assay reported by Houston et al. [30].

PCR amplification

The THCA synthase gene was amplified using the C2 (forward) and E2 (reverse) primers reported by Rotherham and Harbison [14] (Table 5.1). Gradient PCR was performed to optimize the annealing temperature as previously described [31].

Table 5.1 PCR and SBE primers used

Primer	Primer sequence	Primer length (bp)	Final primer concentration (µM)
F PCR primer C2	CAAACTKGTTGYTGTCCCATC	21	0.2
R PCR primer E2	CGTCTTCTTCCCAGCTGATC	20	0.2
SBE primer 8F	GAGTTGGGTATTAAAAAAACTGATTGCAA AGAATT	35	0.2
SBE primer 9F	CAACCATCTTCTACAGTGGTGTTGTAAATT	30	0.2
SBE primer 16R	TCRACTAGACTATCCACTCCACCA	24	0.1
SBE primer 17R	TACTGTAGTCTTATTCTTCCCATGATTATC TGTAATATTC	40	0.4

PCR reactions consisted of 6.25 µL of Type-it® Microsatellite PCR Master Mix (QIAGEN), 1.25 µL of 2 µM C2 and E2 primer mix (Integrated DNA Technologies, Coralville, IA, USA), 1.25 µL of 5x Q-solution (QIAGEN), and 3.75 µL of DNA sample (0.5 ng). A touchdown PCR method was used to account for the active and inactive forms of THCA synthase amplifying best at different temperatures (inactive at 60 °C and active at 65 °C). Cycling conditions consisted of a 5 min enzyme activation at 95 °C; 3 cycles of 30 s at 95 °C, 90 s at 65 °C, and 30 s at 72°C;27 cycles of 30 s at 95 °C, 90 s at 60 °C, and 30 s at 72 °C; and a 30 min final extension at 60 °C. PCR cleanup was achieved using 5 µL calf intestinal alkaline phosphatase (CIAP) (1 U/µL, Promega Corporation, Madison, WI, USA) and 2 µL exonuclease I (10 U/µL, Invitrogen). Samples were incubated at 37 °C for 90 minutes, followed by an enzyme inactivation at 75 °C for 30 minutes.

Single base extension

Single base extension (SBE) was performed using half reactions of the SNaPshot™ Multiplex Kit (Thermo Fisher Scientific), following the manufacturer's protocol [32]. SBE primer sequences and concentrations reported by Rotherham and Harbison [14] were used (Table 5.1). Following thermal cycling according to the manufacturer's guidelines, a

cleanup step was performed using 1 µL CIAP (1 U/µL, Promega Corporation). Samples were incubated at 37 °C for 90 min and 75 °C for 30 min.

Capillary electrophoresis

Capillary electrophoresis (CE) was carried out using a 3500 Genetic Analyzer (Thermo Fisher Scientific). Run parameters consisted of the following: oven 60 °C; pre-run 15kV, 180 s; injection 1.6 kV, 8 s; run 15 kV, 560s; capillary length 36 cm; polymer POP-4™; and dye set E5. Data were analyzed using GeneMapper ID v.5 (Thermo Fisher Scientific) with an analytical threshold of 100 RFU.

Chemical analysis

Chemicals and reagents

Cannabidivarin (CBDV), cannabichromevarin (CBCV), cannabicitran (CBT), tetrahydrocannabivarin (THCV), cannabivarin (CBV), cannabicyclol (CBL), CBD, CBC, Δ8-tetrahydrocannabinol (Δ8-THC), Δ6a-10a-tetrahydrocannabinol (Δ6a-10a-THC), Δ9-THC, Δ10-THC, CBG, cannabinol (CBN), and Δ9-THC-d3 were obtained from Cayman (Ann Arbor, MI, USA). Methanol (LC-MS grade) was obtained from Avantor (Randor, PA, USA). Exo-tetrahydrocannabinol (exo-THC) and phencyclidine (PCP) were obtained from Cerilliant (Round Rock, TX, USA).

Sample collection

Due to legal restrictions on obtaining cannabis, cannabinoid content could only be determined for the hemp flower samples. Additionally, cannabinoid content was not measured for the hemp seed products as little cannabinoid content is expected in seeds. The same fifteen hemp flower strains from The Original Hemp Buds and CBD Hemp Direct that were used for DNA analysis were also evaluated for their cannabinoid content.

Cannabinoid data was provided for each strain in the form of a COA (Appendix A). All strains reported that the hemp flowers were below 0.3% THC content by dry weight. However, THC + THCA concentrations for five of the strains were greater than 0.3%.

Instrumentation

GC-MS analysis was performed using an Agilent 6890N/5975C GC-MS instrument (Santa Clara, CA, USA) equipped with an Agilent 7683B autosampler using the optimum conditions described below.

Preparation of standards

All stock solutions of standards, except CBL, were prepared at 1 mg/mL in methanol. CBL stock solution was prepared at 1 mg/mL in acetonitrile. A standard mixture of all 15 cannabinoids (CBDV, CBCV, CBT, THCV, CBV, CBL, CBD, CBC, exo-THC, Δ8-THC, Δ9-THC, Δ6a-10a-THC, Δ10-THC, CBG, and CBN) was prepared at 0.1 mg/mL for each analyte. An internal standard (ISTD) solution of Δ9-THC-d3 was prepared in methanol at a concentration of 0.02 mg/mL. Additionally, a positive QC consisting of a 1:1 mixture of Δ9-THC (0.1 mg/mL) and Δ9-THC-d3 (0.02 mg/mL) was analyzed during each batch.

Sample preparation and extraction

Prior to GC-MS analysis, 50 mg of dried cannabis plant material (flower) was weighed, and five mL of methanol was added to the sample, followed by vortex mixing (10 s). The mixture was then allowed to stand for 5 min at room temperature and vortexed for 10 s before the supernatant was filtered through cotton using a disposable Pasteur pipette. The filtered supernatant was then mixed 1:1 with 0.02 mg/mL of internal standard (THC-d3) in the autosampler vial prior to GC-MS analysis.

Gas chromatography-mass spectrometry

Analytes were separated using a DB-5MS GC column (30 m x 0.25 mm x 0.25 μm). The optimized GC-MS parameters were as follows: injector port temperature: 250 °C; injection mode: split (20:1); carrier gas: helium (1.5 mL/min); oven program: initial hold at 210 °C, 30 °C/min for 0.5 min, followed by 35 °C/min for 3 min; and ion source temperature: 230°C. Mass spectral data were acquired using full scan (*m/z* 40 to 550 amu) and selected ion monitoring (SIM) modes; qualitative identification of cannabinoids was made using full scan acquisition, while semi-quantitative estimates of Δ9-THC were made using SIM (*m/z* 314, 231 and 271 for THC, and 317, 234 and 274 for THC-d3). All data were analyzed using Agilent ChemStation Software. Analytes were qualitatively identified based upon acceptable full scan spectra (library match >90%) and retention time within ± 1% of the verified analytical standard. The peak area ratio of Δ9-THC to Δ9-THC-d3 of the positive QC was used to estimate the Δ9-THC concentration in the sample. Determinations were not considered quantitative and were performed solely for verification purposes (against the COA).

Comparison of genetic prediction and chemical data

The cannabinoid concentrations reported in the COAs of the hemp samples were compared to those identified using the GC-MS method. Additionally, predictions of high or low THC content from the genetic data were compared to chemical data and the reported crop type (marijuana or hemp) of each sample.

Results and discussion

DNA analysis

Assay optimization

Gradient PCR performed using a hemp sample showed that the “inactive” form of the THCA synthase gene (now believed to be CBCA synthase) [24] amplified best at 60 °C and was not detected above 63 °C. Initially, amplification was performed at 60 °C; however, it became clear that the “active” form of THCA synthase did not amplify well. The annealing temperature was changed to 65 °C, as reported by Rotherham and Harbison [14]; however, little to no amplification of the “inactive” peaks occurred. The authors instead used a touchdown PCR method to amplify the “active” form in the first three cycles of PCR using an annealing temperature of 65 °C, followed by a decrease to 60 °C in the remaining 27 cycles, which resulted in relatively balanced peaks in heterozygous marijuana samples (Fig. 5.2).

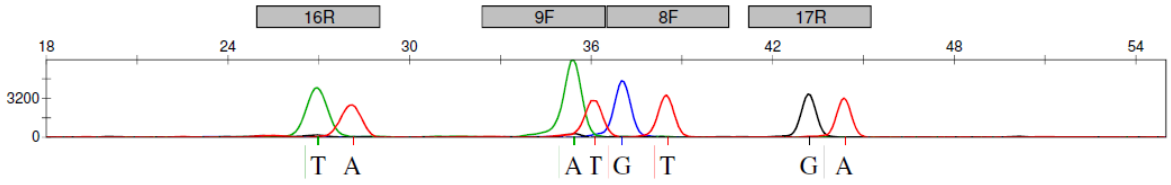


Fig. 5.2 SNaPshot™ result for a marijuana sample heterozygous at the THCA synthase locus

Genotyping

Table 5.2 SNaPshot™ results for THCA synthase polymorphisms. Results are reported as "Active" for the presence of only active THCA synthase extension products, "Inactive" for the presence of only active THCA synthase extension products, or "Heterozygous" for the presence of both active and inactive THCA synthase extension products. Genotypes that are considered incorrect by the Kojoma et al. model are italicized

Sample Name	Case/Strain	Active/Inactive
USA-Mexican Marijuana		
1-D1	Case 1	Heterozygous*
4-J5	Case 4	Heterozygous*
6-D1	Case 6	Heterozygous
9-B2	Case 9	Heterozygous*
11-D1	Case 11	Heterozygous*
13-D3	Case 13	Heterozygous*
16-B1	Case 16	Heterozygous*
18-A8	Case 18	Heterozygous*
20-A1	Case 20	Heterozygous
21-A7	Case 21	Heterozygous*
21-A27	Case 21	*Inactive*
Chilean Marijuana		
35	Case 7	Active
41	Case 9	Heterozygous
50	Case 10	Heterozygous
Chilean Medical Marijuana		
Med 1	Amnesia	Active
Med 4	AK	Active
Med 7	Lemon Haze	Active
Med 10	London Cheese	Active
Canadian Hemp		
H1-2 (seed)	Manitoba Harvest	*Heterozygous*†
H1-3 (seed)	Manitoba Harvest	Inactive
H2-4 (seed)	Badia	Inactive
H3-3 (seed)	Navitas Organics	Inactive
H4-5 (seed)	Manitoba Harvest	Inactive
Hemp 1.C (seed)	Manitoba Harvest	Inactive
USA Hemp		
H5-1 nt (seed)	American Hemp Harvest	Inactive
H5-2 nt (seed)	American Hemp Harvest	*Active*
H5-4 nt (seed)	American Hemp Harvest	*Active*
H5-5 nt (seed)	American Hemp Harvest	*Active*

(continued)

Sample Name	Case/Strain	Active/Inactive
H5-4 (seed)	American Hemp Harvest	*Active*
H5-5 (seed)	American Hemp Harvest	*Active*
H6-2 (flower)	CBD Elektra	Inactive
H7-2 (flower)	CBD Lifter	Inactive
H8-1 (flower)	CBD Elektra	Inactive
H9-1 (flower)	CBD Lifter	Inactive
H10-1 (flower)	CBD Sour Space Candy	Inactive
H11-1 (flower)	CBD Special Sauce	Inactive
H12-1 (flower)	CBD Cascade	Inactive
H13-1 (flower)	CBD Hemp World Haze	Inactive
H14-1 (flower)	CBD Sunset Rd Sherbert #2	Inactive
H15-1 (flower)	CBD Paradise OG	Inactive
H16-1 (flower)	CBD Casino Cookies #2	Inactive
H17-1 (flower)	CBD Durban Potion #2	Inactive
H18-1 (flower)	CBD Tangie	Inactive
H19-1 (flower)	CBD Juicy Fruit #2	Inactive
H20-1 (flower)	CBD Trophy Wife	Inactive
H21-1 (flower)	CBG Desert Snow	*Active*
H22-1 (flower)	CBG Jazzy CBG	*Active*

* indicates heterozygote peak imbalance with much smaller “active” peaks
† indicates heterozygote peak imbalance with much smaller “inactive” peaks

All samples were successfully amplified and produced extension products at all four loci (Table 5.2). Among Mexican marijuana samples (THC-positive), ten were heterozygous for the active and inactive forms of THCA synthase described by Kojoma et al. [13], consistent with their classification as marijuana according to Kojoma et al. [13] and Rotherham and Harbison [14]. One sample, however, resulted only in inactive extension products, suggesting its classification as hemp. Among Chilean marijuana samples, two were heterozygous and one had only active extension products, confirming their classification as marijuana. The Chilean medical marijuana samples all typed as active, as expected. One Canadian hemp sample was shown to be heterozygous for the active and inactive isoforms of THCA synthase, and the other five were inactive. Five of the six USA hemp samples from American Hemp Harvest (hemp seeds) genotyped as

active, while the remaining sample typed as inactive. All CBD hemp flowers were homozygous for the inactive THCA synthase, as expected; however, both CBG strains genotyped as active.

The model of Kojoma et al. [13] considers samples which have at least one copy of the THCA synthase gene to be marijuana (chemotype I), and samples with no active copies are considered to be hemp (chemotype III). Intermediate samples (chemotype II), CBG samples (chemotype IV), and no cannabinoid samples (chemotype V) are not accounted for in this model. If we classify our results accordingly, one marijuana sample was classified incorrectly as hemp, six hemp seeds were classified as marijuana, and two CBG hemp flowers were incorrectly classified as marijuana (overall error rate of approximately 19%).

Laverty et al. [24] showed that the inactive THCA synthase gene identified by Kojoma et al. [13] was actually the CBCA synthase gene. In fact, according to Laverty et al. [24] and other studies which favored the single locus model, no active or inactive form of the THCA synthase gene exists in chemotype III samples, which are instead homozygous at the B locus (B_D/B_D), producing only CBDA synthase. Given this model, we would still expect all hemp samples to be homozygous for the "inactive" gene identified by Kojoma et al. [13] and all marijuana samples to show "active" THCA synthase peaks, so our hemp and marijuana classification results under this model are unchanged.

Under the model of Weiblen et al. [19], both hemp and marijuana samples have active THCA synthase; therefore, we would expect all our samples to have active peaks. This is certainly not the case for the majority of our hemp samples (21 out of 29 of which typed as homozygous inactive).

Additionally, we need to consider chemotype II, which is not accounted for in Kojoma et al.'s model [13]. Chemotype II plants have an intermediate CBD/THC ratio, ranging from 0.5-3.0 [12]. If we consider the heterozygote samples in this study to be chemotype II, they could be classified as either marijuana or hemp by their CBD/THC ratios; only chemical analysis would make their classification certain. However, United States law does not classify hemp and marijuana based on their CBD/THC ratio; it requires direct measurement of THC. Due to the difficulty or inconclusive results of chemical analysis for some sample types (i.e., immature crops, seeds, and root materials naturally contain low amounts of cannabinoids, even in marijuana; and trace residues and small leaf fragments that may not provide sufficient material for chemical analysis), a DNA method to screen for potential THC content is desirable. The active versus inactive THCA synthase model has shown promising results in some studies [13-15, 17]; however, the current study shows that this method results in both false negatives (classifying marijuana as hemp; 2%) and false positives (classifying hemp as marijuana; 17%). It should be noted; however, that the hemp samples which were classified by their THCA synthase gene as marijuana were seeds (N=6) from a common source and CBG hemp flowers (N=2). Since cannabis seeds of either crop type contain only negligible amounts of THC and are sterilized to prevent germination, these seeds are not scheduled under the Controlled Substances Act. Therefore, even seeds which are genetically capable of producing plants with high THC content (marijuana) can be classified as hemp for legal purposes as long as they are sterilized. Disregarding all seed samples, our error rate is 3% for false negatives (one out of 35 samples) and 6% for false positives (two out of 35 samples). All CBD hemp flowers were classified correctly as hemp, but both CBG flower strains were classified as marijuana.

CBG hemp flowers are chemotype IV, which is not accounted for in this assay. As a note, the marijuana samples used in this study were shown to be THC-positive by the forensic laboratories which donated them; however, the THC amount was not quantified (obtained prior to 2018 Farm Bill).

3.2 Chemical analysis

Fifteen hemp flower strains were extracted and analyzed to determine their cannabinoid content using GC-MS. Fifteen cannabinoids were targeted for identification purposes (Fig. 5.3). The GC-MS results indicated that CBC and CBG were present in all 15 hemp flower strains tested, and CBD was present in all CBD flowers but not in CBG flowers. This was in agreement with COAs provided by the supplier (Table 5.3). A low amount of total THC (<1%) was present in all strains (Table 5.3). However, it should be noted that eight of the fifteen strains reported Δ9-THC and THCA concentrations below the limit of quantification (LOQ). Exo-THC, an impurity formed in the synthesis of Δ9-THC, was identified in all strains (Table 5.3). This cannabinoid was not quantified by either supplier in the COA. According to the COAs, THCV was present in 9 of the 15 hemp samples, but it was not qualitatively identified during our analysis. However, CBDV was detected in 13 of the 15 hemp flower strains with our GC-MS method (Table 5.3). CBDV is the precursor of THCV, and studies have shown that CBDV can be isomerized into THCV under acidic conditions or in plant material [33]. Laboratory methods of analysis by the hemp suppliers are unknown, but if extractions were performed under acidic conditions, the inconsistencies between the COAs and our data could be explained. However, little research has been conducted on the relationship between these two cannabinoids to date. CBT, which is an emerging cannabinoid of interest, was identified in

5 strains using our technique [34]. The relative amounts of CBT in the plant are typically low, and it was not a targeted analyte for either hemp supplier. Δ6a-10a-THC and Δ10-THC are not present in plant material but are often found in e-cigarette or e-liquid formulations [35]. CBV, CBL, CBN, and Δ8-THC were not identified among the 15 hemp samples (Table 5.3).

THC concentrations were estimated from the peak area ratio of Δ9-THC to Δ9-THC-d3 in plant extracts relative to the positive QC (0.1 mg/mL THC or 1% THC by dry weight). Results were not quantitative in nature and used solely for verifying COA data from suppliers. Estimated concentrations were within 30% of supplier estimates for the six strains with reported THC and THCA above the LOQ (Elektra, Lifter, Sour Space Candy, Special Sauce, Cascade, and Jazzy CBG) (Table 5.4).

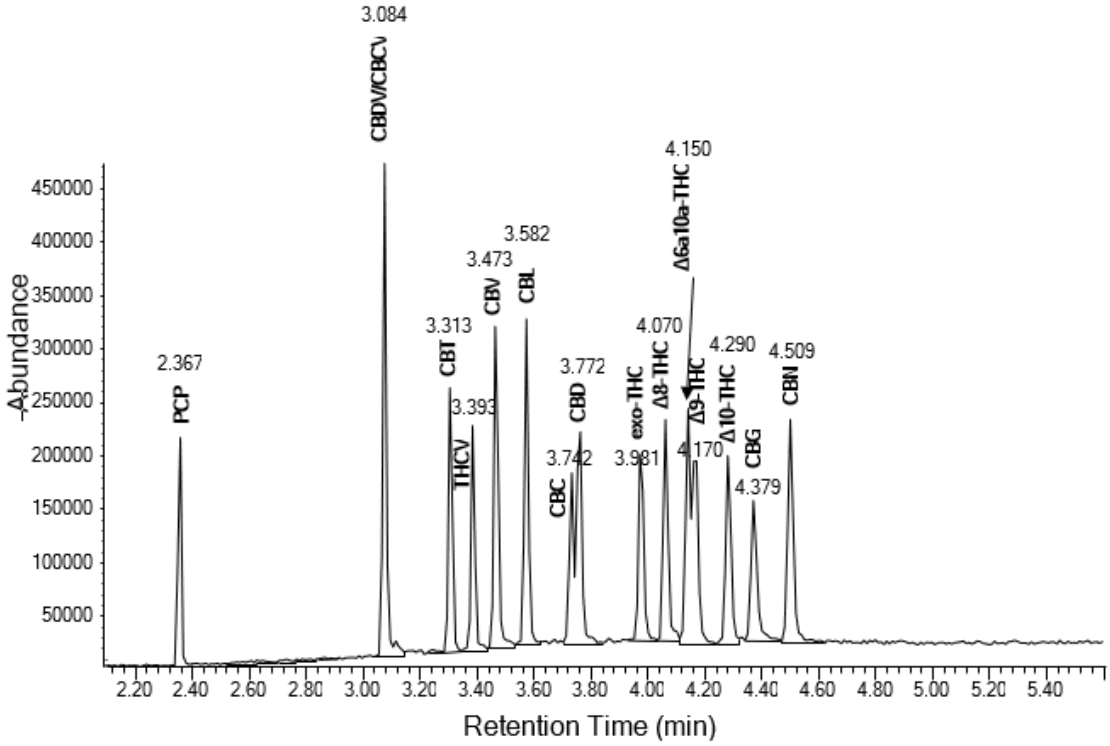


Fig. 5.3 Chromatographic separation of the fifteen cannabinoids using the optimized GC-MC procedure. Phencyclidine (PCP) is shown for reference

Table 5.3 Qualitative identification of seven cannabinoids

Hemp product name	CBDV	CBT	CBC	CBD	CBG	exo-THC	Δ9-THC
CBD Elektra	+	-	+	+	+	+	+
CBD Lifter	+	+	+	+	+	+	+
CBD Sour Space Candy	+	+	+	+	+	+	+
CBD Special Sauce	+	+	+	+	+	+	+
CBD Cascade	+	+	+	+	+	+	+
CBD Hemp World Haze	+	-	+	+	+	+	+
CBD Sunset Rd Sherbert #2	+	-	+	+	+	+	+
CBD Paradise OG	+	-	+	+	+	+	+
CBD Casino Cookies #2	+	-	+	+	+	+	+
CBD Durban Potion #2	+	-	+	+	+	+	+
CBD Tangie	+	-	+	+	+	+	+
CBD Juicy Fruit #2	+	-	+	+	+	+	+
CBD Trophy Wife	+	-	+	+	+	+	+
CBG Desert Snow	-	-	+	-	+	-	+
CBG Jazzy CBG	-	-	+	-	+	-	+

+ indicates present; - indicates not present or below the limit of detection.

Table 5.4 Summary of cannabinoid concentrations from COAs, estimated total THC content based upon GC-MS analysis and THCA synthase activity

Strain name	Certificate of analysis						Analyzed in this study	
	CBD	THC	THCA	Total THC*	CBG+ CBGA	Total Cannabinoid	THC estimate (%)	THCA synthase genotype
CBD Elektra	16.60%	0.06%	0.69%	0.67%	0.98%	20.71%	0.68	Inactive
CBD Lifter	15.35%	0.02%	0.75%	0.68%	2.16%	19.33%	0.60	Inactive
CBD Sour Space Candy	12.96%	0.03%	0.61%	0.56%	1.41%	16.82%	0.53	Inactive
CBD Special Sauce	16.61%	0.02%	0.77%	0.70%	2.14%	21.88%	0.57	Inactive
CBD Cascade	13.63%	0.04%	0.72%	0.67%	2.79%	19.12%	0.71	Inactive
CBD Hemp World Haze	20.54%	<LOQ	<LOQ	<LOQ	0.56%	23.98%	0.95	Inactive
CBD Sunset Rd Sherbert #2	19.49%	<LOQ	<LOQ	<LOQ	0.71%	24.20%	0.58	Inactive
CBD Paradise OG	21.81%	<LOQ	<LOQ	<LOQ	0.73%	25.38%	0.62	Inactive
CBD Casino Cookies #2	18.27%	<LOQ	<LOQ	<LOQ	0.71%	22.82%	0.78	Inactive
CBD Durban Potion #2	14.80%	<LOQ	<LOQ	<LOQ	0.65%	17.52%	0.45	Inactive
CBD Tangie	17.12%	<LOQ	<LOQ	<LOQ	0.19%	19.52%	0.72	Inactive
CBD Juicy Fruit #2	18.28%	<LOQ	<LOQ	<LOQ	0.59%	22.68%	0.84	Inactive
CBD Trophy Wife	17.26%	<LOQ	<LOQ	<LOQ	0.48%	19.86%	0.68	Inactive
CBG Desert Snow	<LOQ	<LOQ	0.24%	0.21%	21.83%	22.24%	0.25	Active
CBG Jazzy CBG	<LOQ	<LOQ	0.24%	0.21%	19.72%	20.143%	0.16	Active

*Total THC from COA calculated: Total THC = THCA * 0.877 + Δ9-THC

Combined genetic and chemical data

As expected, all fifteen hemp strains tested had low levels of THC, and the CBD strains had a corresponding "inactive" THCA synthase genotype (Table 5.4). Since the THCA synthase enzyme is inactive in CBD strains, more of the precursor cannabinoid (CBG) is converted to CBD, resulting in the high levels reported in the COA (12.96-21.81%; Table 5.4). The CBG strains, which had no detectable CBD (Table 5.3) and low amounts of THC, typed as active for THCA synthase. This result may seem contradictory, but other factors also influence cannabinoid content. Expression of THCA and CBDA synthases is variable throughout plant life, and THCA and CBDA concentrations increase over time while flowers develop [28]. Therefore, harvesting the plant early may result in high CBG levels and low THC and CBD levels, which is consistent with the concentrations observed for the two CBG strains.

The COA and qualitative chemical results in this study identify the presence of measurable quantities of THC in hemp. Given the two competing genetic models (single versus dual locus coding of THCA synthase and CBDA synthase), the model of separate but tightly linked genes makes more sense given that under the single locus model, hemp (chemotype III; B_D/B_D) should not produce any form of the THCA synthase enzyme. Under the dual locus model, THCA synthase is still produced, but it is inactive, or more accurately, much less active than the THCA synthase in marijuana, resulting in very low levels of THC.

Conclusions

Under the 2018 Farm Bill, forensic laboratories in the United States are charged with differentiating between marijuana (>0.3% THC) and hemp (<0.3% THC). An accurate

DNA method for classifying samples would be advantageous for sample types that are unsuitable for chemical analysis (i.e., immature crops, seeds, and root material, which typically have low cannabinoid content and are not representative of other parts of the plant; or trace residues and small leaf fragments, which may have an insufficient amount of material for chemical analysis). Several chemotype classification models have been suggested based on the genotyping of polymorphisms in the THCA synthase gene. The SNaPshot™ assay reported by Rotherham and Harbison [14] was evaluated here. It was found to yield both false negatives (one sample from the current study; 2%) and false positives (especially for hemp seeds and CBG hemp flowers; 17%). This method also does not account for chemotype II (intermediate CBD/THC ratio) or chemotype IV (high in CBG and low in THC) plants. Chemical data obtained by GC-MS was used to confirm the presence of common phytocannabinoids in hemp flowers and was consistent with the data on the COA. Factors other than genetics play a role in cannabinoid content. For example, in CBG strains, early harvesting while THCA and CBDA synthase gene expression levels are low may be responsible for low THC and high CBG levels despite the samples having an “active” THCA synthase gene.

References

1. Werf H, Mathussen E, Haverkort A (1996) The potential of hemp (*Cannabis sativa* L.) for sustainable fibre production: a crop physiological appraisal. Ann Appl Biol 129: 109-23
2. ElSohly MA, Radwan MM, Gul W, Chandra S, Galal A. (2017) Phytochemistry of Cannabis sativa L. Phytocannabinoids. Springer. pp. 1-36

3. Substance Abuse and Mental Health Services Administration (2018) Key substance use and mental health indicators in the United States: Results from the 2017 National Survey on Drug Use and Health (HHS Publication No. SMA 18-5068, NSDUH Series H-53). Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration Rockville, MD
4. Mitosinka GT, Thornton JI, Hayes TL (1972) The Examination of Cystolithic Hairs of Cannabis and other Plants by means of the Scanning Electron Microscope. J Forensic Sci Soc 12: 521-9
5. Mikes F, Hofmann A, Waser PG (1971) Identification of (−)-Δ9-6a, 10a-trans-tetrahydro-cannabinol and two of its metabolites in rats by use of combination gas chromatography-mass spectrometry and mass fragmentography. Biochem Pharmacol 20: 2469-76
6. U.S. Department of Agriculture (2019) Testing guidelines for Identifying Delta-9 Tetrahydrocannabinol (THC) Concentration in Hemp
7. Dussy FE, Hamberg C, Luginbühl M, Schwerzmann T, Briellmann TA (2005) Isolation of Δ9-THCA-A from hemp and analytical aspects concerning the determination of Δ9-THC in cannabis products. Forensic Sci Int 149: 3-10
8. Taura F, Morimoto S, Shoyama Y, Mechoulam R (1995) First direct evidence for the mechanism of. DELTA. 1-tetrahydrocannabinolic acid biosynthesis. J Am Chem Soc 117: 9766-7
9. Sirikantaramas S, Morimoto S, Shoyama Y et al (2004) The gene controlling marijuana psychoactivity molecular cloning and heterologous expression of Δ1-

tetrahydrocannabinolic acid synthase from Cannabis sativa L. J Biol Chem 279: 39767-74

10. Baker P, Taylor B, Gough T (1981) The tetrahydrocannabinol and tetrahydrocannabinolic acid content of cannabis products. J Pharm Pharmacol 33: 369-72
11. de Meijer EP, Bagatta M, Carboni A et al (2003) The inheritance of chemical phenotype in Cannabis sativa L. Genetics 163: 335-46
12. Pacifico D, Miselli F, Micheler M, Carboni A, Ranalli P, Mandolino G (2006) Genetics and Marker-assisted Selection of the Chemotype in Cannabis sativa L. Mol Breed 17: 257-68
13. Kojoma M, Seki H, Yoshida S, Muranaka T (2006) DNA polymorphisms in the tetrahydrocannabinolic acid (THCA) synthase gene in “drug-type” and “fiber-type” Cannabis sativa L. Forensic Sci Int 159: 132-40
14. Rotherham D, Harbison S (2011) Differentiation of drug and non-drug Cannabis using a single nucleotide polymorphism (SNP) assay. Forensic Sci Int 207: 193-7
15. Cirovic N, Kecmanovic M, Keckarevic D, Markovic MK (2017) Differentiation of Cannabis subspecies by THCA synthase gene analysis using RFLP. J Forensic Legal Med 51: 81-4
16. Cascini F, Farcomeni A, Migliorini D et al (2019) Highly Predictive Genetic Markers Distinguish Drug-Type from Fiber-Type Cannabis sativa L. Plants 8: 496
17. Staginnus C, Zörntlein S, de Meijer E (2014) A PCR marker Linked to a THCA synthase Polymorphism is a Reliable Tool to Discriminate Potentially THC-Rich Plants of C annabis sativa L. J Forensic Sci 59: 919-26

18. Onofri C, de Meijer EP, Mandolino G (2015) Sequence heterogeneity of cannabidiolic-and tetrahydrocannabinolic acid-synthase in Cannabis sativa L. and its relationship with chemical phenotype. Phytochemistry 116: 57-68
19. Weiblen GD, Wenger JP, Craft KJ et al (2015) Gene duplication and divergence affecting drug content in Cannabis sativa. New Phytologist 208: 1241-50
20. Van Bakel H, Stout JM, Cote AG et al (2011) The draft genome and transcriptome of Cannabis sativa. Genome Biol 12: R102
21. Borna T, Salami SA, Shokrpour M (2017) High resolution melting curve analysis revealed SNPs in major cannabinoid genes associated with drug and non-drug types of cannabis. Biotechnol Biotechnol Equip 31: 839-45
22. Toth JA, Stack GM, Cala AR et al (2020) Development and validation of genetic markers for sex and cannabinoid chemotype in Cannabis sativa L. GCB Bioenergy: 213-22
23. Toth JA, Stack GM, Cala AR et al Development and validation of genetic markers for sex and cannabinoid chemotype in Cannabis sativa L. GCB Bioenergy
24. Laverty KU, Stout JM, Sullivan MJ et al (2019) A physical and genetic map of Cannabis sativa identifies extensive rearrangements at the THC/CBD acid synthase loci. Genome Res 29: 146-56
25. Vergara D, Huscher EL, Keepers KG et al (2019) Gene copy number is associated with phytochemistry in Cannabis sativa. AoB Plants 11: plz074
26. McKernan KJ, Helbert Y, Tadigotla V et al (2015) Single molecule sequencing of THCA synthase reveals copy number variation in modern drug-type Cannabis sativa L. BioRxiv: 028654

27. Grassa CJ, Wenger JP, Dabney C et al (2018) A complete Cannabis chromosome assembly and adaptive admixture for elevated cannabidiol (CBD) content. BioRxiv: 458083

28. Richins RD, Rodriguez-Uribe L, Lowe K, Ferral R, O'Connell MA (2018) Accumulation of bioactive metabolites in cultivated medical Cannabis. PLoS One 13.

29. QIAGEN (2018) DNeasy® Plant Handbook. Hilden, Germany

30. Houston R, Birck M, Hughes-Stamm S, Gangitano D (2016) Evaluation of a 13-loci STR multiplex system for Cannabis sativa genetic identification. Int J Legal Med 130: 635-47

31. Roman MG, Gangitano D, Houston R (2019) Characterization of new chloroplast markers to determine biogeographical origin and crop type of Cannabis sativa. Int J Legal Med 133: 1721-32

32. Thermo Fisher Scientific (2010) ABI PRISM® SNaPshot™ Multiplex Kit. South San Francisco, CA

33. Deiana S, Watanabe A, Yamasaki Y et al (2012) Plasma and brain pharmacokinetic profile of cannabidiol (CBD), cannabidivarine (CBDV), Delta(9)-tetrahydrocannabivarin (THCV) and cannabigerol (CBG) in rats and mice following oral and intraperitoneal administration and CBD action on obsessive-compulsive behaviour. Psychopharmacology (Berl) 219: 859-73. doi: 10.1007/s00213-011-2415-0

34. Prandi C, Blangetti M, Namdar D, Koltai H (2018) Structure-Activity Relationship of Cannabis Derived Compounds for the Treatment of Neuronal Activity-Related Diseases. Molecules 23. doi: 10.3390/molecules23071526

35. (2018) WHO Expert Committee on Drug Dependence Critical Review: Isomers of THC.

CHAPTER VI

Evaluation of 19 short tandem repeat markers for individualization of *Papaver somniferum*

This dissertation follows the style and format of *International Journal of Legal Medicine*.

Young, B, Roman, MG, LaRue, B, Gangitano, D, Houston, R (2020) Evaluation of 19 short tandem repeat markers for individualization of *Papaver somniferum*. Sci Justice. https://doi.org/10.1016/j.scijus.2019.12.002

Abstract

Papaver somniferum, commonly known as opium poppy, is the source of natural opiates, which are used as analgesics or as precursors in the creation of semi-synthetic opioids such as heroin. An increase in opioid addiction in the United States has resulted in high rates of illicit opioid use and overdoses. It has recently been shown that *P. somniferum* DNA suitable for genetic analysis can be recovered from heroin samples. The development of a comprehensive genetic individualization tool for opium poppy could serve to link cases and strengthen programs such as the Drug Enforcement Administration's (DEA) Heroin Signature Program, which seeks to combat rising opioid use.

The purpose of this study was to develop a quantitative real-time PCR (qPCR) method for the quantification of opium poppy DNA, compare three commercial DNA extraction kits for their ability to isolate DNA from poppy seeds, and evaluate nineteen opium poppy short tandem repeat (STR) markers for their use in a forensic identification panel. Such a panel could be used for individualizing samples and determining the geographic origin in heroin or poppy seed tea cases. The qPCR method was proven to be reproducible and reliable, specific for P. somniferum, and sensitive enough for forensic case-type samples. Of the three kits tested, the nexttec™ one-step DNA Isolation Kit for Plants was the optimal method and facilitated rapid extraction of DNA from poppy seeds. The majority of evaluated STR primer sets were unreliable or had low discriminatory power, limiting their use for individualization of poppy samples. A six-locus STR multiplex was developed and evaluated according to Scientific Working Group on DNA Analysis Methods (SWGDAM) and International Society of Forensic Genetics (ISFG) guidelines, including the use of a sequenced allelic ladder. The multiplex was found to

have low discriminatory power, with greater than two-thirds of samples analyzed having just two different genotypes. The multiplex was determined to be unsuitable for individualization; however, a genotype map was developed as a proof of concept that these markers may be useful for determining the biogeographical origin of samples. Searching the poppy genome for new STR markers and developing new primer sets may be necessary for the creation of a powerful genetic tool for the individualization of *P. somniferum*.

Keywords: Forensic DNA, Forensic plant science, *Papaver somniferum*, Short tandem repeats, Heroin

Introduction

Papaver somniferum, commonly known as opium poppy, is the source of opium [1]. Opium is a dried latex obtained from the poppy seedpod that contains opiates, such as morphine. Due to their powerful analgesic effects, opiates and semi-synthetic opioids are commonly used as painkillers for moderate to high levels of pain [2]. Additionally, they are highly addictive [3,4] and also serve as precursors in the production of heroin.

The widespread use and addictiveness of opiates have resulted in an epidemic in the United States [5]. People who become addicted to opiates may turn to illegal and dangerous alternatives such as heroin, a schedule I drug that is twice as potent as morphine [6–8]. Other drug users circumvent the law by using legal methods of obtaining a high, such as drinking poppy seed tea [9]. Since poppy seeds are a common food product and contain only negligible amounts of opium latex, they are not scheduled. However, with enough unwashed seeds, it is possible to brew a tea that contains sufficient opiates to create a high [10]. Two published case studies, one fatal, describe the toxic effects of overdosing on poppy seed tea [11,12]. Determining the origin of the poppy seeds used to brew poppy seed tea may be possible through the packaging of the seeds or other documentation. However, in cases where no packaging is available, there is currently no method to determine the origin of the seeds.

One avenue currently being pursued to address the opioid epidemic is to track or source the drug samples, which could aid in identifying distributors, entry points into the country, or the country of origin. This in turn increases law enforcement's ability to locate and confiscate illegal drugs and block the entry of more drugs. The Drug Enforcement Agency (DEA) is attempting to source heroin through the Heroin Signature Program

(HSP). The goal of this program is to chemically analyze heroin samples, and by determining the purity of the sample and the identity of the impurities, the DEA aims to relate samples back to their country of origin and determine the trafficking routes being used [13–15]. The HSP has successfully provided discriminatory results [16]. However, the inclusion of alternative methods of individualization, such as genetic analysis, could strengthen the program.

A recent study by Marciano et al. developed a method for extracting *P. somniferum* DNA from heroin [17]. Tracking heroin through genetic analysis had not been previously utilized due to the high temperature and pH ranges that heroin samples experience during production. It was hypothesized that the DNA would be severely degraded and not viable for forensic analysis. However, Marciano et al. [17] demonstrated that DNA can survive heroin production, thus introducing the possibility of utilizing DNA to individualize heroin samples and determine their origin.

Using DNA for identification has become the gold standard in forensic science, and using DNA to source heroin could be a viable addition or alternative to the DEA's chemical analysis. A precedent for the use of short tandem repeat (STR) multiplexes in forensic plant science has been set by previous research. One fundamental study developed an STR multiplex to individualize marijuana samples, link cases, and determine whether samples shared a common origin [18]. The highly polymorphic nature of STRs results in multiplex panels with high individualization potential. Additionally, the poppy genome is diploid [2], so an STR multiplex developed for the poppy plant could use the methodology of human identification as a model. To date, there is not a multiplex STR method developed for P. somniferum, and little population data has been collected. Without a database and validated

genetic method, DNA isolated from heroin samples can provide little data about heroin trafficking routes.

In the past decade, efforts to use genetic information to identify opium poppy have had varying success. Multiple STR markers for *P. somniferum* have been mined, developed, and evaluated by Lee et al. [19,20], Masárová et al. [21], Celik et al. [22], Şelale et al. [23], and Mičianová et al. [24]. Many of the STRs developed were dinucleotide repeats, which are prone to increased stutter and result in profiles that are difficult to interpret. Instead, the International Society of Forensic Genetics (ISFG) recommends the use of tetranucleotide repeats to reduce stutter [25]. Additionally, most of the markers developed have not been evaluated for their use in forensic investigations and little is known about their variability in different populations. A more recent study by Mičianová et al. [24] built upon previous studies and determined the variability of twelve previously published STRs in Slovak, Czech, and Austrian cultivars. However, not all of the markers were polymorphic, and there was no attempt to create a multiplex. A related study by Kati et al. [26] found eleven polymorphic microsatellite markers in *Papaver rhoeas* (corn poppy), a relative of P. somniferum. These markers were not tested in *P. somniferum*; however, they were hypothesized to be transferable across species within the same family. The purpose of this study was to evaluate the most promising previously published tri- and tetranucleotide STR markers and develop a genetic tool for forensic analysis of P. somniferum samples that followed Scientific Working Group on DNA Analysis Methods (SWGDAM) and ISFG guidelines [25,27]. This included development and validation of a novel quantitative polymerase chain reaction (qPCR) method specific for *P. somniferum* DNA and a comparison of three commercial extraction kits to extract DNA from poppy

seeds. Additionally, 19 primer sets were evaluated for their use in an STR panel. The six most promising markers were incorporated into a preliminary STR multiplex assay, which ultimately proved to lack sufficient discriminatory ability for forensic purposes. The information discovered during the evaluation of the 19 STR loci demonstrates that the markers in previous literature may be unreliable or have low discriminatory power.

Materials and methods

Sample collection

Poppy seed samples (N = 21 packages) were obtained from various online sources (Table 6.1). Origin countries included the United States, Canada, United Kingdom, Spain, Holland, Turkey, Tasmania, and Afghanistan. Three individual seeds were randomly sampled from each package (N = 63).

Table 6.1 Inventory of poppy samples with product information provided by the vendors

Sample ID	Vendor	Origin claimed on packaging	Plant type claimed on packaging	Other information
1	TerraVita	Canada	California Poppy	Poppy Seed Tea Bags
2	Poppy Seed Wash	Not specified	P. somniferum	Unwashed
3	Dual Spices	Holland	Not specified	
4	Medley Hills Farm	Ohio	Blue	
5	International Spice	Spain	Not specified	Unwashed
6	Frontier Co-Op	Not specified	P. somniferum	
7	Sincerely Nuts	U.K.	Not specified	
8	Bob's Red Mill	Not specified	Not specified	
9	Natural Traders Co	Turkey	Blue	
10	Food to Live	England	Not specified	
11	Anna and Sarah	Holland	Blue	Unwashed
12	Nature's Gourmet Classics	Not specified	Not specified	Unwashed
13	Spicy World	Not specified	Black	
14	We Got Nuts	Not specified	Not specified	
15	Tasmanian Connoisseurs	Tasmania	Tasmanian strain	Unwashed
16	We Know Seeds	Afghanistan	Afghan Blue P. somniferum	Unwashed
17	Nodding Turtle	Afghanistan	Afghan Blue P. somniferum	
18	Sincerely Nuts	U.K.	Not specified	Unwashed
19	International Spice	Spain	Not specified	
20	Sincerely Nuts	U.K.	Not specified	Unwashed
21	Unknown	Not specified	Not specified	Purchased in Woodlands, TX

"Not specified" indicates lack of available information on product packaging. It is important to note that the labeling for sample 1 was misleading; though the plant type was claimed as "California poppy," microscopic and chemical analysis confirmed that seeds were P. somniferum, not Eschscholzia californica.

Comparison of DNA extraction methods

Three commercial kits were compared for their utility in extracting DNA from poppy seeds. Single poppy seeds (N = 90; 9 seeds from 10 packages) were homogenized in liquid nitrogen using a Kimble-Chase Kontes™ Pellet Pestle™ (Kimble-Chase, Rockwood, TN). DNA extraction was performed using the DNeasy® Plant Mini Kit (QIAGEN, Hilden, Germany) (N = 30; 3 seeds per package), nexttec™ 1-Step DNA Isolation Kit for Plants (nexttec, Hilgertshausen, Germany) (N = 30; 3 seeds per package), or the PDQeX phytoGEM Plant Kit (MicroGEM, Charlottesville, VA) (N = 30; 3 seeds per package) [28–30]. The manufacturers' protocols were followed for the DNeasy® and nexttec™ kits. For the phytoGEM kit, the homogenized seed material was resuspended in 100 µL of water, then a modified protocol for extraction of DNA from ground plant tissue (unpublished, provided by the manufacturer) was followed using 40 µL of the resuspended material. All three methods yielded 100 µL of DNA extract. The DNA concentration of each extract was then quantified in duplicate. Data sets were assessed for normality using the Shapiro-Wilk W test and assessed for variance using the Levene's test.

DNA quantitation

Real-time PCR parameters

DNA samples were quantified by qPCR on a StepOne™ Real-Time PCR System (Thermo Fisher Scientific, Waltham, MA) using the *P. somniferum*-specific primer set psgSSR006 [19]. Master mix consisted of 12.5 µL of SYBR® Green Real-Time PCR Master Mix (Thermo Fisher Scientific), 0.5 µL of primer mix (20 µM) (Integrated DNA Technologies, Coralville, IA, USA), 0.8 µL of bovine serum albumin (BSA; 8 mg/mL) (Sigma-Aldrich, St. Louis, MO), and 9.2 µL of deionized water. Reactions included 23 µL

of master mix and 2 µL of sample extract. The PCR cycling conditions were as follows: activation for 10 min at 95 °C, followed by 40 cycles of 15 s at 95 °C and 1 min at 60 °C. A primary DNA standard was created by extracting DNA from 40 seeds together, followed by quantitation using the Qubit™ dsDNA HS Assay (Thermo Fisher Scientific) on a Qubit™ Fluorometer 2.0 (Thermo Fisher Scientific). The primary standard was diluted to obtain calibrators at concentrations of 0.01, 0.04, 0.08, 0.16, 0.63, 1.25, 5, and 10 ng/µL. A minimum R2 value of 0.99 was accepted for quantitation.

Real-time PCR parameters

Validation studies for the real-time PCR method included reproducibility and precision, sensitivity, and species specificity. For reproducibility and precision, 15 separate real-time PCR runs were performed with the 8-point calibrator set, two *P. somniferum* DNA extracts (high and low concentration), and a negative template control, all in duplicate. Efficiency was determined using the slope of the standard curve; efficiency = [10(−1/slope)] − 1. Species specificity was determined by performing the real-time PCR assay on non-P. somniferum samples including *Petroselinum crispum* (parsley), *Ocimum basilicum* (basil), *Bos taurus* (cow), *Origanum vulgare* (oregano), *Rosmarinus officinalis* (rosemary), *Canis lupus familiaris* (dog), *Daucus carota* (carrot), *Nicotiana tabacum* (tobacco), *Felis catus* (cat), *Homo sapiens* (human), *Eucalyptus globulus* (eucalyptus), and *Cannabis sativa* (marijuana). During the species specificity studies, melt curve analysis was utilized to screen for the possibility of non-specific amplification products. Sensitivity and limit of linearity studies were performed by evaluating the assay at a wide range of calibrators from 15 to 0.005 ng/µL and choosing the parameters that resulted in an R-squared value above 0.99.

Screening of STR loci

Prior to multiplex development, 19 previously reported primer sets (PMS051, PMS073, psom2, psom9, psom11, psom12, psom13, psom16, psom17, psom22, psgSSR006, psgSSR022, psgSSR044, psgSSR069, psgSSR080, psgSSR220, psgSSR314, psgSSR488, and psgSSR917) [19,22,26] were screened for potential use in a multiplex. Primer sequences for the psgSSR(#) loci were received from the authors of the original report [22].

DNA extraction

Samples (N = 30) extracted using the nexttec™ 1-step DNA Isolation Kit for Plant during the extraction method comparison were used for screening of STR loci. In addition, three seeds from each of the remaining 11 packages (N = 33) were extracted using the nexttec™ 1-step DNA Isolation Kit for a total of 63 samples.

Annealing temperature determination

The optimal annealing temperature (Ta) of STR primer sets was determined using an Eppendorf Mastercycler® Gradient Thermal Cycler (Eppendorf, Hamburg, Germany). Reactions were prepared with 2 µL template DNA (1 ng), 6.25 µL of 2x HotStarTaq® Plus Master Mix (QIAGEN), 1.25 µL 2 µM primer mix (Integrated DNA Technologies), 1.25 µL 5× Q-solution (QIAGEN), 0.4 µL 8 mg/mL BSA (Sigma-Aldrich), and 1.35 µL deionized water. Cycling conditions were as follows: activation for 5 min at 95 °C, followed by 30 cycles of 30 s at 94 °C, 30 s at a gradient of (60 ± 10 °C; 12 wells), 30 s at 72 °C, and a final extension of 30 min at 60 °C. The optimal annealing temperature was determined via electrophoresis on a 2% agarose gel. The highest temperature that produced a bright band was considered the optimal annealing temperature.

PCR amplification and detection

Following annealing temperature determination, STR loci were analyzed for peak morphology and amplification characteristics (efficiency and number of peaks present). DNA amplification was performed using the Type-it® Microsatellite PCR Kit (QIAGEN) on a Veriti™ 96-Well Thermocycler (Thermo Fisher Scientific). The PCR master mix consisted of 6.25 µL of 2× Type-it® Multiplex PCR Master Mix (QIAGEN), 1.25 µL of 2 µM primer mix (Integrated DNA Technologies), and 1.25 µL 5 × Q-Solution (QIAGEN). A volume of template DNA needed to reach the target DNA amount of 1 ng was added, and deionized water was used to bring the total volume to 12.5 µL. The PCR cycling conditions were as follows: 5 min activation at 95 °C, followed by 30 cycles of 30 s at 95 °C, 90 s at the optimal annealing temperature, and 30 s at 72 °C, then a 30 min final extension at 60 °C. Positive and negative controls were included in every PCR assay.

Amplification products were analyzed either using capillary electrophoresis (CE) on a 3500 Genetic Analyzer (Thermo Fisher Scientific) or the DNA 1000 Kit (Agilent, Santa Clara, CA) on an Agilent 2100 Bioanalyzer (Agilent) using the manufacturer's protocol [31]. For CE, PCR product (0.5 µL) was mixed with 9 µL Hi-Di Formamide® and 0.5 µL LIZ® 600 Size Standard (Thermo Fisher Scientific). Samples were denatured for 5 min at 95 °C before being loaded onto the analyzer. The instrument was run using the following conditions: oven: 60 °C; prerun 15 kV, 180 s; injection 1.6 kV, 8 s; run 19.5 kV, 1330 s; capillary length 50 cm; polymer: POP-7™; and dye set G5. Data were visualized using GeneMapper v.5 software (Thermo Fisher Scientific) [32]. An analytical threshold of 100 relative fluorescence units (RFUs) was used.

Primers that produced products with good peak morphology were used to screen a larger number of samples (N = 48 on CE or N = 22 on Bioanalyzer) to assess allele diversity. This screening process was used to ensure that only optimal primer sets were included in the final multiplex.

Multiplex STR typing

Multiplex design

An STR multiplex was designed using the six most reliable and reproducible microsatellite loci: psom12, psom13, psom16, psom17, psgSSR069, and psgSSR080 [19,22] (Table 6.2). The multiplex was designed using Multiplex Manager software v.1.2 [33], and primer interactions were assessed using AutoDimer software [34]. Loci were configured across four dye channels with a minimum of 20 base pairs between loci on the same dye channel. The maximum amplicon size was approximately 250 base pairs. This small maximum length was chosen to reduce the stochastic effects of DNA degradation, which is common in DNA recovered from heroin [17].

Table 6.2 Details of the six-locus multiplex STR panel

Locus	Repeat motif	Dye	Forward primer sequence	Reverse primer sequence	Annealing temperature (°C)	Primer concentration (10X stock; µM)	Reference
psom12	TTC	PET®	CCCTGCTGCTTTCAAATCTC	ATGGAAGAAAAGGTGCCTGA	61	0.75	[19]
psom13	CAG	VIC®	ACCCCCACAACTTTGTCGTA	CGAGTGTTTGGGAACCTGAT	58	0.25	[19]
psom16	GAA	6-FAM™	AAAATGGGAACAGCCATCAG	TCTGGGCCTTCACCACTTAC	58	1	[19]
psom17	AAT	6-FAM™	CCCAACAATTTGGTGCAGTA	CCCGAGGTGAAACTTCTCTG	53	1	[19]
psgSSR069	CAAT	NED™	ATAGATTTATTTTGGCCACCT	CACCTATTGATTGAGGATGAA	53	0.75	[22]
psgSSR080	GGAA	VIC®	ACAGAGACAGTTCACTTCCAA	ATGAGTCGTTTTTGTTGTTGT	64	0.5	[22]

Multiplex PCR amplification parameters

DNA amplification was performed using the Type-it® Microsatellite PCR Kit [35]. The Veriti™ 96-Well Thermocycler (Thermo Fisher Scientific) was used for amplification. The PCR master mix consisted of 6.25 μL of 2× Type-it® Multiplex PCR Master Mix (QIAGEN), 1.25 μL of 10× primer mix (Integrated DNA Technologies) (Table 6.2), and 1.25 μL 5× Q-Solution (QIAGEN). A volume of template DNA needed to reach the target DNA amount of 1 ng was added, and deionized water was used to bring the total volume to 12.5 μL. A touchdown amplification protocol was used to compensate for differences in optimal annealing temperature (Table 6.2) and to reduce non-specific amplification. The cycling conditions were as follows: 5 min activation at 95 °C; followed by 2 cycles of 30 s at 95 °C, 90 s at 63 °C, and 30 s at 72 °C; 2 cycles of 30 s at 95 °C, 90 s at 58 °C, and 30 s at 72 °C; 27 cycles of 30 s at 95 °C, 90 s at 53 °C, and 30 s at 72 °C; and a 30 min final extension at 60 °C.

Capillary electrophoresis and genotyping

PCR products were separated and detected on a 3500 Genetic Analyzer (Thermo Fisher Scientific) as previously described in Section 2.4.3. Data were visualized using GeneMapper v.5 software (Thermo Fisher Scientific) using a custom panel and bin set [32]. An analytical threshold of 100 RFUs was used.

Allelic ladder design and allele sequencing

An allelic ladder was produced from the alleles observed during microsatellite screening. First, the alleles were sequenced using the BigDye™ Direct Cycle Sequencing Kit (Thermo Fisher Scientific) as per the manufacturer's protocol [36]. Centri-Sep™ Spin columns (Princeton Separations, Freehold, NJ) were used to purify the sequencing

products. Sequencing data was analyzed using Geneious Pro v7.1 (Biomatters, Auckland, New Zealand).

The ladder was created according to previous reports [18,37]. Each sample was amplified in single-plex PCR, and the products representing different alleles for the same locus were combined and balanced using CE. The balanced ladders for each locus were then combined and analyzed for inter-locus balance, creating the final allelic ladder. This ladder was diluted 1:1000 in TE buffer to create a stock which was re-amplified utilizing a touchdown PCR method, as follows: 5 min activation at 95 °C; followed by 2 cycles of 30 s at 95 °C, 90 s at 63 °C, and 30 s at 72 °C; 2 cycles of 30 s at 95 °C, 90 s at 58 °C, and 30 s at 72 °C; 17 cycles of 30 s at 95 °C, 90 s at 53 °C, and 30 s at 72 °C; and a 30 min final extension at 60 °C.

STR validation studies

Studies were performed to analyze sensitivity, specificity, reproducibility, and precision of the multiplex STR assay. The sensitivity of the assay was determined by analyzing DNA at seven different DNA input amounts: 1000, 500, 250, 125, 62.5, 31.2, and 15.6 pg. Three samples were analyzed at each of the input amounts in triplicate, for a total of 63 reactions. The species specificity was analyzed by amplifying DNA from *Petroselinum crispum* (parsley), *Ocimum basilicum* (basil), *Bos taurus* (cow), *Origanum vulgare* (oregano), *Rosmarinus officinalis* (rosemary), *Canis lupus familiaris* (dog), *Daucus carota* (carrot), *Nicotiana tabacum* (tobacco), *Felis catus* (cat), *Homo sapiens* (human), *Eucalyptus globulus* (eucalyptus), *Cannabis sativa* (marijuana), *Escherichia coli*, and *Shewanella xiamenensis* in duplicate. Additionally, poppy samples screened with a 1

ng DNA target were used to assess the average peak height and average peak height ratio of profiles.

Population studies

The validated multiplex was used to screen the entire sample population (N = 63). The discriminatory power of each locus and of the six-locus multiplex were then assessed by analyzing the genotypic diversity of the population. The ability of the multiplex to trace samples to a country of origin was evaluated through the creation of a genotype map.

Statistical analysis

To determine the discriminatory power of the multiplex, population studies were performed to analyze allele and genotype frequency data. Additionally, quality of the multi-locus profiles was analyzed using profile completeness, average peak height, and average peak height ratio.

Results

Comparison of DNA extraction methods

The data sets generated in the extraction kit comparison proved to violate assumptions necessary for statistical analysis. It was found not to be normally distributed by the Shapiro-Wilk W test ($p < 0.05$, $N = 60$) and to have unequal variance by the Levene's test ($p < 0.05$, $N = 60$). Additionally, the data sets were not symmetrical and did not have similar kurtosis or skewness. Due to these issues, it was determined that statistical testing was inappropriate. Instead, the data were visually examined for trends. The nexttec™ 1-step DNA Isolation Kit for Plant (nexttec) yielded the highest DNA concentrations (average: 1.05 ng/µL) of the three kits tested, followed by the PDQeX phytoGEM Plant

Kit (MicroGEM) (average: 0.17 ng/µL), and finally the DNeasy® Plant Mini kit (QIAGEN) (average: 0.07 ng/µL) (Fig. 6.1).

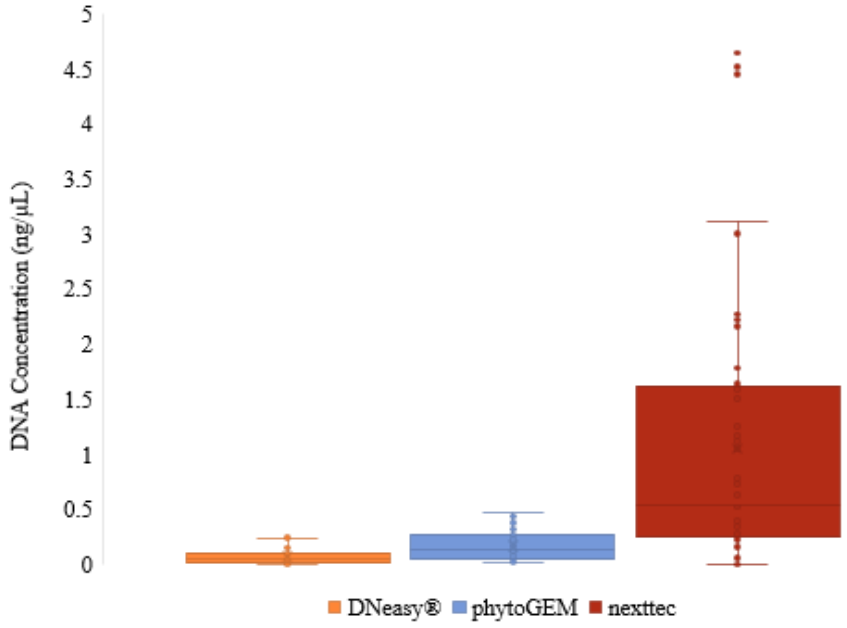


Fig. 6.1 Comparison of DNA concentrations from three extraction methods: DNeasy® Plant Mini kit (QIAGEN), nexttec™ 1-step DNA Isolation Kit for Plant (nexttec), and PDQeX phytoGEM Plant Kit (MicroGEM) (N=30 seeds each, quantified in duplicate)

Real-time PCR development and validation

Of the twelve non-*P. somniferum* species tested, only *C. lupus familiaris* amplified. However, the canine sample quantified at just 0.0015 ng/µL (Ct 28.95), an order of magnitude lower than the lowest standard (0.01 ng/µL), so this amount was considered negligible. Melt curve analysis showed a single peak for poppy samples, indicating that there was no non-specific amplification. Therefore, the qPCR assay was determined to be specific for *P. somniferum*.

The sensitivity study found that the optimal range of calibrator concentrations was 0.01 ng/µL to 10 ng/µL, using an eight-point standard curve (0.01, 0.04, 0.08, 0.16, 0.63,

1.25, 5, and 10 ng/µL). When the assay was tested with higher and lower concentrations, the standard curve lost its linearity ($R2 < 0.99$).

The reproducibility and precision data generated for the eight quantification standards and linear regression of the standard curve are shown in Tables 6.3 and 6.4, respectively. Precision is represented by the percent coefficient of variation of Ct values (%CV = 100 × (standard deviation/mean)), which had an average of 0.64%. Reproducibility is represented by the standard deviation and range of the Ct values, which were consistently under 0.2 and 1, respectively. Amplification efficiency averaged approximately 93%, and R2 averaged 0.999. The coefficient of variation for all linear regression statistics was 2% or lower.

Screening of STR loci

The results of the STR loci screening, including optimal annealing temperatures for each primer set and the number of alleles observed, are shown in Table 6.5. Two primer sets (psom22 and psgSSR044) failed to produce a PCR product. Four markers (psom9, psom11, PMS051, and psgSSR917) had suboptimal amplification, indicated by abnormal peak shape or production of multiple peaks (Fig. 6.2). Additionally, seven markers (psom2, PMS073, psgSSR006, psgSSR022, psgSSR220, psgSSR314, and psgSSR488) lacked genetic diversity, with only a single allele observed after sample screening. The remaining six loci (psom12, psom13, psom16, psom17, psgSSR069, and psgSSR080) were incorporated into a multiplex STR panel.

Table 6.3 Reproducibility data for the eight quantification standards in the real-time PCR assay

Standard	Poppy DNA (ng/μL)	Average Ct	Standard deviation	Minimum	Maximum	Range
1	10.00	15.94	0.14	15.48	16.17	0.69
2	5.00	16.80	0.14	16.33	17.03	0.69
3	1.25	18.82	0.14	18.42	19.13	0.71
4	0.63	19.98	0.14	19.58	20.38	0.80
5	0.16	22.12	0.17	21.75	22.52	0.77
6	0.08	23.17	0.17	22.86	23.61	0.76
7	0.04	24.26	0.16	23.99	24.69	0.71
8	0.01	26.30	0.18	25.82	26.75	0.93

Table 6.4 Linear regression data of the standard curve showing reproducibility and precision of the real-time PCR assay

Run	Slope	Amplification efficiency (%)	R^2	Y-intercept
1	-3.69	86.64	0.999	19.45
2	-3.51	92.74	0.999	19.28
3	-3.47	94.17	0.999	19.35
4	-3.52	92.28	0.999	19.39
5	-3.54	91.57	0.999	19.34
6	-3.55	91.29	0.999	18.93
7	-3.50	93.25	0.999	19.31
8	-3.49	93.40	0.999	19.42
9	-3.47	94.21	0.999	19.39
10	-3.48	93.91	1	19.27
11	-3.43	95.53	0.999	19.21
12	-3.48	93.95	0.999	19.25
13	-3.49	93.58	0.999	19.37
14	-3.46	94.47	0.999	19.39
15	-3.48	93.65	0.999	19.23
Average	-3.50	92.98	0.999	19.30
Standard Deviation	0.06	2.07	0.00026	0.13
Coefficient of Variation	0.02	0.02	0.00026	0.01
Minimum	-3.69	86.64	0.999	18.93
Maximum	-3.43	95.53	1.000	19.45
Range	0.26	8.89	0.001	0.52

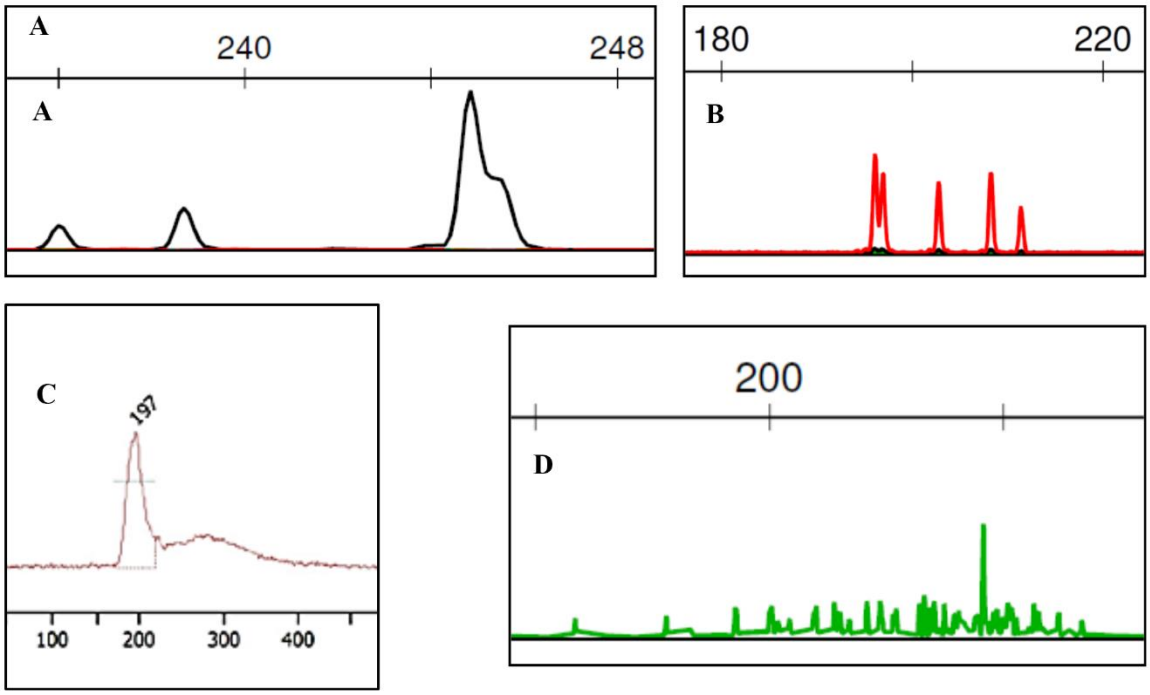


Fig. 6.2 Electropherograms of markers with suboptimal amplification. Markers psom9 (A) and PMS051 (B) consistently produced more than two peaks, which is indicative of a genetic duplication event. Marker psom11 (C), screened on the Agilent 2100 Bioanalyzer, showed suboptimal peak morphology, and marker psgSSR917 (D) had nonspecific amplification. None of these loci are optimal for use in an STR multiplex.

Table 6.5 Results from screening of 19 STR loci

Locus	Reference	Repeat motif	Optimal Ta (°C)	Detection method	Number of alleles	Number of samples screened
psom2	[19]	CTT	55	Bioanalyzer	1	22
psom9	[19]	AGA	64	CE	-	7
psom11	[19]	TCT	64	Bioanalyzer	1	11
*psom12	[19]	TTC	61	Bioanalyzer	2	63
*psom13	[19]	CAG	58	CE	2	63
*psom16	[19]	GAA	58	CE	3	63
*psom17	[19]	AAT	53	CE	3	63
[2]psom22	[19]	TAA	-	-	-	-
PMS051	[26]	GAA	56	CE	-	3
PMS073	[26]	TTC	54	CE	1	48
[2]psgSSR044	[22, 24]	TGA	-	-	-	-
*[1]psgSSR069	[22, 24]	TAAA	53	Bioanalyzer	1	63
psgSSR006	[22]	AACA	60	CE	1	48
psgSSR022	[22]	TGG	63	Bioanalyzer	1	22
*[1]psgSSR080	[22]	GGAA	64	Bioanalyzer	1	63
psgSSR220	[22]	TCA	61	Bioanalyzer	1	22
psgSSR314	[22]	TGAT	61	Bioanalyzer	1	22
psgSSR488	[22]	TTAT	58	CE	1	48
psgSSR917	[22]	GCGG	64	CE	-	7

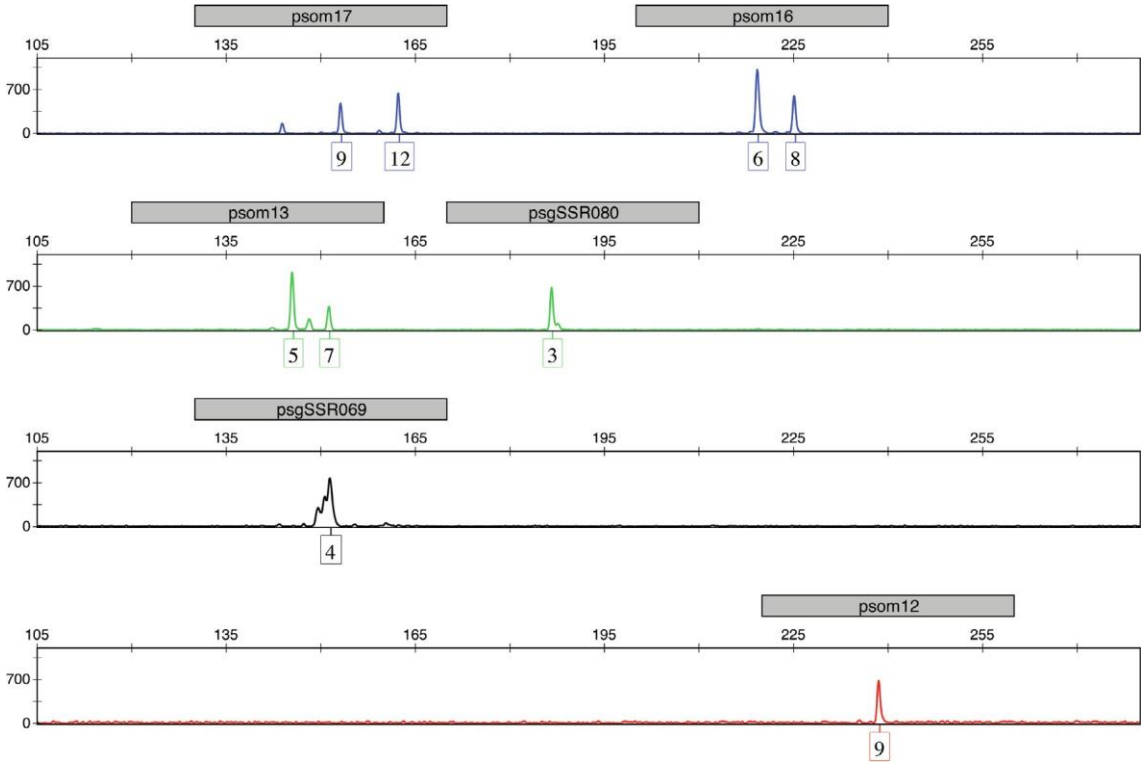


Fig. 6.3 Allelic ladder for the six-locus STR multiplex

Multiplex STR typing

Allelic ladder design and allele sequencing

The multi-locus allelic ladder contained a total of nine alleles (Fig. 6.3). The locus psom13 had two alleles, which were found to contain five and seven repeats of the motif. For psom16, two alleles (six and eight repeats) were found in the initial screening and incorporated into the allelic ladder, and a third allele (seven repeats) was found during the multiplex screen. For psom17, two alleles found in the initial screen (nine and twelve repeats) were incorporated into the ladder, and a third allele (ten repeats) was observed in the multiplex screen. For psom12 and psgSSR080, the samples we believed to contain different alleles actually represented only one allele at each locus; psom12 had a nine repeat allele while psgSSR080 had a three repeat allele. During multiplex screening, an eight

allele was also observed for psom12. For psgSSR069, the sequencing quality was low; however, all three samples sequenced appeared to have four repeats of the STR motif. The international guidelines for naming STR alleles were followed to ensure accurate nomenclature [38–40]. The allele sequences are shown in Fig. 6.4-6.8. Novel sequences were submitted to GenBank (MN271694 – MN271699).

This is a note for Fig. 2.1 through 2.7. In the consensus sequences, the forward and reverse primer binding locations are underlined and the single base extension (SBE) primer binding sites are highlighted.** The location of hSTR repeats are indicated in the consensus sequences as **[REPEAT]**, indels are indicated as **[INDEL]**, and SNPs are indicated by their **nucleotide ambiguity code.** The GenBank accession numbers are also referenced in the table. N refers to the total number of samples with the indicated haplotype.

CCCTGCTGCTTTCAAATCTCTTCGTCTTGTCTTAGTCACTCTTCATTCC
ATGATGCGATCTTCTCTGAGCGG**[REPEAT]**A**R**AAACACCAGACGGAT
ACGTAGTCTTTTTCTTCGCCTTCCGCATGGGTGTATGAGAATGAATAC
CTTTATGCGATTCGTTAAGGGTAACTTGAACCGACTGCCTATCTTTTTG
TAGCTCAGGCACCTTTTCTTCCAT

Allele	[REPEAT]	Post-repeat SNP	N	Accession	Reference
	TTC	R			
8	8	A	1	GU903172, FG610824	[19], Unpublished[1]
8	8	G	0	GU903173	[19]
9	9	A	2	MN271694	This study

Fig. 6.4 Consensus sequence of the psom12 locus, allele nomenclature, and genotypes observed

[1]GenBank sequence submitted by Hagel,J., Zulak,K.G. and Facchini,P.J. Expressed sequence tags from stem and root cDNA libraries of opium poppy (Papaver somniferum).

ACCCCCACAACTTTGTCGTAAACACTGTTGAAGAACGACTGTCTTC**R**GACGTC
YG**Y**T**Y**TTCGTGTTCTT**[REPEAT]**AAACA**R**AGTTT**S**ATCGAACTCTGATTTCTTC
TTCTC**Y**GGCCCTCATCAGGTTCCCAAACACTCG

Allele	Pre-repeat SNPs				[REPEAT]	Post-repeat SNPs			N	Accession	Reference
	R	Y	Y	Y	CAG	R	S	Y			
5	G	T	T	C	5	G	G	T	2	MN271695	This study
6	G	C	C	C	6	G	G	T	0	GU903174	[19]
6	A	C	T	T	6	A	C	C	0	GU903175	[19]
7	G	T	T	C	7	G	G	T	2	FG606729	Unpublish-ed[1]

Fig. 6.5 Consensus sequence of the psom13 locus, allele nomenclature, and genotypes observed

[1]GenBank sequence submitted by Hagel,J., Zulak,K.G. and Facchini,P.J. Expressed sequence tags from stem and root cDNA libraries of opium poppy (Papaver somniferum).

AAAATGGGAACAGCCATCAGGGA**Y**AGAATTCAAGAACGGTTGACGGGACAG
CAGCAATTGCGACTTAAGAATGGTGATCCTGAGAATGGAGAAGAAGAATA**Y**T
ATGACGAAGACGAAGAATACTACTA**Y**TACGAGGATGATGGCGATGATGATG
ATTTT**[REPEAT]**GAGGAGG**M**GAAAGGGAAAGGG**R**AAAAGTAAGTGGTGAAG
GCCCAGA

Allele	Pre-repeat SNPs			[REPEAT]		Post-repeat SNPs		N	Accession	Reference
	Y	Y	Y	GAA	GAG	M	R			
6	C	T	T	5	1	C	A	0	GU903180	[19]
6	T	C	C	6		A	A	2	MN271696	This study
8	T	C	C	8		A	G	2	FE967887	Zulak et al. 2007[1]

Fig. 6.6 Consensus sequence of the psom16 locus, allele nomenclature, and genotypes observed

[1]K.G. Zulak, A. Cornish, T.E. Daskalchuk, M.K. Deyholos, D.B. Goodenowe, P.M. Gordon, D. Klassen, L.E. Pelcher, C.W. Sensen, P.J. Facchini, Gene transcript and metabolite profiling of elicitor-induced opium poppy cell cultures reveals the coordinate regulation of primary and secondary metabolism, Planta, 225 (2007) 1085-1106.

CCCAACAATTTGGTGCAGTAAGAAACAATCACTGACTACTC**RW**AGAAAAGA
CATTTTCCCTTCAAGAGAAAAACAAGAACAC**[REPEAT]**GCTACAAACTCCTA
AAACCCACTACCAACTCAGAGAAGTTTCACCTCGGG

Allele	Pre-repeat SNPs		[REPEAT]		N	Accession	Reference
	R	W	AAT	GAT			
6	A	A	6		0	GU903182	[19]
9	G	A	9		2	MN271697	This study
10	G	A	10		1	FE967526	Zulak et al. 2007[1]
10	G	A	9	1	0	GU903183	[19]
10	G	T	9	1	0	GU903181	[19]
12	G	A	12		2	MN271698	This study

Fig. 6.7 Consensus sequence of the psom17 locus, allele nomenclature, and genotypes observed

[1]K.G. Zulak, A. Cornish, T.E. Daskalchuk, M.K. Deyholos, D.B. Goodenowe, P.M. Gordon, D. Klassen, L.E. Pelcher, C.W. Sensen, P.J. Facchini, Gene transcript and metabolite profiling of elicitor-induced opium poppy cell cultures reveals the coordinate regulation of primary and secondary metabolism, Planta, 225 (2007) 1085-1106.

ACAGAGACAGTTCACTTCCAAGAGACCGCAAATGAC**[REPEAT]**CAAATATTA
CATTTCAAAGCTGTGTCGATTAGCCTTCCAGCAACACATCCTAATGGTATTGA
CCACGGGTAGTCATTGATTTTTATTTTACTAACTACAGCGAACTACGTAATCA
ATTAAGACAACAACAAAAACGACTCAT

Allele	[REPEAT]	N	Accession	Reference
	GGAA			
3	3	3	MN271699	This study

Fig. 6.8 Consensus sequence of the psgSSR080 locus, allele nomenclature, and genotypes observed

STR validation studies

The sensitivity of the multiplex was found to be 0.5 ng (Fig. 6.9). Below this input amount, a full profile was not reliably obtained. Below 125 pg of DNA, severe allele dropout and peak height imbalance were observed. For optimal results, a DNA target of 1 ng was used.

When evaluating species specificity, it was found that there was cross-reactivity with *B. taurus* (cow) and *H. sapiens* (human) DNA (Fig. 6.10).

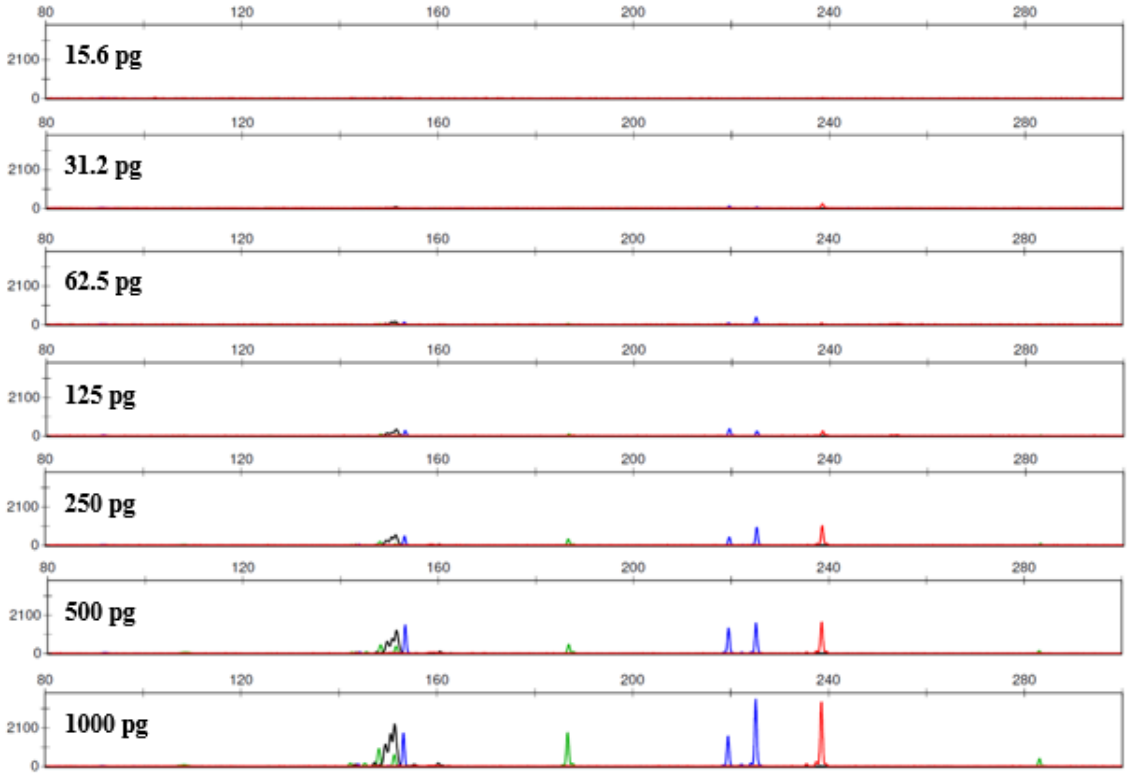


Fig. 6.9 Electropherograms from the STR multiplex sensitivity study using seven DNA targets

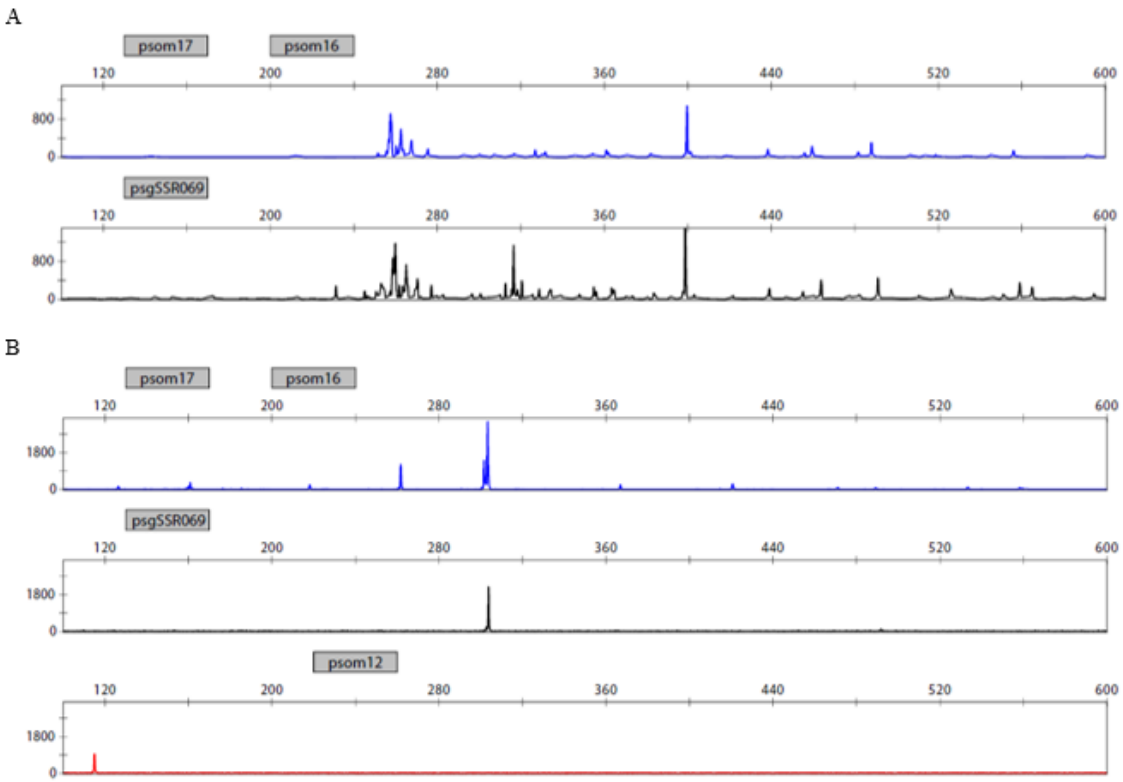


Fig. 6.10 Cross-reactions observed during the species specificity study. (A) Cross-reactions were seen in *B. taurus* (cow) DNA with the primers for psgSSR69 and psom17 and/or psom16. (B) Cross- reactions were seen in *H. sapiens* (human) DNA with the primers for psgSSR69, psom12, and psom17 and/or psom16

Population studies

Full profiles were obtained for 58 out of 63 samples. The majority of samples (N = 42 out of 58) expressed one of two genotypes, termed genotype A and genotype B (Table 6.6). These two genotypes only differed by a single allele at the psom12 locus, which was either a nine repeat or a suspected null allele. Additionally, the remaining genotypes only differed from genotypes A and B by up to three alleles. Only five individuals had unique genotypes not observed in any other sample.

Additionally, the genotypes observed at each locus were analyzed (Fig. 6.11). The marker psgSSR080 was found to be mono-allelic, with every sample (N = 62; dropout

occurred in one sample) having a 3 allele. For psgSSR069, 60 samples had a 4 allele. Three samples (all from a common source) also showed additional peaks at this locus (no sequencing data) (Fig. 6.12B). Regarding single-locus genotypes for psom13, psom16, and psom17, the majority of samples (> 80%) were represented by a single genotype. For psom12, the majority of samples (> 90%) were represented by either the 9-repeat allele or the suspected null allele, with each representing a near equal percentage of the population. Using the reported origin of seed samples, a genotype map was developed (Fig. 6.13).

Overall profile quality for all samples was also assessed. For samples utilizing a 1 ng target, the height of homozygous peaks ranged from 105 to 13,564 RFU, with an average peak height of 2,095 RFU. The height of heterozygous alleles ranged from 258 to 12,654 RFU, with an average of 2,810 RFU. The peak height ratio range for heterozygote loci was 35% to 84% with an average peak height ratio of 53%.

Table 6.6 Allele breakdown of genotypes (N=63) observed

Genotype	psom17	psom16	psom13	psgSSRg080	psgSSR069	psom12	N=
A	9	6,8	7	3	4	9	26
B	9	6,8	7	3	4	Null	16
C	12	6,8	7	3	4	Null	2
D	9	6,7	7	3	4	9	1
E	12	6,8	5	3	4	9	1
F	9	6,8	5	3	4	Null	1
G	12	6	7	3	4	Null	2
H	12	6,8	5	3	4	Null	1
I	9	6,8	7	3	4	8	2
J	10	8	7	3	4	8	1
K	9	8	7	3	4	9	2
L	10	6	5	3	*	Null	3

* Indicates a different but unknown genotype at the psSSR69 locus due to the occurrence of multiple peaks, as seen in Fig. 6.12

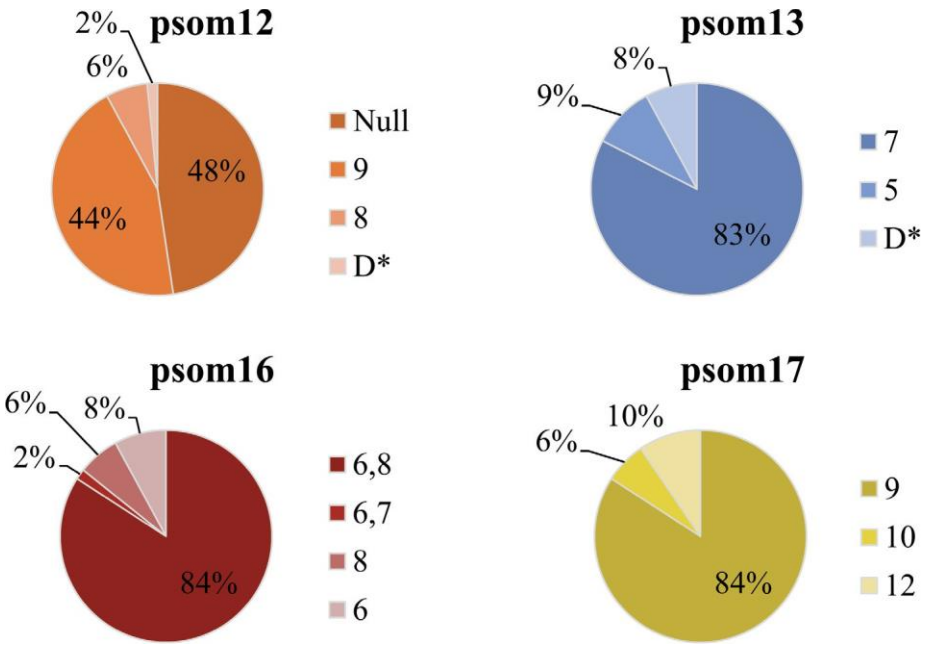


Fig. 6.11 Genotype breakdown of the four loci that showed genetic diversity. D* indicates locus dropout

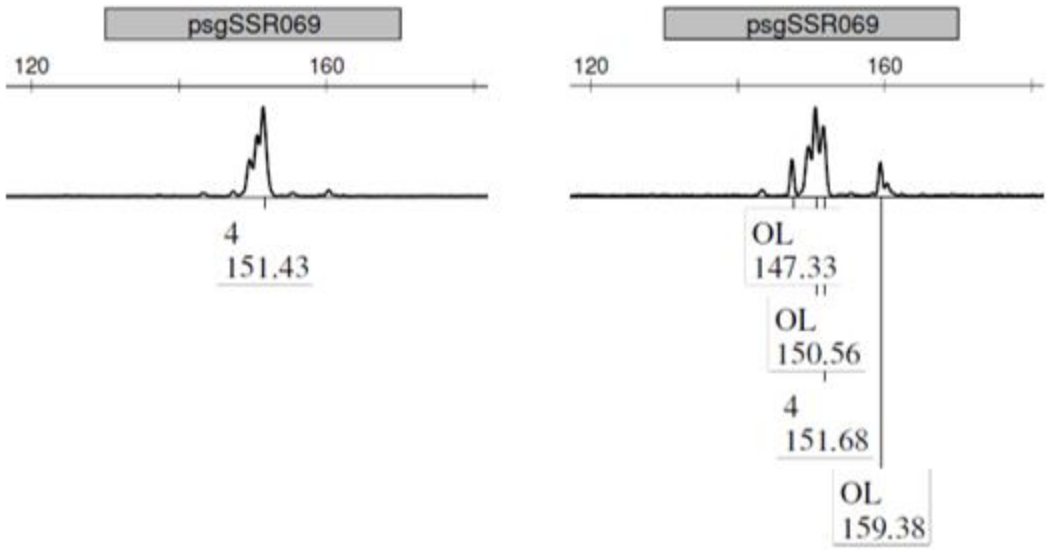


Fig. 6.12 Electropherograms for psgSSR069 showing (A) the 4 allele seen in most samples (N=60) and (B) the alternate genotype (labeled "*" due to lack of sequencing data) observed in three samples

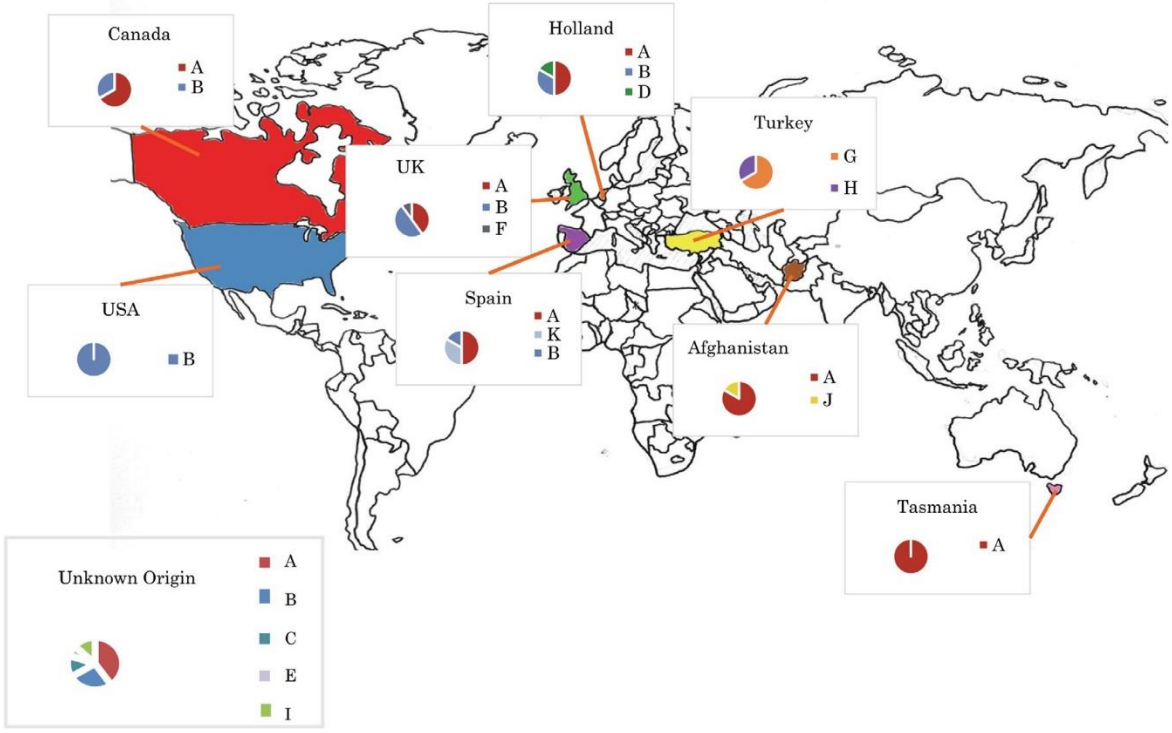


Fig. 6.13 Genotype map displaying the prevalence of observed genotypes in eight countries of origin

Discussion

Comparison of DNA extraction methods

Though statistical testing was determined to be inappropriate, it was decided through visual examination of the data that the nexttec™ extraction kit was the most efficient method. One factor that likely contributed to the issues with normality and unequal variance in these data sets is the inconsistency of seed homogenization. The method of homogenization was not completely consistent, due to some seeds being tougher than others and difficulty in visualizing how well the seed was homogenized. This led to some seeds being homogenized better than others, increasing the variance of DNA yield.

In addition to its higher average DNA yields, the nexttec™ extraction kit was also faster than the DNeasy® kit, taking just 45 min to extract DNA from 12 samples (after homogenization) rather than 90 min. The phytoGEM extraction method was even faster than the nexttec™ kit, at just 20 min after homogenization, but it was decided that the greater potential DNA yield from the nexttec™ kit was more desirable. The multiplex was sensitive to approximately 0.5 ng, and the maximum input volume of DNA extract was 3.75 µL; therefore, dropout would be expected with DNA concentrations under 0.13 ng/µL. Since extracts from the DNeasy® kit averaged only 0.07 ng/µL, allelic dropout would be expected in many of the samples. Because many aspects of these three kits are proprietary, it is difficult to determine exactly why the nexttec™ kit outperformed the other two. The nexttec™ kit does use a different purification system than traditional silica-based extraction methods such as the DNeasy® kit. Silica-based extraction kits use a silica sorbent that binds DNA, followed by several wash steps to get rid of unwanted proteins, cell debris, and other contaminants, and the final step involves elution of the DNA into a clean tube.

DNA can come unbound from the silica and be lost during the wash steps, or it may remain bound after elution, resulting in loss of DNA [41–43]. The nexttec™ kit utilizes a sorbent that retains components other than DNA, and the DNA is eluted in a single step, which reduces DNA loss.

Real-time PCR validation studies

The high precision, R-squared values, efficiency, and reproducibility indicate that the qPCR method developed in this study is a robust and suitable method for quantifying opium poppy DNA. This is the first reported quantitation method specific for *P. somniferum* DNA, and it lays an important groundwork for future studies evaluating new STR markers in opium poppy.

Screening of STR loci

A low amount of genetic diversity was observed in all of the markers examined. The psom markers from Lee et al. [19] were derived from expressed sequence tag (EST) databases, so it is expected that they would have lower diversity than markers derived from genomic sequences, such as those discovered by Celik et al. [22]. Lee et al. [19] tested six loci (psom2, psom4, psom12, psom13, psom17, and psom22) for polymorphism in Korean P. somniferum samples, and found between 2 and 5 alleles per locus. Five of these loci were included in this study (psom4 was excluded due to its dinucleotide repeat motif). We eliminated psom2 and psom22 due to poor amplification, likely due to the large difference in melting temperatures between the forward and reverse primers. For psom12 and psom13, Lee et al. and this study both showed two alleles and low heterozygosity. Additionally, Lee et al. observed null alleles for psom12 in about 40% of samples tested, which is comparable to the frequency of null alleles we observed at this locus (about 48%).

The most polymorphic locus in Lee et al.'s report was psom17, which had five alleles, and we observed three alleles in our samples. In addition, our study included three other loci (psom9, psom11, and psom16) which were proposed by Lee et al. but were only tested on a small sample set in their original report due to multiple peaks (psom11) or lack of variability in the first 32 samples screened (psom9 and psom16). We found that indeed psom11 had suboptimal peak morphology (Fig. 6.2C), which could be due to the amplification of multiple regions in *P. somniferum*, as observed by Lee et al. However, we also observed multiple peaks for psom9 (Fig. 6.2A), which was not seen in the original report. We also found three alleles for psom16, where Lee et al. found only one.

Nine of the loci (psgSSR006, psgSSR022, psgSSR044, psgSSR069, psgSSR080, psgSSR220, psgSSR314, psgSSR488, and psgSSR917) tested in this study came from a report by Celik et al. [22]. These markers were derived from genomic sequences rather than EST data, so we would expect them to be more polymorphic [44,45]. Unfortunately, six of the nine markers were monoallelic (psgSSR006, psgSSR022, psgSSR080, psgSSR220, psgSSR314, and psgSSR488), two had suboptimal or non-specific amplification (psgSSR917, Fig. 6.2D, and psgSSR069, Fig. 6.12), and one failed to amplify (psgSSR044). Celik et al. tested these markers in Turkish samples and observed polymorphism in psgSSR006 and psgSSR022, did not observe polymorphism in psgSSR069, indicated poor amplification in psgSSR044, and did not test the other three loci for variability. The reason psgSSR044 and psgSSR069 were included in the study despite amplifying poorly or being monomorphic in Celik et al.'s report is due to variability seen in these loci when tested in another study [24]. Mičianová et al. [24] observed variability in psgSSR044 (two alleles) and psgSSR069 (three alleles) in Slovak, Czech,

and Austrian cultivars. To note, in their paper, Mičianová reported in a table that they were using the EST-SSRs developed by Şelale et al. [23] (psSSR44 and psSSR69); however, the primer sequences they reported corresponded to the genomic SSRs reported by Celik et al. [22] (psgSSR044 and psgSSR069). This discrepancy could explain the differences observed for these two loci between the Celik et al. [12] and Mičianová et al. [24] studies.

Two P. rhoeas markers (PMS051 and PMS073) discovered by Kati et al. [26] were included in the screening. *P. rhoeas* is a member of the Papaveraceae family and therefore a relative of *P. somniferum*. It was predicted that the markers would be transferable between the two species, which was confirmed by searching the Basic Local Alignment Search Tool (BLAST) (NCBI). Kati et al. observed variability in both loci, with eight alleles for PMS051 and four alleles for PMS073. Though both primer sets amplified in our samples, PMS051 produced multiple peaks (Fig. 6.2B), and PMS073 proved monomorphic. Therefore, neither marker was useful for individualization of *P. somniferum.*

Suboptimal amplification, indicated by abnormal peak shape or the production of multiple (> 2) peaks, was observed in four markers (psom9, psom11, PMS051, and psgSSR917) (Fig. 6.2). The markers psom9 and PMS051 each consistently produced greater than two peaks. This is indicative of a genetic duplication event that is common in plants [46]. BLAST searches of both amplicons revealed partial duplication across multiple chromosomes. This makes these primers unsuitable for use in an STR multiplex. Additionally, psom11 showed suboptimal peak morphology, with an amplification artifact appearing after the main peak. In psgSSR917, large amounts of non-specific amplification made this primer set unusable.

A misidentification of diversity occurred in psom12, psgSSR069, and psgSSR080 when the loci were screened on the Bioanalyzer due to lower accuracy in sizing amplicons compared to the 3500 Genetic Analyzer. In addition, initial screening of psgSSR069 on the Bioanalyzer did not produce abnormal peaks; however, following capillary electrophoresis, a three-peak complex was observed, which can be seen in the allelic ladder (Fig. 6.3).

Multiplex STR typing

Allele sequencing

The sequencing data for psgSSR069 was of poor quality, but it was possible to determine that four repeats of the motif CAAT were present in all samples sequenced. Celik et al. [22] also reported a 4 allele at this locus. An abnormal peak morphology was observed for this marker, indicating suboptimal amplification, which likely affected sequencing.

STR validation studies

The cross-reactions seen when human and cow DNA were amplified are not ideal for forensic analysis due to the increased chance of false positives and difficulty in interpreting profiles containing a mixture of poppy DNA with human or cow DNA. This is especially worrisome with the cross-reaction seen with human samples, where the peaks fall within psom16 and psom17 marker ranges. Since some of the dye channels only contain one locus, we can determine which forward fluorescent primer was a source of the cross reactions. The psgSSR069 forward primer cross-reacts with both cow and human DNA, while the psom12 forward primer cross-reacts only with human. Additionally, psom17 and/or psom16 also cross-react with cow and human.

While performing validation studies for the multiplex, several persistent optimization issues were also seen. In psom13, there was non-specific amplification, producing peaks of similar height to true alleles. This is likely due to the use of a touchdown PCR method; although this method does reduce the amount of non-specific amplification, the majority of the PCR cycles are still below psom13's optimal annealing temperature of 58 °C. Our results suggest that the two cycles of amplification at 58 °C are insufficient to properly suppress the non-specific amplification products. No non-specific amplification was observed when this locus was amplified in single-plex at its optimal annealing temperature. Additionally, psom16 shows minor incomplete adenylation, characterized by a peak one base pair shorter than the allele peak, despite the PCR method utilizing a 30-minute final adenylation at 60 °C; however, this did not affect interpretation of the profiles. Suboptimal amplification, resulting in a three-peak complex, was observed for psgSSR069 (Fig. 6.12A). In three samples, all originating from the same source, two additional peaks were observed at this locus (Fig. 6.12B), which may indicate nonspecific amplification or duplication of this locus. The genotype for these samples has tentatively been listed as "*" (Table 6.6) to represent a different genotype than what was seen in the other 60 samples. Lastly, it was found that psom12 may have a null allele, which may be problematic when dealing with degraded or low copy number samples because it can be difficult to tell the difference between a null allele and dropout. Null alleles occur due to mutations in the primer binding site, which can prevent the primer from annealing and thus prevent amplification. This issue could potentially be avoided by identifying the binding site mutation and designing a degenerate primer for this STR locus.

A wide range of peak heights were observed during analysis of the profiles, despite the input DNA amount being consistent across all samples. This observation is worrisome because variable amplification efficiency among samples could lead to instances of dropout, even in samples with sufficient quantities of DNA. Additionally, peak height ratios were commonly under 70%, a common cutoff for heterozygote determination, potentially leading to single source samples being mistaken for mixtures or complicating mixture interpretation.

Population studies

The low genetic diversity observed in both the multi-locus and single-locus genotypes gives this multiplex little discriminatory power. Low genetic diversity in *P. somniferum* samples could be due to inbreeding of individuals in order to select for desirable genetic characteristics, such as opioid content. The Founder Effect is another possible contributing factor to low diversity and results when populations are originated by a limited number of ancestors. Though the multiplex would not be able to reliably differentiate between individuals chosen at random, it could be used for exclusionary purposes. In addition, we found that the multiplex showed limited discrimination of samples by their country of origin. Genotype B, the second most common genotype, was only seen in western countries. This data was generated from a smaller than ideal sample size, so it should only be considered a proof of concept. Analyzing more diverse loci in a greater number of samples is expected to greatly increase the ability to determine the geographic origin of poppy samples as well as reveal unique genotypes in geographically-isolated areas.

Conclusions

With the continued increase in deaths related to opioid overdose, it becomes increasingly important to use a variety of methods to contain the spread of these dangerous substances. In this study, we developed and validated a powerful qPCR tool that is able to accurately and reliably quantify DNA from opium poppy, *P. somniferum*. We also determined that the most effective extraction method for poppy seeds, of the three commercial kits tested, was the nexttec™ one-step DNA Isolation Kit for Plants. Additionally, 19 previously reported STR loci were analyzed for potential use in an STR multiplex. Most of these primer sets were found to have suboptimal amplification, be unreliable, or have too low discriminatory power to be useful for individualization of samples. The six most promising loci were incorporated into a multiplex STR panel, and samples from the United States, Canada, United Kingdom, Spain, Holland, Turkey, Tasmania, and Afghanistan were genotyped. The multiplex proved to have low discriminatory power, making it unable to reliably individualize samples or determine their country of origin. Future research should make use of the fully sequenced *P. somniferum* genome and focus on finding new STR markers and/or developing new primers that will be more reliable and discriminatory.

References

1. Labanca F, Ovesnà J, Milella L (2018) Papaver somniferum L. taxonomy, uses and new insight in poppy alkaloid pathways. Phytochem Rev 17: 853-71. doi: 10.1007/s11101-018-9563-3
2. Guo L, Winzer T, Yang X et al (2018) The opium poppy genome and morphinan production. Science (New York, NY). doi: 10.1126/science.aat4096

3. Facchini PJ, De Luca V (2008) Opium poppy and Madagascar periwinkle: model non-model systems to investigate alkaloid biosynthesis in plants. Plant J 54: 763-84
4. Ziegler J, Facchini PJ, Geißler R et al (2009) Evolution of morphine biosynthesis in opium poppy. Phytochemistry 70: 1696-707
5. Rudd RA, Aleshire N, Zibbell JE, Matthew Gladden R (2016) Increases in drug and opioid overdose deaths—United States, 2000–2014. Am J Transplant 16: 1323-7
6. Rudd RA, Seth P, David F, Scholl L (2016) Increases in Drug and Opioid-Involved Overdose Deaths - United States, 2010-2015. MMWR Morbidity And Mortality Weekly Report 65: 1445-52. doi: 10.15585/mmwr.mm655051e1
7. CDC. (2018) Opioid Overdose. Centers for disease control and prevention
8. DEA. (2017) Drugs of abuse - A DEA resource guide - 2017 edition. U.S. Department of justice, Drug enforcement administration. pp. 40-2,5
9. Braye K, Harwood T, Inder R, Beasley R (2007) Poppy seed tea and opiate abuse in New Zealand. Drug Alcohol Rev 26: 215-9
10. Powers D, Erickson S, Swortwood MJ (2018) Quantification of Morphine, Codeine, and Thebaine in Home-Brewed Poppy Seed Tea by LC-MS/MS. J Forensic Sci 63: 1229-35. doi: 10.1111/1556-4029.13664
11. Bailey K, Richards-Waugh L, Clay D, Gebhardt M, Mahmoud H, Kraner JC (2010) Fatality involving the ingestion of phenazepam and poppy seed tea. J Anal Toxicol 34: 527-32
12. Seyani C, Green P, Daniel L, Pegden A (2017) An interesting case of opium tea toxicity. BMJ Case Reports 2017: bcr-2016-218971

13. Lurie IS, Sam SC, Sara ED, Sini P (2013) Determination of heroin and basic impurities for drug profiling by ultra-high-pressure liquid chromatography. Forensic Sci Int 231: 300-5. doi: http://dx.doi.org/10.1016/j.forsciint.2013.06.008

14. Morello DR, Cooper SD, Panicker S, Casale JF. (2010) Signature Profiling and Classification of Illicit Heroin by GC-MS Analysis of Acidic and Neutral Manufacturing Impurities. Blackwell Publishing Ltd. United States. pp. 42.

15. Morello DR, Meyers RP. (1995) Qualitative and Quantitative Determination of Residual Solvents in Illicit Cocaine HCI and Heroin HCI. American Society for Testing and Materials. United States. pp. 957.

16. DEA. (2018) The 2016 Heroin Signature Program Report. Drug Enforcement Agency

17. Marciano MA, Panicker SX, Liddil GD, Lindgren D, Sweder KS. (2018) Development of a Method to Extract Opium Poppy (Papaver somniferum L.) DNA from Heroin. Sci Rep 8(1): 1-15

18. Houston R, Birck M, Hughes-Stamm S, Gangitano D (2016) Evaluation of a 13-loci STR multiplex system for Cannabis sativa genetic identification. Int J Legal Med 130: 635-47

19. Lee EJ, Jin GN, Lee KL, Han MS, Lee YH, Yang MS. (2011) Exploiting Expressed Sequence Tag Databases for the Development and Characterization of Gene-Derived Simple Sequence Repeat Markers in the Opium Poppy (Papaver somniferum L.) for Forensic Applications. Blackwell Publishing Ltd. United States. pp. 1131

20. Lee EJ, Hwang IK, Kim NY et al (2010) An assessment of the utility of universal and specific genetic markers for opium poppy identification. J Forensic Sci 55: 1202-8

21. Masárová V, Mihálik D, Kraic J (2015) In silico retrieving of opium poppy (Papaver somniferum L.) microsatellites. Agriculture (Polnohospodárstvo) 61: 149-56

22. Celik I, Gultekin V, Allmer J, Doganlar S, Frary A (2014) Development of genomic simple sequence repeat markers in opium poppy by next-generation sequencing. Mol Breed 34: 323-34

23. Şelale H, Çelik İ, Gültekin V, Allmer J, Doğanlar S, Frary A (2013) Development of EST-SSR markers for diversity and breeding studies in opium poppy. Plant Breed 132: 344-51

24. Mičianová V, Ondreičková K, Muchová D et al (2017) Forensic application of EST-derived STR markers in opium poppy. Biologia 72: 587-94

25. Linacre A, Gusmao L, Hecht W et al (2011) ISFG: recommendations regarding the use of non-human (animal) DNA in forensic genetic investigations. Forensic Sci Int Genet 5: 501-5

26. Kati V, Le Corre V, Michel S, Jaffrelo L, Poncet C, Délye C (2012) Isolation and characterisation of 11 polymorphic microsatellite markers in Papaver rhoeas L. (corn poppy), a major annual plant species from cultivated areas. Int J Mol Sci 14(1): 470-479

27. SWGDAM (2012) Validation guidelines for DNA analysis methods. Scientific working group on DNA analysis methods

28. Nexttec. DNA Isolation from Plant by nexttec 1 -Step. 3.0 ed Hilgertshausen

29. QIAGEN (2018) DNeasy® Plant Handbook. Hilden, Germany
30. ZyGEM. Plant DNA extraction (for use with phytoGEM- kits). Charlottesville, VA
31. Agilent (2016) Agilent DNA 1000. Santa Clara, CA
32. Applied Biosystems (2014) DNA fragment analysis by capillary electrophoresis. Foster City, CA
33. Holleley CE, Geerts PG (2009) Multiplex Manager 1.0: a cross-platform computer program that plans and optimizes multiplex PCR. Biotechniques 46: 511-7
34. Vallone PM, Butler JM (2004) AutoDimer: a screening tool for primer-dimer and hairpin structures. Biotechniques 37: 226-31
35. QIAGEN. (2009) Type-it® Microsatellite PCR Handbook. Hilden, Germany
36. Thermo Fisher Scientific (2011) BigDye® Direct Cycle Sequencing Kit. Foster City, CA
37. Griffiths R, Barber M, Johnson P et al (1998) New reference allelic ladders to improve allelic designation in a multiplex STR system. Int J Legal Med 111: 267-72
38. Mayr W (1995) DNA recommendations-1994 report concerning further recommendations of the DNA commission of the ISFH regarding PCR-based polymorphisms in STR (short tandem repeat) systems. Vox Sanguinis 69: 70-1
39. Gill P, Brinkmann B, d'Aloja E et al (1997) Considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature. Forensic Sci Int 87: 185-92
40. Parson W, Ballard D, Budowle B et al (2016) Massively parallel sequencing of forensic STRs: considerations of the DNA commission of the International Society

for Forensic Genetics (ISFG) on minimal nomenclature requirements. Forensic Sci Int Genet 22: 54-63

41. Katevatis C, Fan A, Klapperich CM (2017) Low concentration DNA extraction and recovery using a silica solid phase. PloS One 12: e0176848
42. Moore D, Dowhan D (2002) Purification and concentration of DNA from aqueous solutions. Current protocols in molecular biology 59: 2.1. -2.1. 10
43. Lloyd KG, MacGregor BJ, Teske A (2010) Quantitative PCR methods for RNA and DNA in marine sediments: maximizing yield while overcoming inhibition. FEMS Microbiol Ecol 72: 143-51
44. Thiel T, Michalek W, Varshney R, Graner A (2003) Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (Hordeum vulgare L.). Theor Appl Genet 106: 411-22
45. Parthiban S, Govindaraj P, Senthilkumar S (2018) Comparison of relative efficiency of genomic SSR and EST-SSR markers in estimating genetic diversity in sugarcane. 3 Biotech 8: 144
46. Panchy N, Lehti-Shiu M, Shiu S-H (2016) Evolution of gene duplication in plants. Plant Physiol 171: 2294-316

CHAPTER VII

Use of Eucalyptus DNA profiling in a case of illegal logging

This dissertation follows the style and format of *International Journal of Legal Medicine*.

Roman, MG, Gangitano, D, Figueroa A, Solano J, Anabalon, L, Houston, R (2020) Use of Eucalyptus DNA profiling in a case of illegal logging. Submitted to Sci Justice

Abstract

Eucalyptus is grown world-wide for paper pulp, solid wood, and other industries. Theft or illegal cutting of the trees causes hardship to owners of plantations and countries whose economies rely on the sale and export of eucalyptus products. Unfortunately, many of these crimes are not adjudicated due to lack of validated techniques to provide forensic evidence.

Over 1,200 short tandem repeat (STR) markers have been identified in the genomes of genus *Eucalyptus* and related species. However, their importance and utility in aiding forensic investigations of wood theft have not been explored. This study evaluated nine STRs for diversity and applied them to a case involving suspected wood theft.

As expected, three dinucleotide STR markers showed greater variability but resulted in harder to interpret profiles. Four STR tetranucleotide markers evaluated in this study were found to contain additional repeat structures (dinucleotide or trinucleotide) that enhanced their variability but resulted in profiles with peaks at multiple stutter positions and heterozygote peak imbalance. The most promising STR markers were EGM37 and EMBRA 1374. Though less variable, they yielded robust and reproducible DNA profiles.

All nine STR markers were applied to a case involving suspected wood theft. Samples were collected from seized wood and from remaining stumps in a plantation. No DNA match was found, thus eliminating the evidence samples as having originated from the forest. Dendrochronology analysis also resulted in an exclusion. This case study represents the first report using STR markers in any eucalyptus species to provide DNA evidence in a case of suspected wood theft.

Keywords: Forensic plant science, *Eucalyptus*, Illegal logging, Environmental crimes, Short tandem repeat (STR), Forensic genetics

Introduction

Eucalyptus is a genus of long-lived, evergreen trees in the family *Myrtaceae* [1]. Though native to the Southern hemisphere, eucalypts are grown in plantations world-wide and known for their adaptability, fast growth, and important industrial uses [2]. Eucalyptus is also an important tree in Chile, which has two forest regions, one with 13.4 million hectares (ha) of native forests and another comprised of more than 2 million ha of plantations containing exotic species such as *Eucalyptus nitens* and *E. globulus* [3]. In Chile, *E. globulus* is used in agriculture, as firewood, in the fabrication of parquets and foils, and in the production of pulp and paper. It has also a high content of cineol, and its flowers produce nectar and pollen for quality honey [4].

Illegal logging is a key environmental crime involving the harvest, transport, processing, or sale of timber products in violation of national laws [5]. An estimated 15 to 30 percent of timber traded around the world is obtained through illegal logging activities, generating revenues of between 50 to 150 billion USD annually. These actions lead to forest loss and degradation, habitat destruction, and consequent species extinction [6]. According to the National Forest Corporation (CONAF), there were 495 reported cases of wood theft in Chile between 2004 and 2010, and nearly one-third of these crimes occurred on plantations of *E. globulus* [7]. Indeed, the area of native forest affected by illegal logging was estimated to be 10,668 ha during the period 2013-2019. There was an increase in the area affected by illegal activities: from 1,193 ha in 2017 to 1,815 ha in 2018, and a total of 1,359 ha has been affected as of October 2019 [8, 9]. It is important to note, however, that the vast majority of illegal logging around the word is occurring in natural forests. These crimes result in deforestation, damage to local communities, and affect the economies of

producer countries. Unfortunately, cases involving suspected illegal logging are often dismissed without charges due to a lack of evidence to support the crime [10]. However, many genetic markers have been reported in the literature and could be used to provide evidence in illegal logging cases.

Researchers have identified over 1,200 short tandem repeat (STR) markers in various species in the *Eucalyptus* genus and species in the closely related genus *Corymbia* [11-36]. One major resource is the collection of EMBRA STR markers [13, 18, 21, 27, 29, 36], which includes over 800 genomic and expressed sequence tag (EST)-derived STR markers developed in *E. grandis* and *E. urophylla*. Additionally, 40 EMCRC EST-STRs were developed in *E. globulus* [16], *Corymbia variegata* [15], and *Corymbia citriodora* subspecies *variegata* [22]. These markers have been shown to be transferable across species within the *Eucalyptus* and *Corymbia* genera [16, 17, 20, 22, 27, 29, 31, 32, 34, 37-44].

The majority of the developed STR markers contain dinucleotide repeat motifs, and they have been used for diverse applications, including research in conservation and ecology, breeding studies, diversity assessment, parentage testing, and mapping of quantitative trait loci [11]. However, no previous studies have reported the use of STR markers in *Eucalyptus* for forensic analyses related to illegal logging. The usefulness of STR markers in combatting illegal logging has previously been shown in *Neobalanocarpus heimii* (Dipterocarpaceae) [45], *Aquilaria crassna* (agarwood) [46], and *Intsia palembanica* (Leguminosae) [47], in which unique DNA fingerprints were developed to identify individual logs or to identify the subpopulation the logs most likely originated from.

In forensics, guidelines for the analysis of human and non-human DNA have been set by the International Society for Forensic Genetics (ISFG). One such guideline discourages the use di- and trinucleotide STR markers due to increased stutter, which complicates profile interpretation. Instead, ISFG recommends the use of tetranucleotide STR markers to produce more robust and reproducible DNA profiles. Additionally, the development and use of a sequenced allelic ladder is recommended to ensure accurate allele calling between runs and between forensic laboratories [48].

The purpose of this study was to evaluate nine previously reported STR markers in *E. globulus* and *E. nitens* and to demonstrate the utility of STR analysis in a case study involving suspected illegal logging in Chile. Three dinucleotide STR markers were included in the study due to their proven variability in multiple eucalyptus species and the abundance of this type of marker in the literature. In addition, four tetranucleotide, one pentanucleotide, and one hexanucleotide STR loci were assessed.

Methods

Assay development and optimization

Sample collection

One bag of *E. globulus* (Sheffield's Seed Company contained ~ 112 seeds) and two bags of *E. nitens* (purchased from Amazon.com contained approximately 40 seeds, and Sheffield's Seed Company, approximately 112 seeds) were used for method development and optimization at a United States University (blinded). These two species were chosen because they represented the two species present in the forest for the suspected illegal logging case.

DNA extraction and quantification

In order to assess initial variability and amplification of STR markers, DNA was extracted from six randomly selected *E. globulus* seeds and seven randomly selected *E. nitens* seeds (N=13 total) using the DNeasy® Plant Mini Kit (QIAGEN, Hilden, Germany) according to the manufacturer's protocol [49].

DNA was quantified using the Qubit™ dsDNA HS Assay Kit (Thermo Fisher Scientific, South San Francisco, CA, USA) on a Qubit™ 2.0 Fluorometer (Thermo Fisher Scientific) according to the manufacturer's protocol [50].

Dinucleotide STR optimization and amplification

Three dinucleotide STR markers (EMCRC 9, EMCRC 10, and EMCRC 11) [16] were evaluated for genotyping performance. These dinucleotides were chosen because they were previously used to genotype the case samples using gel electrophoresis at the Temuco Regional Crime Laboratory. Amplification conditions were optimized using gradient PCR to determine the optimal annealing temperature, and primer titrations were performed to balance peak heights.

Gradient PCR

In order to determine the optimal annealing temperature by gradient PCR, samples were amplified using the HotStarTaq *Plus* Master Mix Kit (QIAGEN) on an Eppendorf Mastercycler® Gradient Thermal Cycler (Eppendorf, Hamburg, Germany). PCR reactions consisted of 6.25 µL of HotStarTaq *Plus* Master Mix (QIAGEN), 1.25 µL 2 µM primer mix (Integrated DNA Technologies, Coralville, IA, USA), 1.25 µL 5X Q-solution (QIAGEN), 0.4 µL of bovine serum albumin (BSA, 8 mg/mL) (Sigma-Aldrich, St. Louis, MO), 2 µL of DNA, and 1.35 µL of deionized nuclease-free water.

Amplification parameters consisted of a 5 min activation step at 95 °C, followed by 30 cycles of 30 s at 94 °C, 90 s at a gradient of 60 °C ± 10 °C, and 30 s at 72 °C; and a final extension of 30 min at 60 °C. PCR products were visualized on a 2% agarose gel. The highest temperature resulting in a bright band was considered the optimal annealing temperature.

Amplification conditions

Due to timing of extractions, only eleven of the thirteen samples (five *E. globulus* and six *E. nitens*) were amplified in single PCR reactions using the HotStarTaq *Plus* Master Mix Kit (QIAGEN) for the dinucleotide markers. The PCR master mix consisted of 6.25 µL of HotStarTaq *Plus* Master Mix (QIAGEN), 1.25 µL 10X primer mix (Integrated DNA Technologies, Coralville, IA, USA) (Table 7.1), 1.25 µL 5X Q-solution (QIAGEN), and 0.4 µL of bovine serum albumin (BSA; 8 mg/mL) (Sigma-Aldrich, St. Louis, MO). An aliquot of DNA (0.5 ng) was added to the master mix, and deionized nuclease-free water was used to bring the final reaction volume to 12.5 µL. Primer concentrations (10X) for EMCRC 9, EMCRC 10, and EMCRC 11 were 1.8 µM, 1.2 µM, and 1.8 µM, respectively. Annealing temperatures of 61 °C, 56 °C, and 61 °C were used for EMCRC 9, EMCRC 10, and EMCRC 11, respectively. Amplification parameters consisted of a 5 min activation step at 95 °C, followed by 30 cycles of 30 s at 94 °C, 90 s at the optimal annealing temperature, and 30 s at 72 °C; followed by a final extension of 30 min at 60 °C. Positive and negative DNA controls were included in every PCR run.

Table 7.1 STR markers analyzed in this study. N_a indicates the number of alleles observed. T_a indicates the optimal annealing temperature determined by gradient PCR

Marker name	Repeat motif	Allele size range (bp)	N_a	Forward primer	Reverse primer	Reference	T_a (°C)
EMCRC 9	TG	301-337	8	CTGGGCTGTGCATCTCTGAAA	GACCCGGTCAACTCCTCAAGA	[16]	61
EMCRC 10	(GT)(GA)	311-332	6	GCTTGGTCGGGTAGGAA	TCGGGTTGATGTCCTTATTGT	[16]	56
EMCRC 11	(TC)(AC)	226-270	12	AACTGACTGTGGATTTGAAGC	GTGAGTCATTATTTGGCAACC	[16]	61
EMBRA 813	CTCC	77-102	11	ATCTCTCTCGCCGATCTCAA	CGGAGAGATCAAAGGCATGT	[29]	60
EMBRA 925	TCCT	232-253	14	ATCCATCCCACCAAGGAAAT	CGTAGAACTTGGCGAGGAAG	[29]	54
EMBRA 1008	ATCG	154-171	6	AAGCTCGCAGCTCAGAAAAA	GTACTTGTCCTCCGCCATGT	[29]	62
EMBRA 1364	CTCC	313-339	10	CGTTTTCGCTCCTCTCTCTC	TGTAGAGATCGGGGTCCTTG	[29]	57
EGM37	GCTTA	256-281	7	TGAGGTCACTTCAAGCACCAAGA	GGAAGCGGCAACAACCTTAACA	[34]	64
EMBRA 1374	CGCCGT	335-365	6	GTCTGAACTCGGCTTCCTTG	TTCTTCCCGTTGTAAATCCG	[29]	59

Capillary electrophoresis

Capillary electrophoresis (CE) was performed on a 3500 Genetic Analyzer (Thermo Fisher Scientific) with a 50 cm column and POP-7™ polymer. An aliquot of 0.5 µL of PCR product was added to 9 µL of Hi-Di™ Formamide and 0.5 µL GeneScan™ 600 LIZ® Size Standard (Thermo Fisher Scientific). The samples were denatured for 5 min at 95 °C, chilled on ice, then run using the following instrument conditions: oven: 60 °C; pre-run 15 kV, 180 s; injection 1.6 kV, 8 s; run 19.5 kV, 1330 s; and G5 dye set. GeneMapper v.5 software (Thermo Fisher Scientific) was used for data analysis with an analytical threshold of 100 relative fluorescence units (RFU). A custom panel and bin set was developed for each STR marker.

Tetranucleotide STR multiplex development

Amplification conditions

A multiplex assay consisting of four tetranucleotide STR markers (EMBRA 813, EMBRA 925, EMBRA 1008, and EMBRA 1364) [29] was also developed for genotyping of samples. These four markers were chosen for their tetranucleotide nature, amplicons of different sizes (to fit on a single dye channel when multiplexed), and absence of primer interactions, verified with *in-silico* with AutoDimer software [51]. Gradient PCR was performed as described in "2.1.3.1 Gradient PCR", resulting in optimal annealing temperatures of 60 °C, 54 °C, 62 °C, and 57 °C for EMBRA 813, EMBRA 925, EMBRA 1008, and EMBRA 1364, respectively. Primer titrations were carried out to balance peaks across the loci. Multiplex amplification was performed using the Type-it® Microsatellite PCR Kit (QIAGEN). The PCR master mix consisted of 6.25 µL of Type-it® Multiplex PCR Master Mix (QIAGEN), 1.25 µL of 10X primer mix (1 µM EMBRA 813, 1 µM

EMBRA 925, 5 µM EMBRA 1008, and 0.8 µM EMBRA 1364) (Integrated DNA Technologies) (Table 7.1), and 1.25 µL 5× Q-Solution (QIAGEN). An aliquot of 0.5 ng DNA was added, along with deionized nuclease-free water to bring the reaction volume to 12.5 µL. Cycling was carried out on a Veriti™ 96-Well Thermocycler (Thermo Fisher Scientific). To compensate for the differences in annealing temperatures, a touchdown PCR method was used for multiplex amplification. Cycling conditions consisted of a 5 min activation at 95 °C; followed by 2 cycles of 30 s at 95 °C, 90 s at 62 °C, and 30 s at 72 °C; 2 cycles of 30 s at 95 °C, 90 s at 57 °C, and 30 s at 72 °C; 27 cycles of 30 s at 95 °C, 90 s at 54 °C, and 30 s at 72 °C; and a final extension step for 30 min at 60 °C. Positive and negative DNA controls were included in every PCR run. CE was performed as described above in the "Capillary Electrophoresis" section, and data were analyzed with GeneMapper v.5 software using a custom panel and bin set and an analytical threshold of 100 RFU.

Development of an allelic ladder

Homozygous DNA samples with different allele sizes were used to develop an allelic ladder. Similar to previous studies [52, 53], an allelic ladder was prepared by first amplifying samples in single reactions, then combining the amplification products of samples with different alleles for a single locus. Once a ladder was designed for each of the four loci, these were combined and balanced to develop an allelic ladder for the multiplex assay. This ladder was included in every CE run to ensure accurate allele calling.

Penta- and hexanucleotide STR optimization and amplification

Amplification conditions

Single PCR assays were designed to amplify a pentanucleotide STR marker (EGM37) [34] and a hexanucleotide STR marker (EMBRA1374) [29]. Gradient PCR was

performed as described previously, and the optimal annealing temperatures for EGM37 and EMBRA1374 were 64 °C and 59 °C, respectively. Amplification was performed using the HotStarTaq *Plus* Master Mix Kit (QIAGEN). Each reaction consisted of 6.25 µL of HotStarTaq *Plus* Master Mix (QIAGEN), 1.25 µL 10X primer mix (2 µM; Integrated DNA Technologies), 1.25 µL 5X Q-solution (QIAGEN), 0.4 µL of bovine serum albumin (BSA; 8 mg/mL) (Sigma-Aldrich), and 3.35 µL of DNA (0.5 ng).

Amplification parameters on a Veriti™ 96-Well Thermocycler (Thermo Fisher Scientific) consisted of a 5 min activation step at 95 °C, followed by 30 cycles of 30 s at 94 °C, 90 s at the optimal annealing temperature, and 30 s at 72 °C; followed by a final extension of 30 min at 60 °C. Thirteen samples (six *E. globulus* and seven *E. nitens*) were amplified using the optimized conditions. Positive and negative template controls were included in every PCR run. CE was performed as described above, and data were analyzed with GeneMapper v.5 software using a custom panel and bin set and an analytical threshold of 100 RFU.

Allele sequencing

Alleles were sequenced using the BigDye™ Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific). Homozygous samples were amplified with the HotStarTaq *Plus* Master Mix Kit (QIAGEN) as described in the previous section "Amplification Conditions." PCR products were purified using the QIAquick PCR Purification Kit (QIAGEN) according to the manufacturer's instructions [54]. Cycle sequencing was performed according to the BigDye™ Terminator v3.1 Cycle Sequencing Kit user guide [55], and sequencing products were purified using Centri-Sep™ Spin columns (Princeton Separations, Freehold, NJ). Sequence data were analyzed and proof-read using Geneious

Pro v7.1 (Biomatters, Auckland, New Zealand). Allele sequences were submitted to GenBank (accession numbers MN917828-MN917832).

Case study

Case background

On May 20, 2013, local police seized wood logs from trucks in Traiquen, Araucaria region, Chile. The Local Prosecutor's Office requested that Temuco Regional Crime Laboratory establish whether the seized wood came from a private property known as El Avellano. This property contained a *Eucalyptus* forest (*E. globulus* and *E. nitens*), in which wood was processed for commercial purposes. The case involved a criminal investigation into possible wood theft by the defendants, who were accused of illegally cutting wooden logs from *Eucalyptus* trees in El Avellano.

Sample collection

Samples from the two cut trees in *Eucalyptus* forest at El Avellano were collected. Samples were collected for both dendrochronology and DNA analysis. The two trees were morphologically identified as being from two different species: *E. nitens* and *E. globulus.* Samples included a transverse slice of each tree for dendrochronology comparison, a stem still attached to the cut *E. nitens* tree (EN212) for DNA comparison, and a leaf still attached to the cut *E. globulus* tree (EG214). Subsequently, wood transverse slices from the seized eucalyptus logs morphologically identified as being *E. globulus* and *E. nitens* found in the trucks were collected for dendrochronology comparison. Additionally, samples for DNA comparison were taken from the truck: slice of wood from *E. nitens* tree (EN219) and stem still attached to *E. globulus* tree in truck (EG2023). To note, stems and leaves were preferentially chosen when possibly as it is easier to extract DNA from those sources of

the tree compared to the trunk. *E. nitens* (EN-213) and *E. globulus* (EG-213) control samples were kindly provided to the Chilean Regional Crime Laboratory by a local eucalyptus grower, and morphological analysis was performed at the laboratory for species identification.

Dendrochronology analysis

Physical characterization of the collected wood material and the analysis of annual tree-ring widths and other related parameters (minimum, maximum, and average density, earlywood width, and more) were performed at the Chilean Regional Crime Laboratory using the Windendro software (Regent Instruments Inc., Canada).

DNA extraction and quantification

DNA extraction from leaf, stem, and transverse slices was performed using the DNeasy® Plant Mini Kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions [49] at Temuco Regional Crime Laboratory. DNA extracts were quantified by fluorometry on a Qubit™ 4.0 Fluorometer (Thermo Fisher Scientific) using the Qubit™ dsDNA HS Assay Kit (Thermo Fisher Scientific) according to the manufacturer's protocol [50].

Amplification of dinucleotide STR markers

Analysis of the three dinucleotide STR markers (EMCRC 9, EMCRC 10, and EMCRC 11) occurred at Temuco Regional Crime Laboratory. A triplex PCR was performed using the Type-it® Multiplex PCR Master Mix (QIAGEN). A 2 µL aliquot of sample containing 1 ng of DNA was added to 13 µL of master mix. PCR amplification was carried out on a in a Rotor-Gene® PCR Thermocycler (QIAGEN). PCR products were run on a 4% high-resolution agarose gel (Sigma-Aldrich).

Amplification of tetra-, penta-, and hexanucleotide STR markers

Six DNA extracts were received for genotyping at a United States University (blinded). The extracts were representative of all individuals tested by the Chilean Regional Crime Laboratory; however, duplicate samples from the same individual were not included. The extracts were labeled with the species name (three *E. nitens* and three *E. globulus*) but were otherwise processed in a blind fashion. Tetranucleotide STR markers were analyzed as described in the section "Tetranucleotide STR Multiplex Development." Penta- and hexanucleotide STR markers were analyzed as described in the section "Penta- and Hexanucleotide STR Optimization and Amplification" with the exception that a multiplex amplification was performed. The primer mix consisted of 2 µM of each primer, and an annealing temperature of 59 °C was used.

Results

Assay development

Dinucleotide STRs

Only four out of eleven samples produced a full genotype (Fig. 7.1, Table 7.2). Five *E. nitens* samples failed to produce a peak for EMCRC 10. EMCRC 11 was the most polymorphic of the three loci with 12 alleles ranging from 226-270 bp, followed by EMCRC 9 with 8 alleles ranging from 301-337 bp, and EMCRC 10 was the least polymorphic with 6 alleles ranging from 311-332 bp (Table 7.1). Additionally, the presence of high stutter ratios made profile interpretation difficult (Fig. 7.1).

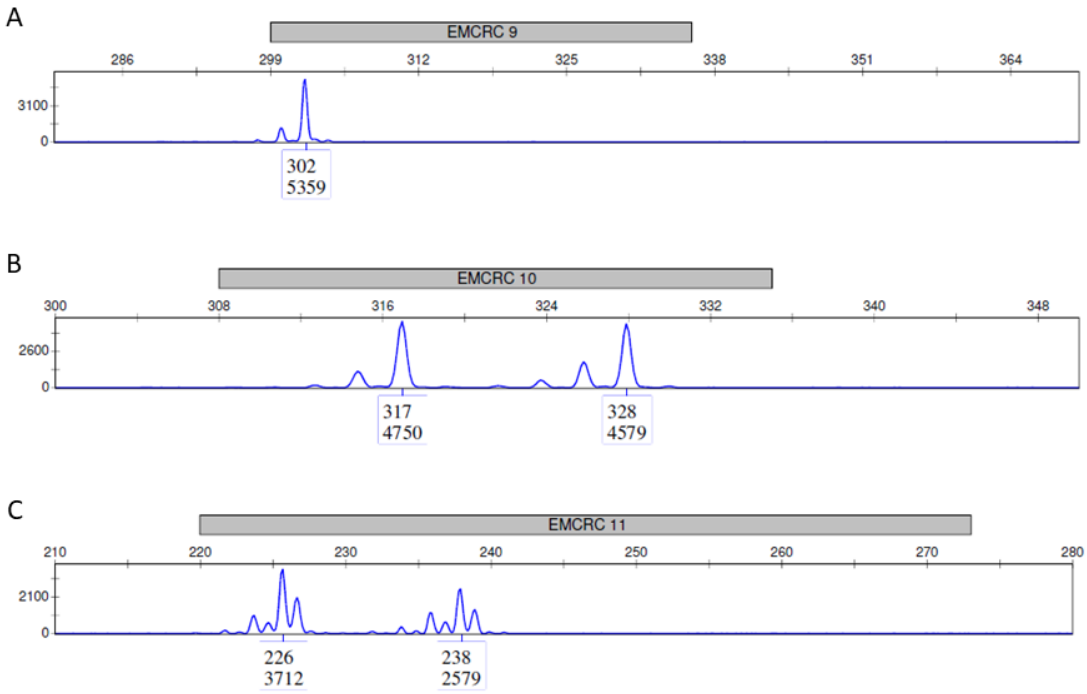


Fig. 7.1 Example dinucleotide STR profiles of an *E. globulus* sample for EMCRC 9 (A), EMCRC 10 (B), and EMCRC 11 (C)

Tetranucleotide STRs

All thirteen samples produced full profiles for the tetranucleotide EMBRA markers (Fig. 7.2, Table 7.2). However, many of the observed alleles differed by only one to three base pairs (bp) instead of the expected four bp apart for tetranucleotide markers. Due to the multiple repeat structures within these EMBRA markers, the authors chose to name the alleles by their bp size rather than the number of tetranucleotide repeats. Alleles for EMBRA 813 ranged from 77-102 bp (11 alleles), EMBRA 925 ranged from 232-253 bp (14 alleles), EMBRA 1008 ranged from 154-171 bp (6 alleles), and EMBRA 1364 ranged from 313-339 bp (10 alleles) (Table 7.1).

Table 7.2 Genotypes of developmental (seed) samples. U indicates a profile which was uninterpretable, and X indicates an unknown allele

Sample	Source	EMCRC 9	EMCRC 10	EMCRC 11	EMBRA 813	EMBRA 925	EMBRA 1008	EMBRA 1364	EGM37	EMBRA 1374
EN-1	Amazon.com	322	No peaks	250, 252	87, 90	244, 250	171	332, 339	256, 266	341
EN-2	Amazon.com	308, 334	327	234, 270	87, 94	241	154	331	256, 266	341, 359
EN-3	Amazon.com	302, 322	No peaks	234, 242	99, 102	241	159, 171	325, 332	266, 281	353, 365
EN-5	Amazon.com	Not tested	Not tested	Not tested	81, 92	240, 248	165	331	266, 276	347, 353
EN-B4	Sheffield's Seed Co.	337	No peaks	234, 260	90, 94	237, 253	U	313, 321	261, 266	347, 365
EN-B5	Sheffield's Seed Co.	320	No peaks	238, 254	90	239, 241	165	321	256, 261	353
EN-B6	Sheffield's Seed Co.	325	No peaks	236, 238	87, 90	237	165, 169	X, 327	261	347
EG-B1	Sheffield's Seed Co.	No peaks	311, 317	226, 230	94, 96	242, 245	162	327, 330	266	341, 347
EG-B2	Sheffield's Seed Co.	302	317, 328	226, 238	83, 96	245	162	329	266	335, 353
EG-B3	Sheffield's Seed Co.	301, 302	315, 328	234, 246	81, 94	241	165	327	267, 271	353
EG-B4	Sheffield's Seed Co.	301	328	230, 234	81, 96	238, 245	162, 165	327, 337	266, 267	335, 353
EG-B5	Sheffield's Seed Co.	No peaks	315, 332	226, 230	85, 94	X, 238	165	321, 329	266, 276	353, 365
EG-B6	Sheffield's Seed Co.	Not tested	Not tested	Not tested	96	241, 245	162, 165	329	267	335, 341

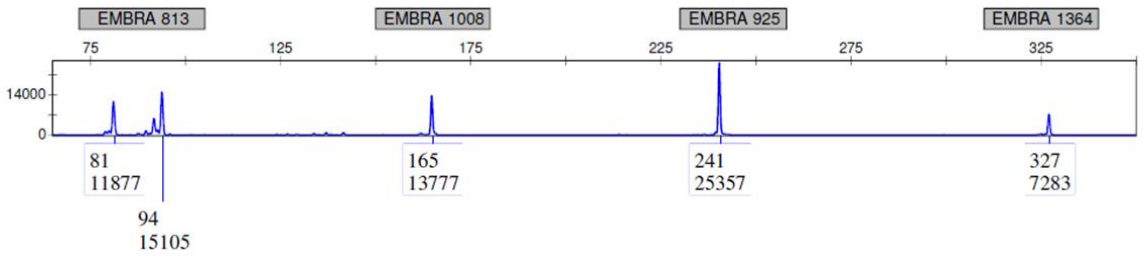


Fig. 7.2 Example tetranucleotide multiplex STR profile of an *E. globulus* sample

Allelic ladder

An allelic ladder was developed (Fig. 7.3) which contained ten alleles for EMBRA 813, eight alleles for EMBRA 925, five alleles for EMBRA 1008, and eight alleles for EMBRA 1364. Due to multiple repeat structures, some of which contained di- or trinucleotide motifs, elevated stutter peaks are apparent in the ladder electropherogram.

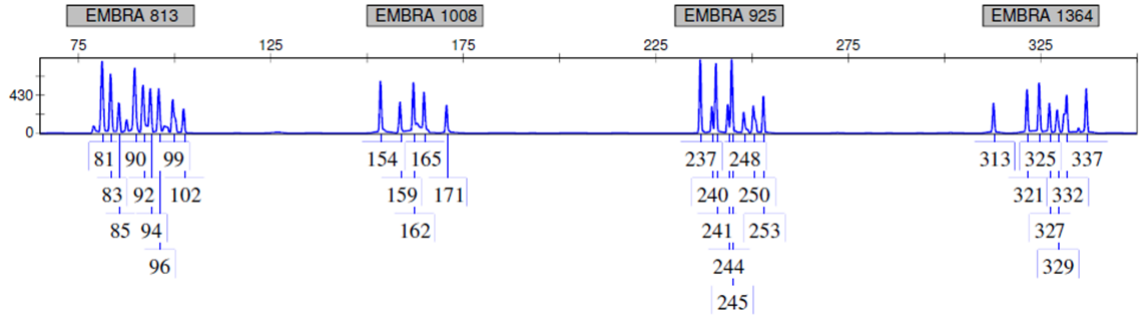


Fig. 7.3 Allelic ladder for the tetranucleotide STR multiplex

Penta- and hexanucleotide STRs

All thirteen samples produced full genotypes for both EGM37 and EMBRA 1374 (Fig. 7.4). Seven alleles were observed for EGM37 (256-281 bp), and six alleles were observed for EMBRA 1374 (335-365 bp) (Table 7.1). Profile interpretation for these markers was considerably easier due to low stutter produced during amplification. Allele sequencing was performed on three homozygous samples per locus. For EGM37,

sequencing of the 261 bp and 266 bp alleles revealed four and five repeats of the GCTTA motif. Sequencing of the 267 bp allele revealed a complex repeat with the structure $(GCTTA)_2(GCTTTA)_1(GCTTA)_2$ (Fig. 7.5; GenBank accessions: MN917828-MN917830). EMBRA 1374 alleles 347 bp and 353 bp (two samples) were also sequenced, revealing three and four repeats of the motif CGCCGT, respectively (Fig. 7.6; GenBank accessions: MN917831 and MN917832).

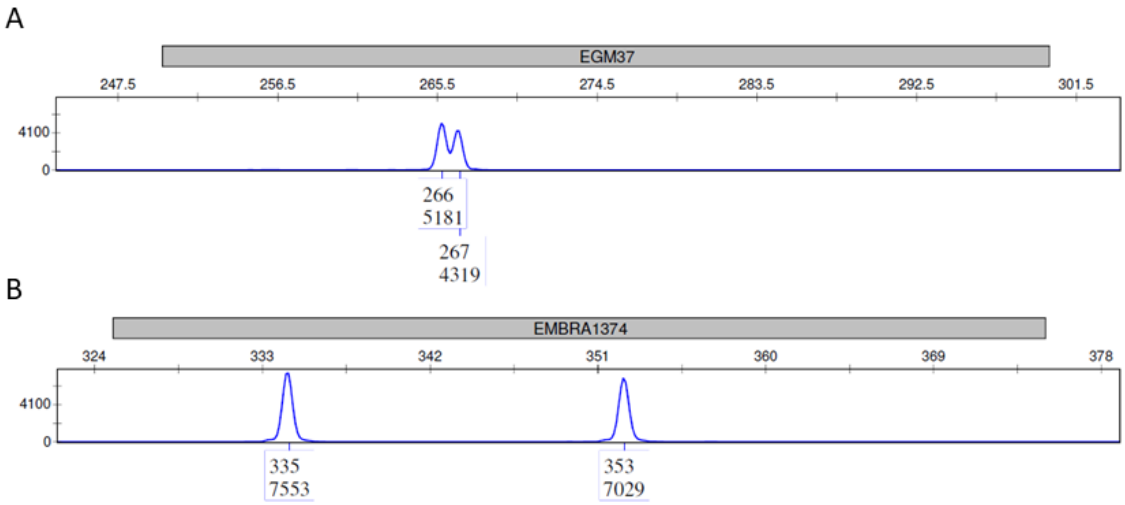


Fig. 7.4 Example profiles of an *E. globulus* sample for EGM37 (A) and EMBRA 1374 (B) loci

This is a note for Fig. 7.5 and 7.6: In the consensus sequences, the forward and reverse primer binding sites are underlined, and the location of the repeat structure is indicated as **[REPEAT]. The variable repeat structure is described in the table. The sequence data taken from the literature is also referenced in the table. N refers to the total number of samples sequenced in this study bearing the genotype described.

TGAGGTCACTTCAAGCACCAAGAATCAACAAAACTCTATTGTTCTTCCATCTCCCACC
AATTCACACACAATT**[REPEAT]**GCTTTTGCTTATTCAATGGCTTCCGACTCAGATGTC
ACCATCCCCTCGAAGATCAAGGCTTGGGTCTATGCCGAACGTGGGAACCCCGCCGAT
GTCTTGAAGCTGGTTCCAGACGTTGGGGTTCCTGAGATTAAGGAAGACCAGGTCCTT
GTTAAGGTTGTTGCCGCTTCC

Allele	[REPEAT]	N	Accession No.	Reference
261	$(GCTTA)_4$	1	MN917828	This study
266	$(GCTTA)_5$	1	MN917829	This study, [REF]*
267	$(GCTTA)_2(GCTTTA)(GCTTA)_2$	1	MN917830	This study

Fig. 7.5 Consensus sequence of the EGM37 locus, allele nomenclature, and genotypes observed
*Allele reported by Bradbury, Smithson, and Krauss (2013) (33) but not previously submitted to GenBank

GTCTGAACTCGGCTTCCTTGCTTCCGCAGCTCCGGCCACTTCC**[REPEAT]**CGGACTCA
AGGTCCGTCGACCTTTCTCTCTCAATCTTCTTCTCGCCGTTCGGCCGCTGAACAGCGG
CTCTGTTGTCCTTCTTTTCCTCTTCTTTCAGTTCCGGGAAGCAAGGGCCCGAACCGAA
GGCGAAGGAATGACGCGATCCTTGTGGAAGGGCATCTTCGTCGACGCGTTCCTGTCG
AAGATCAAGGACCGGAGGGACCTCCTCTCGAACAAGAAGATCTGGTCGCGCCGGTC
CTCCATCCTGCCGGAGTTCGTCAACACCACCATGCGGATTTACAACGGGAAGAA

Allele	[REPEAT]	N	Accession No.	Reference
347	(CGCCGT)3	1	MN917831	This study
353	(CGCCGT)4	2	MN917832	This study

Fig. 7.6 Consensus sequence of the EMBRA 1374 locus, allele nomenclature, and genotypes observed
*Allele reported by Faria et al. (2010) (28) is GenBank accession GF101859, but sequence differs by 87 bases which are missing in GF101859 (marked in red in the above sequence).

Case study

Dendrochronology

The transverse wood samples taken from tree stumps in the private forest showed a uniform growth pattern with a regular age, estimated to be 8 years old, consistent with a commercial plantation such as El Avellano. Material obtained from the seized wood logs showed an irregular growth pattern with an age estimated to be from 5 to 19 years, consistent with a wild forest. Therefore, the dendrochronology results preliminarily exclude the El Avellano trees as being the source of the wood logs found in the seized truck.

Dinucleotide STRs

DNA extracted from the samples in truck and private forest were genotyped at the Temuco Regional Crime Laboratory using the poorly performing dinucleotide STR markers. The DNA profiles produced from the wood in the truck did not match the band patterns from those produced by the leaf and stem samples from the private forest, resulting in an exclusion (data not shown). Though these markers were not analyzed with capillary electrophoresis, it is evident from the large number of bands observed that high levels of stutter occurred during PCR amplification.

Tetranucleotide STRs

One wood slice (EN-219) and a stem (EG-2023) were analyzed from the seized truck, and a stem (EN-212) and leaf (EG-214) were analyzed from the private forest. Additionally, two control samples (EN-213 and EG-213) were genotyped. All six samples resulted in complete STR genotypes (Table 7.3). Profile interpretation was complicated by a high stutter ratio and the presence of microvariants, which were also observed during

marker screening. All four loci were polymorphic within the case study samples; EMBRA 813, EMBRA 925, and EMBRA 1364 each showed five alleles per locus, and EMBRA 1008 showed two alleles. No DNA match was found between reference (private forest – El Avellano) and evidence samples (seized logs from truck), confirming the results obtained by the Temuco Regional Crime Laboratory.

Table 7.3 Genotypes of the six case study samples

Sample ID	Sample type	EMBRA 813	EMBRA 925	EMBRA 1008	EMBRA 1364	EGM37	EMBRA 1374
EN-213	Control DNA	85, 94	247	165	325, 329	261	347
EG-213	Control DNA	85, 96	238	162, 165	332	266, 276	353, 365
EN-219	Slice from truck	92	241	165	313, 321	261	341, 347
EG-2023	Stem from truck	94	234, 238	165	321, 327	266, 276	335, 365
EN-212	Stem from forest	77, 94	232, 241	165	321	261	347
EG-214	Leaf from forest	94, 96	238, 241	162, 165	321, 327	266, 276	341, 353

*EN indicates *E. nitens* and EG indicates *E. globulus.*

Penta- and hexanucleotide STRs

The same six samples genotyped with the tetranucleotide STRs were also genotyped with EGM37 and EMBRA 1374. Two samples shared a common genotype (EN-213, a control, and EN-212, a sample from the private forest), indicating that these two markers alone are not sufficient for individualization of samples. In particular, the locus EGM37 only produced two different genotypes for the case study samples. However, none of the samples from El Avellano or the seized wood produced matching profiles, confirming the exclusionary results obtained by dendrochronology and use of the dinucleotide and tetranucleotide STR markers.

Discussion

Assay development

Dinucleotide STRs

The EMCRC markers were originally developed in *E. globulus*, and only EMCRC 9 and EMCRC 11 were demonstrated to be transferable to *E. nitens* [16]. Therefore, it is not surprising that EMCRC 10 performed poorly in *E. nitens* samples. Additionally, EMCRC 9 failed to amplify in two *E. globulus* samples, which could be due to a mutation in the primer-binding site for these samples, resulting in a null allele. Previous literature has shown that null alleles are frequently encountered when analyzing STR markers in eucalyptus [11, 17, 21], and the potential for null alleles increases when using these markers across species [27, 34, 43, 44]. Because dinucleotides exhibit high stutter, ISFG guidelines suggest the use of tetranucleotide STR markers for forensic analysis to reduce the effect of stutter on profile interpretation [48]. Due to the unreliability and difficult interpretation guidelines for these dinucleotide markers, they were not ideal for forensic use.

Tetranucleotide STRs

Although all samples yielded full profiles, internal dinucleotide repeat structures resulted in harder to interpret profiles. Using tetranucleotide markers, alleles are expected to be a multiple of four bp apart. Faria et al. [29] published their sequences to GenBank (NCBI): accessions GF101851, GF101853, GF101855, and GF101858. Though not commented on by Faria et al., the sequences of all four loci displayed additional repeat structures containing dinucleotide or trinucleotide motifs. EMBRA 813 contained a dinucleotide repeat with the motif $(CT)_{15}$; EMBRA 925 contained a dinucleotide repeat

with the motif $(CT)_{10}$; EMBRA 1008 contained a complex trinucleotide repeat with the motif $(CGG)_6(CAG)(CGG)$; and EMBRA 1364 contained a trinucleotide repeat with the motif $(CGG)_7$ as well as a dinucleotide repeat partially within the forward primer-binding region with the motif $(CT)_6$. Because dinucleotide and trinucleotide STRs are more variable by nature [56], it is likely these repeat structures account for most of the variation observed.

Many of the STR loci from existing literature also contain additional repeat motifs, and these markers tend to be the most polymorphic. For example, the two most polymorphic loci from Glaubitz et al. [17], which were dinucleotide STRs, also contained homopolymeric stretches, causing many of the alleles to differ by a single bp. This study was able to overcome the challenges in allele scoring by incorporating internal reference alleles into each sample analyzed; however, this would not overcome the challenge introduced by elevated stutter.

Penta- and hexanucleotide STRs

Profile interpretation for these markers was considerably easier due to low stutter produced during amplification. Faria et al. [29] previously reported the sequence of EMBRA 1374 (GenBank GF101859); although our sequences were similar, we observed an 87 bp stretch following the repeat motif that was not observed in their original sequence (marked in red in Fig. 7.6). The authors recommend that EGM37 and EMBRA 1374 be further explored in future studies, along with several other promising markers reported in the literature [14, 25, 27, 31].

Case study

The dendrochronology results in this case preliminarily ruled out the felled trees in El Avellano as the source of the seized logs. However, the reliability and limitations of

using tree ring patterns for this purpose requires further investigation [10, 57]. The United Nations Office on Drugs and Crime's (UNODC) Best Practice Guide for Forensic Timber Identification [58] recommends the use of dendrochronology to provide information about growing conditions and minimum tree age, with the possibility of also being able to provide information about the provenance of the tree in some cases. Pattern matching of a log with a stump may be useful for individualization purposes; however, it is unclear whether the ring pattern is consistent between pieces of wood from a single individual [10] and trees of the same species in the same geographical area that are influenced by the same environmental conditions tend to show similar ring patterns [59].

DNA is a more robust, specific, and reliable tool for individualization purposes, and the UNODC's guide also refers to using STR loci for individualization of timber to link seized materials with an illegally felled tree or remaining stump, or for verifying chain of custody along the supply route. Methods should mimic the model of human identification when possible, with highly degraded DNA and polyploidy being complications to consider when analyzing DNA in wood [58].

Of the three marker systems used for individualization of case study samples (dinucleotide, tetranucleotide, and penta- and hexanucleotide markers), the dinucleotide markers showed the most variability, but the penta- and hexanucleotide markers produced the best quality profiles. ISFG recommends that to ensure quality results, non-human forensic DNA analysis should conform to standards similar to human identification, including the use of STR markers with at least four units in the repeat motif (tetranucleotide or larger) [48]. Therefore, the authors recommend further analysis of EGM37 and

EMBRA1374 loci in future studies, as well as the evaluation of additional tetra-, penta-, and hexanucleotide STR loci.

Conclusions

The results of this project highlight the utility of STR analysis in *Eucalyptus* as evidence to combat illegal logging. Nine STR markers (three dinucleotide, four tetranucleotide, one pentanucleotide, and one hexanucleotide) were evaluated. Although the STR markers containing dinucleotide repeat motifs were more variable than others (6-14 alleles per locus), they also produced high stutter peaks and unbalanced heterozygote alleles, making profile interpretation for these markers difficult. Additionally, EMCRC 10, originally developed in *E. globulus*, was not transferable to *E. nitens*. The tetranucleotide markers EMBRA 813, EMBRA 925, and EMBRA 1364 were found to also include dinucleotide STR motifs within the amplicons, and it is likely that these dinucleotide motifs account for the majority of the diversity seen in these markers. Similarly, the tetranucleotide markers EMBRA 1008 and EMBRA 1364 contained trinucleotide motifs within the amplicons (EMBRA 1364 contained three repeat motifs in total). The pentanucleotide marker EGM37 produced seven alleles, and the hexanucleotide marker EMBRA 1374 produced six alleles; though they were less variable, the profiles were much easier to interpret due to low levels of stutter, a single repeat motif, and more balanced heterozygote peaks.

The three sets of STR markers (dinucleotide markers EMCRC 9, EMCRC 10, and EMCRC 11; tetranucleotide markers EMBRA 813, EMBRA 925, EMBRA 1008, and EMBRA 1364; and penta- and hexanucleotide markers EGM37 and EMBRA 1374) were applied to case study samples involving suspected wood theft from a private eucalyptus

forest. All three sets of markers resulted in an exclusionary result, indicating that the evidence samples did not match the samples from the private forest. These results strengthened the exclusionary result obtained from dendrochronology analysis, which indicated that the evidence samples showed different ages and growth patterns compared to the forest samples.

In future studies, additional STRs should be tested which have motifs of at least four nucleotides and do not contain repeat structures in the flanking regions in order to ensure accurate allele calling and avoid multiple stutter peaks. Additional STR markers are available in the literature which may merit additional testing for forensic use [14, 25, 27, 31].

References

1. Ladiges PY, Udovicic F, Nelson G (2003) Australian biogeographical connections and the phylogeny of large genera in the plant family Myrtaceae. Journal Biogeogr 30: 989-98
2. Potts BM, Vaillancourt RE, Jordan G et al (2004) Exploration of the Eucalyptus globulus gene pool
3. Lara A, Little C, Urrutia R et al (2009) Assessment of ecosystem services as an opportunity for the conservation and management of native forests in Chile. For Ecol Manag 258: 415-24
4. Goic P, Balaguer T, Bustamante J et al. (1994) Atlas forestal de Chile. Publicacion CIREN-INFOR 103 Santiago de Chile
5. World Wide Fund for Nature (2019) Illegal Logging

6. Koch GT (2019) Best practices on tools for combating and preventing illegal logging activities and associated trade. Asia-Pacific Economic Cooperation
7. CONAF (2014) Memoria CONAF 2010-2014. Corporación Nacional Forestal Ministerio de Agricultura 235
8. CONAF (2018) Memoria CONAF 2014-2018. Corporación Nacional Forestal Ministerio de Agricultura
9. (2019) Chile has lost more than 10,000 hectares of native forest due to illegal logging. Diario Concepción Santiago, Chile
10. Dormontt EE, Boner M, Braun B et al (2015) Forensic timber identification: It's time to integrate disciplines to combat illegal logging. Biol Conserv 191: 790-8
11. Sumathi M, Yasodha R (2014) Microsatellite resources of Eucalyptus: current status and future perspectives. Bot Stud 55: 73
12. Byrne M, Marquezgarcia M, Uren T, Smith D, Moran G (1996) Conservation and genetic diversity of microsatellite loci in the genus Eucalyptus. Aust J Bot 44: 331-41
13. Brondani RPV, Brondani C, Tarchini R, Grattapaglia D (1998) Development, characterization and mapping of microsatellite markers in Eucalyptus grandis and E. urophylla. Theor Appl Genet 97: 816-27
14. Van der Nest M, Steenkamp E, Wingfield B, Wingfield M (2000) Development of simple sequence repeat (SSR) markers in Eucalyptus from amplified inter-simple sequence repeats (ISSR). Plant Breed 119: 433-6

15. Jones ME, Stokoe RL, Cross MJ, Scott LJ, Maguire TL, Shepherd M (2001) Isolation of microsatellite loci from spotted gum (Corymbia variegata), and cross-species amplification in Corymbia and Eucalyptus. Mol Ecol Notes 1: 276-8

16. Steane D, Vaillancourt R, Russell J, Powell W, Marshall D, Potts B (2001) Development and characterisation of microsatellite loci in Eucalyptus globulus (Myrtaceae). Silvae Genet 50: 89-91

17. Glaubitz J, Emebiri L, Moran G (2001) Dinucleotide microsatellites from Eucalyptus sieberi: inheritance, diversity, and improved scoring of single-base differences. Genome 44: 1041-5

18. Brondani R, Brondani C, Grattapaglia D (2002) Towards a genus-wide reference linkage map for Eucalyptus based exclusively on highly informative microsatellite markers. Mol Genet Genom 267: 338-47

19. Thamarus K, Groom K, Murrell J, Byrne M, Moran G (2002) A genetic linkage map for Eucalyptus globulus with candidate loci for wood, fibre, and floral traits. Theor Appl Genet 104: 379-87

20. Ottewell K, Donnellan S, Moran G, Paton D (2005) Multiplexed microsatellite markers for the genetic analysis of Eucalyptus leucoxylon (Myrtaceae) and their utility for ecological and breeding studies in other Eucalyptus species. J Hered 96: 445-51

21. Brondani RP, Williams ER, Brondani C, Grattapaglia D (2006) A microsatellite-based consensus linkage map for species of Eucalyptus and a novel set of 230 microsatellite markers for the genus. BMC Plant Biol 6: 20

22. Shepherd M, Kasem S, Lee D, Henry R (2006) Construction of microsatellite linkage maps for Corymbia. Silvae Genet 55: 228-38

23. Payn KG, Dvorak WS, Janse BJ, Myburg AA (2008) Microsatellite diversity and genetic structure of the commercially important tropical tree species Eucalyptus urophylla, endemic to seven islands in eastern Indonesia. Tree Genet Genom 4: 519-30

24. Yasodha R, Sumathi R, Chezhian P, Kavitha S, Ghosh M (2008) Eucalyptus microsatellites mined in silico: survey and evaluation. J Genet 87: 21-5

25. da Silva JM, de Sousa ACB, de Souza AP et al (2009) Development and characterization of 14 microsatellite loci from an enriched genomic library of Eucalyptus camaldulensis Dehnh. Conserv Genet Res 1: 465

26. Sexton TR, Henry RJ, McManus LJ, Bowen S, Shepherd M (2010) Capture of assay template by multiplex PCR of long amplicons for genotyping SNPs and InDels with MALDI-TOF mass spectrometry. Mol Breed 25: 471-80

27. Faria DA, Mamani EMC, Pappas MR, Pappas Jr GJ, Grattapaglia D (2010) A selected set of EST-derived microsatellites, polymorphic and transferable across 6 species of Eucalyptus. J Hered 101: 512-20

28. Zhou C, Li F, Weng Q, Yu X, Li M, Gan S (2010) Comparison between direct sequencing and pool-cloning-based sequencing of PCR products in EST-SSR marker development in Eucalyptus. Mol Plant Breed 8: 1-10

29. Faria DA, Mamani EMC, Pappas GJ, Grattapaglia D (2011) Genotyping systems for Eucalyptus based on tetra-, penta-, and hexanucleotide repeat EST microsatellites and

their use for individual fingerprinting and assignment tests. Tree Genet Genom 7: 63-77

30. Acuña C, Villalba P, García M, Hopp E, Marcucci S. (2011) Functional markers development and genetic diversity analysis in Eucalyptus globulus with emphasis in wood quality candidate genes. BMC Proceedings. BioMed Central. pp. P154
31. Acuña CV, Villalba PV, García M, Pathauer P, Hopp HE, Marcucci Poltri SN (2012) Microsatellite markers in candidate genes for wood properties and its application in functional diversity assessment in Eucalyptus globulus. Electron J Biotechnol 15: 2
32. He X, Wang Y, Li F et al (2012) Development of 198 novel EST-derived microsatellites in Eucalyptus (Myrtaceae). Am J Bot 99: e134-e48
33. Nevill PG, Bradbury D, Jørgensen T, Krauss S, Samaraweera S, Gardner MG (2013) Microsatellite primers identified by 454 sequencing in the floodplain tree species Eucalyptus victrix (Myrtaceae). Appl Plant Sci 1: 1200402
34. Bradbury D, Smithson A, Krauss SL (2013) Development and testing of new gene-homologous EST-SSRs for Eucalyptus gomphocephala (Myrtaceae). Appl Plant Sci 1: 1300004
35. Zhou C, He X, Li F et al (2014) Development of 240 novel EST-SSRs in Eucalyptus L'Hérit. Mol Breed 33: 221-5
36. Grattapaglia D, Mamani EM, Silva-Junior OB, Faria DA (2015) A novel genome-wide microsatellite resource for species of Eucalyptus with linkage-to-physical correspondence on the reference genome sequence. Mol Ecol Res 15: 437-48
37. Bundock PC, Hayden M, Vaillancourt RE (2000) Linkage maps of Eucalyptus globulus using RAPD and microsatellite markers. Silvae genet 49: 223-32

38. Arumugasundaram S, Ghosh M, Veerasamy S, Ramasamy Y (2011) Species discrimination, population structure and linkage disequilibrium in Eucalyptus camaldulensis and Eucalyptus tereticornis using SSR markers. PLoS One 6: e28252

39. Subashini V, Shanmugapriya A, Yasodha R (2014) Hybrid purity assessment in Eucalyptus F 1 hybrids using microsatellite markers. 3 Biotech 4: 367-73

40. Neves LG, Mamani EM, Alfenas AC, Kirst M, Grattapaglia D (2011) A high-density transcript linkage map with 1,845 expressed genes positioned by microarray-based Single Feature Polymorphisms (SFP) in Eucalyptus. Bmc Genom 12: 189

41. Hudson CJ, Freeman JS, Kullan AR et al (2012) A reference linkage map for Eucalyptus. BMC Genom 13: 240

42. Petroli CD, Sansaloni CP, Carling J et al (2012) Genomic characterization of DArT markers based on high-density linkage analysis and physical mapping to the Eucalyptus genome. PLoS One 7: e44684

43. Bradbury D, Krauss SL (2013) Limited impact of fragmentation and disturbance on the mating system of tuart (Eucalyptus gomphocephala, Myrtaceae): implications for seed-source quality in ecological restoration. Aust J Bot 61: 148-60

44. Bradbury D, Smithson A, Krauss SL (2013) Signatures of diversifying selection at EST-SSR loci and association with climate in natural Eucalyptus populations. Mol Ecol 22: 5112-29

45. Tnah L, Lee S, Ng K, Faridah Q, Faridah-Hanum I (2010) Highly variable STR markers of Neobalanocarpus heimii (Dipterocarpaceae) for forensic DNA profiling. J Trop For Sci: 214-26

46. Eurlings MC, van Beek HH, Gravendeel B (2010) Polymorphic microsatellites for forensic identification of agarwood (Aquilaria crassna). Forensic Sci Int 197: 30-4

47. Lowe AJ, Wong K-N, Tiong Y-S, Iyerh S, Chew F-T (2010) A DNA method to verify the integrity of timber supply chains; confirming the legal sourcing of merbau timber from logging concession to sawmill. Silvae Genet 59: 263-8

48. Linacre A, Gusmao L, Hecht W et al (2011) ISFG: recommendations regarding the use of non-human (animal) DNA in forensic genetic investigations. Forensic Sci Int Genet 5: 501-5. doi: 10.1016/j.fsigen.2010.10.017

49. QIAGEN (2018) DNeasy® Plant Handbook. Hilden, Germany

50. Thermo Fisher Scientific (2015) Qubit® dsDNA HS assay kits protocol

51. Vallone PM, Butler JM (2004) AutoDimer: a screening tool for primer-dimer and hairpin structures. Biotechniques 37: 226-31

52. Griffiths R, Barber M, Johnson P et al (1998) New reference allelic ladders to improve allelic designation in a multiplex STR system. Int J Legal Med 111: 267-72

53. Houston R, Birck M, Hughes-Stamm S, Gangitano D (2016) Evaluation of a 13-loci STR multiplex system for Cannabis sativa genetic identification. Int J Legal Med 130: 635-47

54. QIAGEN (2018) QIAquick® Spin Handbook. Hilden, Germany

55. Thermo Fisher Scientific (2016) BigDye™ Terminator v3.1 Cycle Sequencing Kit User Guide. South San Francisco, CA

56. Chakraborty R, Kimmel M, Stivers DN, Davison LJ, Deka R (1997) Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. Proc Natl Acad Sci 94: 1041-6

57. Dormontt E, Jardine D, van Dijk K-J et al (2020) Forensic validation of a SNP and INDEL panel for individualisation of timber from bigleaf maple (Acer macrophyllum Pursch). Forensic Sci Int Genet 46: 102252

58. Wiedenhoeft A (2016) Best practice guide for forensic timber identification. In: United Nations Office on Drugs and Crime: International Consortium on Combating Wildlife Crime Vienna, Austria: Laboratory and Scientific Section, Global Programme for Combating Wildlife and Forest Crime: 1-226

59. Creber G (1977) Tree rings: a natural data-storage system. Biol Rev 52: 349-81

CHAPTER VIII

Conclusions

The applications of forensic plant science in investigating drug trafficking and providing evidence of illegal logging have not been fully realized due to lack of research. Marijuana (*Cannabis sativa*, L.) is the most commonly used illicit drug in the United States, and law enforcement is tasked with investigating trafficking of the drug into the country and across state lines. This project explored the potential of chloroplast DNA barcoding markers and nuclear genes controlling an enzyme involved in cannabinoid production to determine crop type and biogeographical origin of *C. sativa*. Due its high abuse potential, heroin (derived from *Papaver somniferum*) is also commonly trafficked in the U.S., and this work evaluates short tandem repeat (STR) markers for individualization and biogeographical origin determination of *P. somniferum* samples. Finally, illegal logging costs the world economy billions of dollars annually and results in deforestation, but these crimes are rarely prosecuted due to lack of forensic evidence. STR markers for *Eucalyptus* species were evaluated for their ability to individualize *Eucalyptus globulus* and *Eucalyptus nitens* samples and were applied to a case study involving suspected wood theft.

A total of 58 chloroplast polymorphisms (31 homopolymeric STRs, hSTRs; 23 single nucleotide polymorphisms, SNPs; and four insertion-deletions, INDELs) were identified by comparing database sequences of the chloroplast genome of *C. sativa*. Seven hotspots with three or more polymorphisms within 1,600 bp were selected, and four of these regions (*rpl32-trnL, trnS-trnG, rps16,* and *clpP*) were evaluated for variation among hemp samples from Canada and the U.S., marijuana from the U.S.-Mexico border and Chile, and medical marijuana from Chile. Capillary electrophoresis (CE)-based fragment

analysis and single base extension (SBE) assays were developed to easily and economically analyze the variability of these polymorphic regions and determine whether they could be used to differentiate between sample populations based on crop type and biogeographical origin. Sequenced allelic ladders were created to ensure accurate allele calling in fragment analysis assays, and the sequences of alleles were reported to GenBank.

Haplotype analysis of the nine polymorphisms in *rpl32-trnL* and *trnS-trnG* hotspots (N=152) resulted in eight haplogroups. Four additional polymorphisms (an hSTR, two SNPs, and an INDEL) and sequence variants (isoalleles) for *rpl32-trnL* hSTR3 were also detected by Sanger sequencing but could not be identified using the fragment-based assays. Haplotype 1 was by far the most common and was observed in all sample groups except Canadian hemp, which was completely differentiated from the other four groups. Phylogenetic analysis using the neighbor-joining method and pairwise comparisons using F_{ST} as genetic distance revealed close genetic relatedness between the USA-Mexico marijuana and Chilean medical marijuana groups and the Chilean medical marijuana, Chilean marijuana, and USA hemp groups. Overall, these two hotspot regions provided a limited ability to differentiate between sample groups based on crop type and biogeographical origin. It was expected that the use of a sequencing assay and analysis of all seven hotspot regions would provide better discrimination.

Next, nine loci within *rps16* and *clpP* hotspots were evaluated. Two SNPs were found to be monomorphic, but eight haplotypes consisting of the other seven loci were observed following genotyping of 166 samples from the same five sample populations as the previous study. Haplotypes 1 and 2 were observed in 46% and 40% of samples tested, respectively. Similar to the previous study with *rpl32-trnL* and *trnS-trnG*, only Canadian

hemp samples could be completely differentiated from the other four groups. Chilean marijuana showed a close relatedness to USA-Mexico marijuana and Chilean medical marijuana, and Chilean medical marijuana and USA hemp were also closely related according to phylogenetic analysis and pairwise comparisons. When the haplotypes at all four hotspots are combined, ten haplogroups are formed, indicating that a combination of the markers provides better differentiation between populations, as expected. While these markers alone are not enough to determine biogeographical origin or crop type of an unknown sample, they do show promise in differentiating between populations. The results suggest that the addition of haplotypes at the other three hotspot regions would likely add to the discriminatory power of the analysis. Additionally, the true discriminatory power of these hotspots will not be known until an extensive database can be created by genotyping diverse samples from around the world.

A massively parallel sequencing (MPS) assay was designed using the MiSeq FGx to simultaneously analyze all seven hotspot regions. Data was collected from 14 samples, and MPS and CE data were concordant at all shared loci. In addition to the 30 loci analyzed by CE, 19 additional loci were analyzed, 16 of which had not been previously reported. Isoalleles were identified at one locus, further increasing the discriminatory power of the assay. The high throughput capability of this MPS assay provides a tool for analyzing a large number of samples in future studies to build a world-wide database, which is necessary before unknown samples can be assigned a crop type or biogeographical origin.

Another approach for determining crop type of *C. sativa* samples is to genotype polymorphisms in the gene coding for tetrahydrocannabinolic acid (THCA) synthase, which plays a role in biosynthesis of the psychoactive chemical of marijuana, delta-9-

tetrahydrocannabinol (THC). A previously reported SBE assay was optimized and used to genotype four SNPs in 47 hemp and marijuana samples from the previous studies. The method resulted in one false negative and several false positives, with an overall error rate of 19%. False positives occurred in nearly all USA hemp samples from a single supplier and strains of USA hemp that were marketed as high in cannabigerol (CBG), indicating that this type of analysis has limitations and should be substituted by other analyses, such as using the DNA barcoding markers described in Chapters II-IV of this dissertation.

A novel quantitative real-time PCR (qPCR) assay was developed for *P. somniferum* DNA. It was proven to be reproducible and precise, sensitive, and specific for *P. somniferum*. Following a comparison of three commercial DNA extraction kits, the nexttec™ one-step DNA Isolation Kit for Plants was chosen as the optimal extraction method for poppy seeds due to a higher average DNA yield and faster processing time. Many STR markers for opium poppy are available in the scientific literature; however, few studies have been performed to test their variability or forensic utility. In this work, nineteen STR markers were evaluated for their individualization potential. Two failed to amplify, three produced nonspecific amplification products, two had suboptimal peak morphology, and eight were monoallelic. The remaining four loci (psom12, psom13, psom16, and psom17) had 2-3 alleles each. A preliminary multiplex using six loci was developed and validated according to the standards of the International Society for Forensic Genetics (ISFG). A sequenced allelic ladder was used to ensure accurate allele calling, and validation parameters included sensitivity, specificity, reproducibility, and precision. The multiplex was used to genotype 63 poppy seed DNA samples from around the world. It was found to lack sufficient discriminatory power to individualize samples or determine

their biogeographical origin. The authors recommend that the now fully sequenced poppy genome be searched for new STR markers.

Nine STR markers were evaluated for individualization of *E. globulus* and *E. nitens*, including three dinucleotide, four tetranucleotide, one pentanucleotide, and one hexanucleotide markers. As expected, the dinucleotide loci showed the highest variability but produced high stutter and unbalanced heterozygote peaks, making profile interpretation more difficult. The tetranucleotide loci were shown to be highly variable due to additional di- or trinucleotide repeat motifs within the amplified regions, which resulted in high stutter and alleles that varied by only 1-3 base pairs (bp). The penta- and hexanucleotide loci were somewhat less variable but produced high-quality profiles. All nine loci were used to genotype samples from a case study involving suspected wood theft, and the results excluded the fallen logs found in a private forest as the source of the seized logs. This study represented the first case report using STR markers in *Eucalyptus* to provide evidence of illegal logging activities. Future studies should focus on finding additional STR loci with tetranucleotide or larger repeat motifs and evaluating their variability and transferability among relevant *Eucalyptus* species.

In summary, this work explores the use of forensic plant science and genetic techniques to provide investigative leads for drug trafficking investigations and evidence of environmental crimes. Forensic plant science is an underutilized resource in the forensic science community, and its applications need to be brought to the attention of the forensic community as well as law enforcement personnel and lawyers through peer-reviewed research, publications, and presentations at professional conferences. The DNA analysis of plants uses similar principles, methodologies, and equipment as human DNA analysis, but

it is seldom used in crime laboratories due to a lack of knowledge and research in the field. Forensic plant science relies on principles well known and accepted by ecologists, botanists, molecular biologists, and other relevant scientific experts, and additional research on forensically relevant species will help to satisfy the Daubert standard requirements and provide valuable evidence in criminal proceedings.

Future research areas related to this dissertation include building a worldwide database of *C. sativa* samples using the MPS assay described in Chapter IV and identifiying more reliable and discriminatory STR markers for individualizing *P. somniferum* and *Eucalyptus*. Other plants which merit further research for forensic drug trafficking investigations include *Erythroxylum coca* and *novogranatense* (cocaine) and *Mitragyna speciosa* (kratom). Full genome sequences have been reported for *E. novogranatense* and *M. speciosa*. STR markers for *M. speciosa* would be of particular forensic interest, as the DEA lists kratom as a drug of concern, and though it is not currently scheduled, several states have already banned its use. Discovery of chloroplast polymorphisms for determining the biogeographical origin of these plants, particularly *P. somniferum* (heroin), would also benefit the forensic community by aiding in drug trafficking investigations. Some researchers have discovered STR markers in trees that are subject to illegal logging, including *Aquilaria crassna* (agarwood), *Intsia palembanica* (Leguminosae), and others. STRs for these species could be used to verify chain of custody for legal trees and to prevent trade of illegal timber products, as well as to provide evidence of illegal logging in court. However, more research is needed to develop robust STR assays for forensic purposes.

REFERENCES

1. (2009) CU Student Dies Brewing Poppy Tea. The Denver Channel
2. (2009) Mother warns of poppy tea 'risks' after son's overdose. BBC News
3. (2012) N.S. family warns of poppy seed tea dangers. CBC News
4. (2016) Friends recall Boulder man who died after drinking poppy tea. The Denver Post Denver, CO
5. (2018) Agriculture Improvement Act of 2018 (115-334)
6. (2018) WHO Expert Committee on Drug Dependence Critical Review: Isomers of THC
7. (2019) Chile has lost more than 10,000 hectares of native forest due to illegal logging. Diario Concepción Santiago, Chile
8. Acharya HS, Sharma V (2009) Molecular Characterization of Opium Poppy. American Journal of Infectious Diseases 5: 155-60
9. Acuña C, Villalba P, García M, Hopp E, Marcucci S. (2011) Functional markers development and genetic diversity analysis in Eucalyptus globulus with emphasis in wood quality candidate genes. BMC Proceedings. BioMed Central. pp. P154
10. Acuña CV, Fernandez P, Villalba PV, García MN, Hopp HE, Poltri SNM (2012) Discovery, validation, and in silico functional characterization of EST-SSR markers in Eucalyptus globulus. Tree Genet Genomes 8: 289-301
11. Acuña CV, Villalba PV, García M, Pathauer P, Hopp HE, Marcucci Poltri SN (2012) Microsatellite markers in candidate genes for wood properties and its application in functional diversity assessment in Eucalyptus globulus. Electron J Biotechn 15: 2-2
12. Substance Abuse and Mental Health Services Administration (2018) Key substance

use and mental health indicators in the United States: results from the 2017 National Survey on Drug Use and Health (HHS Publication No. SMA 18-5068, NSDUH Series H-53). Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration Rockville, MD

13. Agilent (2016) Agilent DNA 1000. Santa Clara, CA
14. U.S. Department of Agriculture (2019) Testing guidelines for identifying delta-9 tetrahydrocannabinol (THC) concentration in hemp
15. Alghanim H, Almirall J (2003) Development of microsatellite markers in Cannabis sativa for DNA typing and genetic relatedness analyses. Anal Bioanal Chem 376:1225–1233
16. Ali B, Al-Wabel NA, Shams S, Ahamad A, Khan SA, Anwar F (2015) Essential oils used in aromatherapy: A systemic review. Asian Pac J Trop Biomed 5: 601-11
17. Allgeier L, Hemenway J, Shirley N, LaNier T, Coyle HM (2011) Field testing of collection cards for Cannabis sativa samples with a single hexanucleotide DNA marker. J Forensic Sci 56: 1245-9
18. Anderson S, Bankier AT, Barrell BG et al (1981) Sequence and organization of the human mitochondrial genome. Nature 290: 457-65
19. Applied Biosystems (2014) DNA fragment analysis by capillary electrophoresis. Foster City, CA
20. Armenian P, Vo KT, Barr-Walker J, Lynch KL (2018) Fentanyl, fentanyl analogs and novel synthetic opioids: a comprehensive review. Neuropharmacol 134: 121-32

21. Arumugasundaram S, Ghosh M, Veerasamy S, Ramasamy Y (2011) Species discrimination, population structure and linkage disequilibrium in Eucalyptus camaldulensis and Eucalyptus tereticornis using SSR markers. PLoS One 6: e28252
22. Ashton CH (2001) Pharmacology and effects of cannabis: a brief review. Br J Psychiatry 178: 101-6
23. Bachir RG, Benali M (2012) Antibacterial activity of the essential oils from the leaves of Eucalyptus globulus against Escherichia coli and Staphylococcus aureus. Asian Pac J Trop Biomed 2: 739-42
24. Bailey K, Richards-Waugh L, Clay D, Gebhardt M, Mahmoud H, Kraner JC (2010) Fatality involving the ingestion of phenazepam and poppy seed tea. J Anal Toxicol 34: 527-32
25. Baker P, Taylor B, Gough T (1981) The tetrahydrocannabinol and tetrahydrocannabinolic acid content of cannabis products. J Pharm Pharmacol 33: 369-72
26. Barnfield C, Burns S, Byrom DL, Kemmenoe AV (1988) The routine profiling of forensic heroin samples. Forensic Sci Int 39: 107-17
27. Basak SS, Candan F (2010) Chemical composition and in vitro antioxidant and antidiabetic activities of Eucalyptus camaldulensis Dehnh. essential oil. J Iran Chem Soc 7: 216-26
28. Batish DR, Singh HP, Kohli RK, Kaur S (2008) Eucalyptus essential oil as a natural pesticide. For Ecol Manag 256: 2166-74
29. Bell KL, De Vere N, Keller A et al (2016) Pollen DNA barcoding: current applications and future prospects. Genome 59: 629-40

30. Bentley DR, Balasubramanian S, Swerdlow HP et al (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456: 53-9
31. Benyahia S, Benayache S, Benayache F et al (2004) Isolation from eucalyptus o ccidentalis and identification of a new Kaempferol derivative that induces apoptosis in human myeloid leukemia cells. J Nat Prod 67: 527-31
32. Birky CW (1995) Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. Proc Natl Acad Sci 92: 11331-8
33. Bock JH, Norris DO (1997) Forensic botany: an under-utilized resource. J Forensic Sci 42: 364-7
34. Bock JH, Norris DO. (2015) Forensic plant science. Academic Press
35. Booth M. (2013) Opium: a history. St. Martin's Griffin
36. Borna T, Salami SA, Shokrpour M (2017) High resolution melting curve analysis revealed SNPs in major cannabinoid genes associated with drug and non-drug types of cannabis. Biotechnol Biotechnol Equip 31: 839-45
37. Bradbury D, Krauss SL (2013) Limited impact of fragmentation and disturbance on the mating system of tuart (Eucalyptus gomphocephala, Myrtaceae): implications for seed-source quality in ecological restoration. Aust J Bot 61: 148-60
38. Bradbury D, Smithson A, Krauss SL (2013) Signatures of diversifying selection at EST-SSR loci and association with climate in natural Eucalyptus populations. Mol Ecol 22: 5112-29
39. Braye K, Harwood T, Inder R, Beasley R (2007) Poppy seed tea and opiate abuse in New Zealand. Drug Alcohol Rev 26: 215-9

40. Brondani R, Brondani C, Grattapaglia D (2002) Towards a genus-wide reference linkage map for Eucalyptus based exclusively on highly informative microsatellite markers. Mol Genet and Genom 267: 338-47

41. Brondani RP, Williams ER, Brondani C, Grattapaglia D (2006) A microsatellite-based consensus linkage map for species of Eucalyptus and a novel set of 230 microsatellite markers for the genus. BMC Plant Biol 6: 20

42. Brondani RPV, Brondani C, Tarchini R, Grattapaglia D (1998) Development, characterization and mapping of microsatellite markers in Eucalyptus grandis and E. urophylla. Theor Appl Genet 97: 816-27

43. Bruijns B, Tiggelaar R, Gardeniers H (2018) Massively parallel sequencing techniques for forensics: a review. Electrophoresis 39: 2642-54

44. Budowle B, Schmedes SE, Wendt FR (2017) Increasing the reach of forensic genetics with massively parallel sequencing. Forensic Science, Med Pathol 13: 342-9

45. Bundock PC, Hayden M, Vaillancourt RE (2000) Linkage maps of Eucalyptus globulus using RAPD and microsatellite markers. Silvae Genetica 49: 223-32

46. Butler J, Vaillancourt R, Potts B et al (2017) Comparative genomics of Eucalyptus and Corymbia reveals low rates of genome structural rearrangement. BMC Genom 18: 397

47. Butler JM. (2011) Advanced topics in forensic DNA typing: methodology. Academic Press

48. Byrne M, Hankinson M (2012) Testing the variability of chloroplast sequences for plant phylogeography. Aust J Bot 60:569–574

49. Byrne M, Marquezgarcia M, Uren T, Smith D, Moran G (1996) Conservation and genetic diversity of microsatellite loci in the genus Eucalyptus. Aust J Bot 44: 331-41

50. Cannabis, Coca, & Poppy: Nature's Addictive Plants. DEA Museum and Visitors Center Arlington, VA

51. Cascini F, Farcomeni A, Migliorini D et al (2019) Highly Predictive Genetic Markers Distinguish Drug-Type from Fiber-Type Cannabis sativa L. Plants 8: 496

52. CDC (2018) Opioid Overdose. Centers for Disease Control and Prevention

53. Celik I, Gultekin V, Allmer J, Doganlar S, Frary A (2014) Development of genomic simple sequence repeat markers in opium poppy by next-generation sequencing. Mol Breed 34: 323-34

54. Centers for Disease Control and Prevention (2018) Opioid Overdose. Centers for Disease Control and Prevention

55. Chakraborty R, Kimmel M, Stivers DN, Davison LJ, Deka R (1997) Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. Proc Natl Acad Sci 94: 1041-6

56. Chandra R, Sharma V (2014) Forensic botany: an emerging discipline of plant sciences. Indian Botanists Blog-o-Journal: 1-8

57. Chen P, Braithwaite R, George C et al (2014) The poppy seed defense: a novel solution. Drug Test Anal 6: 194-201

58. Ciccarone D (2009) Heroin in brown, black and white: structural factors and medical consequences in the US heroin market. Int Drug Policy 20: 277-82

59. Cirovic N, Kecmanovic M, Keckarevic D, Markovic MK (2017) Differentiation of Cannabis subspecies by THCA synthase gene analysis using RFLP. J Forensic Legal Med 51: 81-4

60. CONAF (2014) Memoria CONAF 2010-2014. Corporación Nacional Forestal Ministerio de Agricultura 235

61. CONAF (2018) Memoria CONAF 2014-2018. Corporación Nacional Forestal Ministerio de Agricultura

62. Cone EJ, Dickerson S, Paul BD, Mitchell JM (1993) Forensic drug testing for opiates. V. Urine testing for heroin, morphine, and codeine with commercial opiate immunoassays. J Anal Toxicol 17: 156-64

63. Coyle HM, Ladd C, Palmbach T, Lee HC (2001) The green revolution: botanical contributions to forensics and drug enforcement. Croat Med J 42: 340-5

64. Coyle HM, Palmbach T, Juliano N, Ladd C, Lee HC (2003) An overview of DNA methods for the identification and individualization of marijuana. Croat Med J 44: 315-21

65. Creber G (1977) Tree rings: a natural data-storage system. Biol Rev 52: 349-81

66. Csaikl U, Bastian H, Brettschneider R et al (1998) Comparative analysis of different DNA extraction protocols: a fast, universal maxi-preparation of high quality plant DNA for genetic evaluation and phylogenetic studies. Plant Mol Biol Rep 16: 69-86

67. da Silva JM, de Sousa ACB, de Souza AP et al (2009) Development and characterization of 14 microsatellite loci from an enriched genomic library of Eucalyptus camaldulensis Dehnh. Conserv Genet Resour 1: 465

68. Darben T, Cominos B, Lee C (1998) Topical eucalyptus oil poisoning. Australas J Dermatol 39: 265-7

69. Darling AE, Mau B, Perna NT (2010) Progressivemauve: multiple genome alignment with gene gain, loss and rearrangement. PLOS One 5:e11147

70. Datwyler SL, Weiblen GD (2006) Genetic variation in hemp and marijuana (Cannabis sativa L.) according to amplified fragment length polymorphisms. J Forensic Sci 51:371–375

71. Davis C, Peters D, Warshauer D, King J, Budowle B (2015) Sequencing the hypervariable regions of human mitochondrial DNA using massively parallel sequencing: enhanced data acquisition for DNA samples encountered in forensic testing. Legal Med 17: 123-7

72. Dawoud A, Shayoub M (2015) Effects of ethanolic leaf extract of Eucalyptus camaldulensis on oral glucose tolerance test in Type-2 Model diabetic rats. J Netw Commun Emerg Technol (JNCET) www jncet org 2

73. de Meijer EP, Bagatta M, Carboni A et al (2003) The inheritance of chemical phenotype in Cannabis sativa L. Genetics 163: 335-46

74. De Pauw MA, Vidmar JJ, Collins J, Bennett RA, Deyholos MK (2007) Microarray analysis of bast fibre producing tissues of Cannabis sativa identifies transcripts associated with conserved and specialised processes of secondary wall development. Funct Plant Biol 34: 737-49

75. DEA. (2017) Drugs of abuse - A DEA resource guide - 2017 edition. U.S. Department of justice, Drug Enforcement Administration. pp. 40-2,5

76. DEA (2018) The 2016 Heroin Domestic Monitor Program (DEA-DCW-DIR-026-18). Drug Enforcement Agency

77. DEA (2018) The 2016 Heroin Signature Program Report. Drug Enforcement Agency

78. Deiana S, Watanabe A, Yamasaki Y et al (2012) Plasma and brain pharmacokinetic profile of cannabidiol (CBD), cannabidivarine (CBDV), Delta(9)-tetrahydrocannabivarin (THCV) and cannabigerol (CBG) in rats and mice following oral and intraperitoneal administration and CBD action on obsessive-compulsive behaviour. Psychopharmacology (Berl) 219: 859-73. doi: 10.1007/s00213-011-2415-0

79. Demesure B, Sodzi N, Petit R (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Mol Ecol 4:129–134

80. Dermen H (1940) Colchicine polyploidy and technique. Bot Rev 6: 599-635

81. Devane WA, Dysarz Fr, Johnson MR, Melvin LS, Howlett AC (1988) Determination and characterization of a cannabinoid receptor in rat brain. Mol Pharmacol 34: 605-13

82. Dey B, Mitra A (2013) Chemo-profiling of eucalyptus and study of its hypoglycemic potential. World J Diabetes 4: 170

83. Dhakad AK, Pandey VV, Beg S, Rawat JM, Singh A (2018) Biological, medicinal and toxicological significance of Eucalyptus leaf essential oil: a review. J Sci Food Agr 98: 833-48

84. Dias V, Ribeiro A, Mello I et al (2015) Genetic identification of Cannabis sativa using chloroplast trnL-F gene. Forensic Sci Int Genet 14: 201-2

85. Dong W, Liu J, Yu J, Wang L, Zhou S (2012) Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. PLoS One 7:e35071

86. Dormontt E, Jardine D, van Dijk K-J et al (2020) Forensic validation of a SNP and INDEL panel for individualisation of timber from bigleaf maple (Acer macrophyllum Pursch). Forensic Sci Int Genet 46: 102252

87. Dormontt EE, Boner M, Braun B et al (2015) Forensic timber identification: It's time to integrate disciplines to combat illegal logging. Biol Conserv 191: 790-8

88. Doyle J. (1991) DNA protocols for plants. Molecular techniques in taxonomy. Springer. pp. 283-93

89. Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue

90. Drug Enforcement Administration (2010) Pharmacist's manual: an informational outline of the Controlled Substances Act. Springfield: United States Department of Justice, Drug Enforcement Administration, Office of Diversion Control

91. Drug Enforcement Administration (2017) Drugs of abuse - A DEA resource guide - 2017 edition. U.S. Department of justice, Drug Enforcement Administration. pp. 40-2,5

92. Drug Enforcement Administration (2018) 2018 National Drug Threat Assessment. In: U.S. Department of Justice, ed.

93. Drug Enforcement Administration (2019) Unwashed Poppy Seed

94. Duarte D (2005) Opium and opioids: a brief history. Revista Brasileira de Anestesiologia 55: 135-46

95. Dufresnes C, Jan C, Bienert F, Goudet J, Fumagalli L (2017) Broad-scale genetic diversity of Cannabis for forensic applications. PloS One 12

96. Duke J, Lind C, Mackiewicz K et al (2015) Towards allele-level human leucocyte antigens genotyping–assessing two next-generation sequencing platforms: Ion Torrent Personal Genome Machine and Illumina MiSeq. Int J Immunogenetics 42: 346-58

97. Dumolin-Lapegue S, Pemonge M, Petit R (1997) An enlarged set of consensus primers for the study of organelle DNA in plants. Mol Ecol 6:393–397

98. Dunbar-Co S, Wieczorek AM, Morden CW (2008) Molecular phylogeny and adaptive radiation of the endemic Hawaiian Plantago species (Plantaginaceae). Am J Bot 95:1177–1188

99. Dussy FE, Hamberg C, Luginbühl M, Schwerzmann T, Briellmann TA (2005) Isolation of Δ9-THCA-A from hemp and analytical aspects concerning the determination of Δ9-THC in cannabis products. Forensic Sci Int 149: 3-10

100. History.com Editors. (2017) Heroin, Morphine and Opiates. A&E Television Networks

101. Eduardoff M, Santos C, De La Puente M et al (2015) Inter-laboratory evaluation of SNP-based forensic identification by massively parallel sequencing using the Ion PGM™. Forensic Sci Int Genet 17: 110-21

102. ElSohly MA, Mehmedic Z, Foster S, Gon C, Chandra S, Church JC (2016) Changes in cannabis potency over the last 2 decades (1995–2014): analysis of current data in the United States. Biol Psychiatry 79: 613-9

103. ElSohly MA, Radwan MM, Gul W, Chandra S, Galal A. (2017) Phytochemistry of Cannabis sativa L. Phytocannabinoids. Springer. pp. 1-36

104. Erixon P, Oxelman B (2008) Whole-gene positive selection, elevated synonymous substitution rates, duplication, and indel evolution of the chloroplast clpP1 gene. PLoS One 3:e1386

105. Eurlings MC, van Beek HH, Gravendeel B (2010) Polymorphic microsatellites for forensic identification of agarwood (Aquilaria crassna). Forensic Sci Int 197: 30-4

106. Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564–567

107. F Kelly J, Addanki K, Bagasra O (2012) The non-specificity of the Duquenois-Levine field test for marijuana. The Open Forensic Science Journal 5

108. Facchini PJ, De Luca V (2008) Opium poppy and Madagascar periwinkle: model non-model systems to investigate alkaloid biosynthesis in plants. Plant J 54: 763-84

109. Faeti V, Mandolino G, Ranalli P (1996) Genetic diversity of Cannabis sativa germplasm based on RAPD markers. Plant Breed 115: 367-70

110. Fan H, Chu J-Y (2007) A brief review of short tandem repeat mutation. Genomics, Proteomics & Bioinformatics 5: 7-14

111. Faria DA, Mamani EMC, Pappas GJ, Grattapaglia D (2011) Genotyping systems for Eucalyptus based on tetra-, penta-, and hexanucleotide repeat EST microsatellites and their use for individual fingerprinting and assignment tests. Tree Genet Genom 7: 63-77

112. Faria DA, Mamani EMC, Pappas MR, Pappas Jr GJ, Grattapaglia D (2010) A selected set of EST-derived microsatellites, polymorphic and transferable across 6 species of Eucalyptus. J Heredity 101: 512-20
113. FBI. (2020) Quality assurance standards for forensic DNA testing laboratories
114. Federal Bureau of Investigation. Lindbergh Kidnapping
115. Ferri G, Alù M, Corradini B, Beduschi G (2009) Forensic botany: species identification of botanical trace evidence using a multigene barcoding approach. Int J Legal Med 123: 395-401
116. Ferri G, Corradini B, Ferrari F, Santunione AL, Palazzoli F, Alu M (2015) Forensic botany II, DNA barcode for land plants: Which markers after the international agreement? Forensic Sci Int Genet 15: 131-6
117. Flachowsky H, Schumann E, Weber W, Peil A (2001) Application of AFLP for the detection of sex-specific markers in hemp. Plant Breed 120: 305-9
118. Forapani S, Carboni A, Paoletti C, Moliterni V, Ranalli P, Mandolino G (2001) Comparison of hemp varieties using random amplified polymorphic DNA markers. Crop Sci 41: 1682-9
119. Foster D. (1997) DEA Tries to Weed Out Opium Poppy Seeds From Gardens. Los Angeles Times
120. Gao C, Xin P, Cheng C et al (2014) Diversity analysis in Cannabis sativa based on large-scale development of expressed sequence tag-derived simple sequence repeat markers. PloS One 9
121. Gaoni Y, Mechoulam R (1964) Isolation, structure, and partial synthesis of an active constituent of hashish. J Am Chem Soc 86: 1646-7

122. Ghalem BR, Mohamed B (2008) Antibacterial activity of leaf essential oils of Eucalyptus globulus and Eucalyptus camaldulensis. Afr J Pharm Pharmaco 2: 211-5

123. Gigliano GS (1998) Identification of Cannabis sativa L. (Cannabaceae) using restriction profiles of the Internal Transcribed Spacer II (ITS2). Sci Justice 38:225–230

124. Gigliano GS, Caputo P, Cozzolino S (1997) Ribosomal DNA analysis as a tool for the identification of Cannabis sativa L. specimens of forensic interest. Sci Justice 37:171–174

125. Gigliano SS (1999) Preliminary data on the usefulness of internal transcribed spacer I (ITS1) sequence in Cannabis sativa L. identification. J Forensic Sci 44: 475-7

126. Gill P, Brinkmann B, d'Aloja E, Andersen J, Bar W, Carracedo A, Dupuy B, Eriksen B, Jangblad M, Johnsson V, Kloosterman AD, Lincoln P, Morling N, Rand S, Sabatier M, Scheithauer R, Schneider P, Vide MC (1997) Considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature. Forensic Sci Int 87:185–192

127. Gill P, Brinkmann B, d'Aloja E et al (1997) Considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature. Forensic Sci Int 87: 185-92

128. Gillan R, Cole M, Linacre A, Thorpe J, Watson N (1995) Comparison of Cannabis sativa by random amplification of polymorphic DNA (RAPD) and HPLC of cannabinoids: a preliminary study. Science & justice: journal of the Forensic Science Society 35: 169-77

129. Gilmore S, Peakall R (2003) Isolation of microsatellite markers in Cannabis sativa L.(marijuana). Mol Ecol Res 3: 105-7

130. Gilmore S, Peakall R, Robertson J (2003) Short tandem repeat (STR) DNA markers are hypervariable and informative in Cannabis sativa: implications for forensic investigations. Forensic Sci Int 131: 65-74

131. Gilmore S, Peakall R, Robertson J (2007) Organelle DNA haplotypes reflect crop-use characteristics and geographic origins of Cannabis sativa. Forensic Sci Int 172:179–190

132. Gitzendanner MA (2012) Use and guidelines for plant DNA analyses in forensics. Forensic Botany: A Practical Guide: 93-106

133. Glaubitz J, Emebiri L, Moran G (2001) Dinucleotide microsatellites from Eucalyptus sieberi: inheritance, diversity, and improved scoring of single-base differences. Genome 44: 1041-5

134. Godwin ID, Aitken EA, Smith LW (1997) Application of inter simple sequence repeat (ISSR) markers to plant genetics. Electrophoresis 18: 1524-8

135. Goic P, Balaguer T, Bustamante J et al. (1994) Atlas forestal de Chile. Publicacion CIREN-INFOR 103 Santiago de Chile

136. Graham SA (1997) Anatomy of the Lindbergh kidnapping. J Forensic Sci 42: 368-77

137. Grassa CJ, Wenger JP, Dabney C et al (2018) A complete Cannabis chromosome assembly and adaptive admixture for elevated cannabidiol (CBD) content. BioRxiv: 458083

138. Grattapaglia D, Bradshaw Jr H (1994) Nuclear DNA content of commercially important Eucalyptus species and hybrids. Can J For Res 24: 1074-8

139. Grattapaglia D, Mamani EM, Silva-Junior OB, Faria DA (2015) A novel genome-wide microsatellite resource for species of Eucalyptus with linkage-to-physical correspondence on the reference genome sequence. Mol Ecol Res15: 437-48

140. Griffiths R, Barber M, Johnson P et al (1998) New reference allelic ladders to improve allelic designation in a multiplex STR system. Int J Legal Med 111:267–272

141. Güçlü GB, Gürkök T, Koyuncu M, Arslan N, Parmaksız İ (2014) Genetic Characterization of Turkish Commercial Opium Poppy (Papaver somniferum L.) Cultivars Using ISSR and SSR Markers. J New Results Sci 7

142. Guo F, Yu J, Zhang L, Li J (2017) Massively parallel sequencing of forensic STRs and SNPs using the Illumina® ForenSeq™ DNA signature prep kit on the MiSeq FGx™ forensic genomics system. Forensic Sci Int Genet 31: 135-48

143. Guo L, Winzer T, Yang X et al (2018) The opium poppy genome and morphinan production. Science (New York, NY). doi: 10.1126/science.aat4096

144. Gustavsson E, Andersson M, Stephanson N, Beck O (2007) Validation of direct injection electrospray LC-MS/MS for confirmation of opiates in urine drug testing. J Mass Spectrom 42: 881-9

145. Hagelberg E, Clegg JB (1991) Isolation and characterization of DNA from archaeological bone. Proceedings of the Royal Society of London Series B: Biol Sci 244: 45-50

146. Haider N (2011) Chloroplast-specific universal primers and their uses in plant studies. Biol Plant 55:225–236

147. Hakki EE, Kayis SA, Pinarkara E, Sag A (2007) Inter simple sequence repeats separate efficiently hemp from marijuana (Cannabis sativa L.). Electron J Biotechn 10: 570-81

148. Hall DW (1988) The contributions of the forensic botanist to crime scene investigations. The Prosecutor 22: 35-8

149. Hallerman EM (2003) Population genetics: principles and applications for fisheries scientists

150. Hamilton M (1999) Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. Mol Ecol 8: 521–523

151. Hamilton MB (1999) Tropical tree gene flow and seed dispersal. Nature 401:129–130

152. Hammer Ø, Harper DA, Ryan PD (2001) PAST: paleontological statistics software package for education and data analysis. Palaeontol Electron 4:9

153. Handal KA, Schauben JL, Salamone FR (1983) Naloxone. Ann Emerg Med 12: 438-45

154. He X, Li F, Shi J, Gan S (2011) Short Note: Seven Genomic SSRs Revealed in Eucalyptus by Re-sequencing of DNA Sequences from GenBank. Silvae Genetica 60: 92-4

155. He X, Wang Y, Li F et al (2012) Development of 198 novel EST-derived microsatellites in Eucalyptus (Myrtaceae). Am J Bot 99: e134-e48

156. Hebert PD, Ratnasingham S, De Waard JR (2003) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proceedings of the Royal Society of London Series B: Biological Sciences 270: S96-S9

157. Heinze B (2007) A database of PCR primers for the chloroplast genomes of higher

plants. Plant Methods 3:4

158. Hendry E, Worthington T, Conway BR, Lambert P (2009) Antimicrobial efficacy of eucalyptus oil and 1, 8-cineole alone and in combination with chlorhexidine digluconate against microorganisms grown in planktonic and biofilm cultures. J Antimicrob Chemother 64: 1219-25
159. Henry RJ (2001) Plant DNA extraction. Plant Genotyping: the DNA fingerprinting of plants: 239-49
160. Herkenham M (1995) Localization of cannabinoid receptors in brain and periphery. Cannabinoid receptors
161. Hill KP (2015) Medical marijuana for treatment of chronic pain and other medical and psychiatric problems: a clinical review. Jama 313: 2474-83
162. Holleley CE, Geerts PG (2009) Multiplex Manager 1.0: a cross-platform computer program that plans and optimizes multiplex PCR. BioTechniques 46: 511-7
163. Hollingsworth PM, Graham SW, Little DP (2011) Choosing and using a plant DNA barcode. PLOS One 6: e19254
164. Hollingsworth PM, Li D-Z, van der Bank M, Twyford AD (2016) Telling plant species apart with DNA: from barcodes to genomes. Phil Trans R Soc B 371: 20150338
165. Hong C-Z, Shellock FG (1991) Effects of a topically applied counterirritant (Eucalyptamint) on cutaneous blood flow and on skin and muscle temperatures. A placebo-controlled study. Am J Phys Med Rehabil 70: 29-33
166. Horrocks M, Walsh KA (1998) Forensic palynology: assessing the value of the evidence. Rev Palaeobot Palyno 103: 69-74

167. Houston R, Birck M, Hughes-Stamm S, Gangitano D (2016) Evaluation of a 13-loci STR multiplex system for Cannabis sativa genetic identification. Int J Legal Med 130:635–647

168. Houston R, Birck M, Hughes-Stamm S, Gangitano D (2017) Developmental and internal validation of a novel 13 loci STR multiplex method for Cannabis sativa DNA profiling. Legal Med 26: 33–40

169. Houston R, Birck M, LaRue B, Hughes-Stamm S, Gangitano D (2018) Nuclear, chloroplast, and mitochondrial data of a US cannabis DNA database. Int J Legal Med 1–13

170. Houston R, Mayes C, King JL, Hughes-Stamm S, Gangitano D (2018) Massively parallel sequencing of 12 autosomal STRs in Cannabis sativa. Electrophoresis 39: 2906-11

171. Howard C, Gilmore S, Robertson J, Peakall R (2008) Developmental validation of a Cannabis sativa STR multiplex system for forensic analysis. J Forensic Sci 53: 1061-7

172. Hsieh H-M, Hou R-J, Tsai L-C, Wei CS, Liu SW, Huang LH, Kuo YC, Linacre A, Lee JCI (2003) A highly polymorphic STR locus in Cannabis sativa. Forensic Sci Int 131:53–58

173. Hu C, Zheng X, Teng Y (2011) Characterization and phylogenetic utility of non-coding chloroplast regions trnL-trnF and accD-psaI in Pyrus. Acta Horticulturae Sinica 38:2261–2272

174. Hu Z-G, Guo H-Y, Hu X-L et al (2012) Genetic diversity research of hemp (Cannabis sativa L) cultivar based on AFLP analysis. J Plant Genet Resour 13: 555-61

175. Hubbard M, Kelly J, Rajapakse S, Abbott A, Ballard R (1992) Restriction fragment length polymorphisms in rose and their use for cultivar identification. HortScience 27: 172-3

176. Hudson CJ, Freeman JS, Kullan AR et al (2012) A reference linkage map for Eucalyptus. BMC Genomics 13: 240

177. Hudson CJ, Kullan AR, Freeman JS et al (2012) High synteny and colinearity among Eucalyptus genomes revealed by high-density comparative genetic mapping. Tree Genet Genom 8: 339-52

178. Hughes R, Warner V (1978) A study of false positives in the chemical identification of marihuana. J Forensic Sci 23: 304-10

179. Ikawati Z, Wahyuono S, Maeyama K (2001) Screening of several Indonesian medicinal plants for their inhibitory effect on histamine release from RBL-2H3 cells. J Ethnopharmacol 75: 249-56

180. Illumina (2018) Nextera XT DNA Library Prep Kit Reference Guide. San Diego, CA

181. Illumina (2019) MiSeq System Denature and Dilute Libraries Guide. San Diego, CA

182. Ingman M, Gyllensten U (2006) mtDB: Human Mitochondrial Genome Database, a resource for population genetics and medical sciences. Nucleic Acids Res 34: D749-D51

183. Ishfaq M, Saleem M (2016) Application of molecular markers in forensic botany. Bull Biol Allied Sci Res 1: 1-7

184. Jacobs AD, Steiner RR (2014) Detection of the Duquenois–Levine chromophore in a marijuana sample. Forensic Sci Int 239: 1-5

185. Jagadish V, Robertson J, Gibbs A (1996) RAPD analysis distinguishes Cannabis sativa samples from different sources. Forensic Science Int 79: 113-21

186. Jeffreys AJ, Wilson V, Thein SL (1985) Hypervariable 'minisatellite'regions in human DNA. Nature 314: 67-73

187. Johnson LA, Soltis DE (1994) matK DNA sequences and phylogenetic reconstruction in Saxifragaceae s. str. Syst Bot 143-56

188. Jones ME, Stokoe RL, Cross MJ, Scott LJ, Maguire TL, Shepherd M (2001) Isolation of microsatellite loci from spotted gum (Corymbia variegata), and cross-species amplification in Corymbia and Eucalyptus. Molecular ecology notes 1: 276-8

189. Juergens UR, Stöber M, Vetter H (1998) Inhibition of cytokine production and arachidonic acid metabolism by eucalyptol (1.8-cineole) in human blood monocytes in vitro. Eur J Med Res 3: 508-10

190. Kalia RK, Rai MK, Kalia S, Singh R, Dhawan A (2011) Microsatellite markers: an overview of the recent progress in plants. Euphytica 177: 309-34

191. Kang HW, Cho YG, Yoon UH, Eun MY (1998) A rapid DNA extraction method for RFLP and PCR analysis from a single dry seed. Plant Mol Biol Rep 16: 90

192. Kapadia GJ, Fayez M (1970) Peyote constituents: chemistry, biogenesis, and biological effects. J Pharm Sci 59: 1699-727

193. Katevatis C, Fan A, Klapperich CM (2017) Low concentration DNA extraction and recovery using a silica solid phase. PloS One 12: e0176848

194. Kati V, Le Corre V, Michel S, Jaffrelo L, Poncet C, Délye C (2012) Isolation and Characterisation of 11 Polymorphic Microsatellite Markers in Papaver rhoeas L.

(Corn Poppy), a Major Annual Plant Species from Cultivated Areas. Int J Mol Sci 14(1):470-9

195. Kayis SA, Hakki EE, Pinarkara E (2010) Comparison of effectiveness of ISSR and RAPD markers in genetic characterization of seized marijuana (Cannabis sativa L.) in Turkey. Afr J Agric Res 5:2925–2933
196. Kim EH, Lee HY, Yang IS, Jung S-E, Yang WI, Shin K-J (2016) Massively parallel sequencing of 17 commonly used forensic autosomal STRs and amelogenin with small amplicons. Forensic Sci Int Genet 22: 1-7
197. Koch GT (2019) Best Practices on Tools for Combating and Preventing Illegal Logging Activities and Associated Trade. Asia-Pacific Economic Cooperation
198. Kohjyouma M, Lee I-J, Iida O et al (2000) Intraspecific variation in Cannabis sativa L. based on intergenic spacer region of chloroplast DNA. Biol Pharm Bull 23:727–730
199. Köhnemann S, Nedele J, Schwotzer D, Morzfeld J, Pfeiffer H (2012) The validation of a 15 STR multiplex PCR for Cannabis species. Int J Legal Med 126:601–606
200. Kojoma M, Iida O, Makino Y, Sekita S, Satake M (2002) DNA fingerprinting of Cannabis sativa using inter-simple sequence repeat (ISSR) amplification. Planta Med 68:60–63
201. Kojoma M, Seki H, Yoshida S, Muranaka T (2006) DNA polymorphisms in the tetrahydrocannabinolic acid (THCA) synthase gene in “drug-type” and “fiber-type” Cannabis sativa L. Forensic Sci Int 159: 132-40
202. Koressaar T, Remm M (2007) Enhancements and modifications of primer design program Primer3. Bioinformatics 23: 1289-91

203. Labanca F, Ovesnà J, Milella L (2018) Papaver somniferum L. taxonomy, uses and new insight in poppy alkaloid pathways. Phytochem Rev 17: 853-71. doi: 10.1007/s11101-018-9563-3

204. Lachenmeier DW, Sproll C, Musshoff F (2010) Poppy Seed Foods and Opiate Drug Testing-Where Are We Today? Ther Drug Monit 32: 11-8

205. Ladiges PY, Udovicic F, Nelson G (2003) Australian biogeographical connections and the phylogeny of large genera in the plant family Myrtaceae. J Biogeogr 30: 989-98

206. Lanfranco D, Dungey H (2001) Insect damage in Eucalyptus: a review of plantations in Chile. Austral Ecol 26: 477-81

207. Lao O, Lu TT, Nothnagel M et al (2008) Correlation between genetic and geographic structure in Europe. Curr Biol 18: 1241-8

208. Lara A, Little C, Urrutia R et al (2009) Assessment of ecosystem services as an opportunity for the conservation and management of native forests in Chile. For Ecol Manag 258: 415-24

209. Lavania U, Srivastava S (1999) Quantitative delineation of karyotype variation in Papaver as a measure of phylogenetic differentiation and origin. Curr Sci: 429-35

210. Laverty KU, Stout JM, Sullivan MJ et al (2019) A physical and genetic map of Cannabis sativa identifies extensive rearrangements at the THC/CBD acid synthase loci. Genom Res 29: 146-56

211. Lazaruk K, Walsh PS, Oaks F et al (1998) Genotyping of forensic short tandem repeat (STR) systems based on sizing precision in a capillary electrophoresis instrument. Electrophoresis 19: 86-93

212. Lee EJ, Hwang IK, Kim NY et al (2010) An assessment of the utility of universal and specific genetic markers for opium poppy identification. J Forensic Sci 55: 1202-8

213. Lee EJ, Jin GN, Lee KL, Han MS, Lee YH, Yang MS (2011) Exploiting expressed sequence tag databases for the development and characterization of gene-derived simple sequence repeat markers in the opium poppy (Papaver somniferum L.) for forensic applications. J Forensic Sci 56: 1131-5

214. Lewis P (2001) Genetic data analysis: computer program for the analysis of allelic data. Version 1.0 (d16c). http://lewiseebuconnedu/lewishome/softwarehtml

215. Li H-L (1974) The origin and use of Cannabis in eastern Asia linguistic-cultural implications. Econ Bot 28: 293-301

216. Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S (2015) Plant DNA barcoding: from gene to genome. Biol Rev 90: 157-66

217. Linacre A, Gusmao L, Hecht W et al (2011) ISFG: recommendations regarding the use of non-human (animal) DNA in forensic genetic investigations. Forensic Sci Int Genet 5: 501-5

218. Linacre A, Gusmao L, Hecht W et al (2011) ISFG: recommendations regarding the use of non-human (animal) DNA in forensic genetic investigations. Forensic Sci Int Genet 5: 501-5. doi: 10.1016/j.fsigen.2010.10.017

219. Linacre A, Thorpe J (1998) Detection and identification of cannabis by DNA. Forensic Sci Int 91:71–76

220. Lloyd KG, MacGregor BJ, Teske A (2010) Quantitative PCR methods for RNA and DNA in marine sediments: maximizing yield while overcoming inhibition. FEMS Microbiol Ecol 72: 143-51

221. Lloyd-Jones D, Bonomo Y (2006) Unusual presentations for pharmacotherapy—poppy seed dependence. Drug Alcohol Rev 25: 375-6

222. Loman NJ, Misra RV, Dallman TJ et al (2012) Performance comparison of benchtop high-throughput sequencing platforms. Nat Biotechnol 30: 434

223. López P, Pereboom-de Fauw DP, Mulder PP et al (2018) Straightforward analytical method to determine opium alkaloids in poppy seeds and bakery products. Food Chem 242: 443-50

224. Lowe AJ, Wong K-N, Tiong Y-S, Iyerh S, Chew F-T (2010) A DNA method to verify the integrity of timber supply chains; confirming the legal sourcing of merbau timber from logging concession to sawmill. Silvae Genet 59: 263-8

225. Luo J, Arnold R, Ren S et al (2013) Veneer grades, recoveries, and values from 5-year-old eucalypt clones. Ann For Sci 70: 417-28

226. Lurie IS, Sam SC, Sara ED, Sini P (2013) Determination of heroin and basic impurities for drug profiling by ultra-high-pressure liquid chromatography. Forensic Sci Int 231: 300-5. doi: http://dx.doi.org/10.1016/j.forsciint.2013.06.008

227. Marciano MA, Panicker SX, Liddil GD, Lindgren D, Sweder KS. (2018) Development of a Method to Extract Opium Poppy (Papaver somniferum L.) DNA from Heroin. Sci Rep 8(1): 1-15

228. Mars SG, Bourgois P, Karandinos G, Montero F, Ciccarone D (2016) The textures of heroin: User perspectives on “black tar” and powder heroin in two US cities. J Psychoact Drugs 48: 270-8

229. Masárová V, Mihálik D, Kraic J (2015) In Silico Retrieving of Opium Poppy (Papaver Somniferum L.) Microsatellites. Agriculture (Polnohospodárstvo) 61: 149-56

230. Matielo CBDO, Lemos RPM, Sarzi DS et al (2020) Whole Plastome Sequences of Two Drug-Type Cannabis: Insights Into the Use of Plastid in Forensic Analyses. J Forensic Sci 65: 259-65

231. Mayr W (1995) DNA Recommendations-1994 report concerning further recommendations of the DNA commission of the ISFH regarding PCR-based polymorphisms in STR (short tandem repeat) systems. Vox Sang 69:70–71

232. McElhoe JA, Holland MM, Makova KD et al (2014) Development and assessment of an optimized next-generation DNA sequencing approach for the mtgenome using the Illumina MiSeq. Forensic Sci Int Genet 13: 20-9

233. McKernan KJ, Helbert Y, Tadigotla V et al (2015) Single molecule sequencing of THCA synthase reveals copy number variation in modern drug-type Cannabis sativa L. BioRxiv: 028654

234. Mehmedic Z, Chandra S, Slade D et al (2010) Potency trends of Δ9-THC and other cannabinoids in confiscated cannabis preparations from 1993 to 2008. J Forensic Sci 55: 1209-17

235. Mehta B, Daniel R, Phillips C, Doyle S, Elvidge G, McNevin D (2016) Massively parallel sequencing of customised forensically informative SNP panels on the MiSeq. Electrophoresis 37: 2832-40

236. Mello I, Ribeiro A, Dias V et al (2016) A segment of rbcL gene as a potential tool for forensic discrimination of Cannabis sativa seized at Rio de Janeiro, Brazil. Int J Legal Med 130: 353-6

237. Mendoza M (2003) Losing ground to timber thieves: Illegal logging chips away at for ests, but one court puts foot down. Associated Press, May

238. Mendoza MA, Mills DK, Lata H, Chandra S, ElSohly MA, Almirall JR (2009) Genetic individualization of Cannabis sativa by a short tandem repeat multiplex system. Anal Bioanal Chem 393: 719-26

239. Mestel R (1993) Murder trial features tree's genetic fingerprint. New Scientist 138: 6.

240. Mičianová V, Ondreičková K, Muchová D et al (2017) Forensic application of EST-derived STR markers in opium poppy. Biologia 72: 587-94

241. Mikes F, Hofmann A, Waser PG (1971) Identification of (−)-Δ9-6a, 10a-trans-tetrahydro-cannabinol and two of its metabolites in rats by use of combination gas chromatography-mass spectrometry and mass fragmentography. Biochem Pharmacol 20:2469–2476

242. Milhau G, Valentin A, Benoit F et al (1997) In vitro antimalarial activity of eight essential oils. J Essent Oil Res 9: 329-33

243. Miller HC, Shutler G, Abrams S et al (2003) A simple DNA extraction method for marijuana samples used in amplified fragment length polymorphism (AFLP) analysis. J Forensic Sci 48: 343-7

244. Miller KW, Budowle B (2001) A compendium of human mitochondrial DNA control region: development of an international standard forensic database. Croat Med J 42: 315-27

245. Mitosinka GT, Thornton JI, Hayes TL (1972) The examination of cystolithic hairs of Cannabis and other plants by means of the scanning electron microscope. J Forensic Sci Soc 12:521–529

246. Moliterni VC, Cattivelli L, Ranalli P, Mandolino G (2004) The sexual differentiation of Cannabis sativa L.: a morphological and molecular study. Euphytica 140: 95-106

247. Moore D, Dowhan D (2002) Purification and concentration of DNA from aqueous solutions. Curr Protoc Mol Biol 59: 2.1. -2.1. 10

248. Moorthie S, Mattocks CJ, Wright CF (2011) Review of massively parallel DNA sequencing technologies. HUGO J 5: 1-12

249. Morello DR, Cooper SD, Panicker S, Casale JF. (2010) Signature Profiling and Classification of Illicit Heroin by GC-MS Analysis of Acidic and Neutral Manufacturing Impurities. Blackwell Publishing Ltd. United States. pp. 42

250. Morello DR, Meyers RP. (1995) Qualitative and Quantitative Determination of Residual Solvents in Illicit Cocaine HCI and Heroin HCI. American Society for Testing and Materials. United States. pp. 957

251. Mueller UG, Wolfenbarger LL (1999) AFLP genotyping and fingerprinting. Trends Ecol Evol 14: 389-94

252. Munro S, Thomas KL, Abu-Shaar M (1993) Molecular characterization of a peripheral receptor for cannabinoids. Nature 365: 61

253. Myburg AA, Grattapaglia D, Tuskan GA et al (2014) The genome of Eucalyptus grandis. Nature 510: 356-62

254. Myburg AA, Potts BM, Marques CM et al. (2007) Eucalypts. Forest Trees. Springer. pp. 115-60

255. Nanjayya SB, Murthy P, Chand PK et al (2010) A case of poppy tea dependence in an octogenarian lady. Drug Alcohol Rev 29: 216-8

256. National Institute on Drug Abuse (2019) NIDA's role in providing marijuana for research

257. Neves LG, Mamani EM, Alfenas AC, Kirst M, Grattapaglia D (2011) A high-density transcript linkage map with 1,845 expressed genes positioned by microarray-based Single Feature Polymorphisms (SFP) in Eucalyptus. Bmc Genomics 12: 189

258. Nevill PG, Bradbury D, Jørgensen T, Krauss S, Samaraweera S, Gardner MG (2013) Microsatellite primers identified by 454 sequencing in the floodplain tree species Eucalyptus victrix (Myrtaceae). Appl Plant Sci 1: 1200402

259. Nexttec. DNA Isolation from Plant by nexttec 1 -Step. 3.0 ed Hilgertshausen

260. Novembre J, Johnson T, Bryc K et al (2008) Genes mirror geography within Europe. Nature 456: 98-101

261. Nybom H, Weising K, Rotter B (2014) DNA fingerprinting in botany: past, present, future. Investig Genet 5: 1

262. Ogden R, Linacre A (2015) Wildlife forensic science: a review of genetic geographic origin assignment. Forensic Science International: Genetics 18: 152-9

263. Oh H, Seo B, Lee S et al (2016) Two complete chloroplast genome sequences of Cannabis sativa varieties. Mitochondrial DNA Part A 27: 2835-7

264. Olson ME (2002) Intergeneric relationships within the Caricaceae-Moringaceae clade (Brassicales) and potential morphological synapomorphies of the clade and its families. Int J Plant Sci 163:51–65

265. Onofri C, de Meijer EP, Mandolino G (2015) Sequence heterogeneity of cannabidiolic-and tetrahydrocannabinolic acid-synthase in Cannabis sativa L. and its relationship with chemical phenotype. Phytochem 116: 57-68

266. Osawa T, Ide A, Su JD, Namiki M (1987) Inhibition of in vitro lipid peroxidation by ellagic acid. J Agric Food Chem 35: 808-12

267. O'Shaughnessy WB (1843) On the preparations of the Indian hemp, or Gunjah: Cannabis indica their effects on the animal system in health, and their utility in the treatment of tetanus and other convulsive diseases. Provincial Medical Journal and Retrospect of the Medical Sciences 5: 363

268. Otsuka H, Tsukui M, Matsuoka T, Goto M, Fujimura H (1974) Studies on anti-inflammatory agents. Anti-inflammatory screening by fertile egg method (author's transl). Yakugaku zasshi: J Pharm Soc Jpn 94: 796

269. Ottewell K, Donnellan S, Moran G, Paton D (2005) Multiplexed microsatellite markers for the genetic analysis of Eucalyptus leucoxylon (Myrtaceae) and their utility for ecological and breeding studies in other Eucalyptus species. J Hered 96: 445-51

270. Oudjehih B, Abdellah B (2006) Chromosome numbers of the 59 species of Eucalyptus L'Herit.(Myrtaceae). Caryologia 59: 207-12

271. Oxelman B, Lidén M, Berglund D (1997) Chloroplastrps16 intron phylogeny of the tribeSileneae (Caryophyllaceae). Plant Syst Evol 206:393–410

272. Pacifico D, Miselli F, Carboni A, Moschella A, Mandolino G (2008) Time course of cannabinoid accumulation and chemotype development during the growth of Cannabis sativa L. Euphytica 160: 231-40

273. Pacifico D, Miselli F, Micheler M, Carboni A, Ranalli P, Mandolino G (2006) Genetics and marker-assisted selection of the chemotype in Cannabis sativa L. Mol Breed 17: 257-68

274. Pal L, Lucia LA (2019) Renaissance of industrial hemp: A miracle crop for a multitude of products. BioResources 14: 2460-4

275. Panchy N, Lehti-Shiu M, Shiu S-H (2016) Evolution of gene duplication in plants. Plant Physiol 171: 2294-316

276. Parks M, Cronn R, Liston A (2009) Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. Bmc Biol 7: 84

277. Parson W, Ballard D, Budowle B et al (2016) Massively parallel sequencing of forensic STRs: considerations of the DNA commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements. Forensic Sci Int Genet 22: 54-63

278. Parthiban S, Govindaraj P, Senthilkumar S (2018) Comparison of relative efficiency of genomic SSR and EST-SSR markers in estimating genetic diversity in sugarcane. 3 Biotech 8: 144

279. Payn KG, Dvorak WS, Janse BJ, Myburg AA (2008) Microsatellite diversity and genetic structure of the commercially important tropical tree species Eucalyptus urophylla, endemic to seven islands in eastern Indonesia. Tree Genet Genom 4: 519-30

280. Peil A, Flachowsky H, Schumann E, Weber W (2003) Sex-linked AFLP markers indicate a pseudoautosomal region in hemp (Cannabis sativa L.). Theor Appl Genet 107: 102-9

281. Pert CB, Pasternak G, Snyder SH (1973) Opiate agonists and antagonists discriminated by receptor binding in brain. Science 182: 1359-61

282. Pert CB, Snyder SH (1973) Opiate receptor: demonstration in nervous tissue. Science 179: 1011-4

283. Petroli CD, Sansaloni CP, Carling J et al (2012) Genomic characterization of DArT markers based on high-density linkage analysis and physical mapping to the Eucalyptus genome. PLoS One 7: e44684

284. Pinarkara E, Kayis SA, Hakki EE, Sag A (2009) RAPD analysis of seized marijuana (Cannabis sativa L.) in Turkey. Electron J Biotechn 12: 5-6

285. Porebski S, Bailey LG, Baum BR (1997) Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Mol Biol Rep 15: 8-15

286. Potts BM, Vaillancourt RE, Jordan G et al (2004) Exploration of the Eucalyptus globulus gene pool

287. Powers D, Erickson S, Swortwood MJ (2018) Quantification of Morphine, Codeine, and Thebaine in Home-Brewed Poppy Seed Tea by LC-MS/MS. J Forensic Sci 63: 1229-35. doi: 10.1111/1556-4029.13664

288. Prandi C, Blangetti M, Namdar D, Koltai H (2018) Structure-Activity Relationship of Cannabis Derived Compounds for the Treatment of Neuronal Activity-Related Diseases. Molecules 23. doi: 10.3390/molecules23071526

289. QIAGEN (2018) DNeasy® Plant Handbook. Hilden, Germany

290. QIAGEN (2018) QIAquick® Spin Handbook. Hilden, Germany

291. QIAGEN (2009) Type-it® Microsatellite PCR Handbook. Hilden, Germany

292. Quail MA, Smith M, Coupland P et al (2012) A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC Genom 13: 341

293. Retrieved February 29, 2020 from the Integrated Taxonomic Information System on-line database, http://www.itis.gov

294. Richins RD, Rodriguez-Uribe L, Lowe K, Ferral R, O'Connell MA (2018) Accumulation of bioactive metabolites in cultivated medical Cannabis. PLoS One 13

295. Robinson JT, Thorvaldsdóttir H, Winckler W et al (2011) Integrative genomics viewer. Nature Biotechnol 29: 24-6

296. Röhl A, Brinkmann B, Forster L, Forster P (2001) An annotated mtDNA database. Int J Legal Med 115: 29-39

297. Roman MG, Gangitano D, Houston R (2019) Characterization of new chloroplast markers to determine biogeographical origin and crop type of Cannabis sativa. Int J Legal Med 133: 1721-32

298. Rothberg JM, Hinz W, Rearick TM et al (2011) An integrated semiconductor device enabling non-optical genome sequencing. Nature 475: 348-52

299. Rotherham D, Harbison S (2011) Differentiation of drug and non-drug Cannabis using a single nucleotide polymorphism (SNP) assay. Forensic Sci Int 207: 193-7

300. Royal CD, Novembre J, Fullerton SM et al (2010) Inferring genetic ancestry: opportunities, challenges, and implications. Am J Hum Genet 86: 661-73

301. Rudd RA, Aleshire N, Zibbell JE, Matthew Gladden R (2016) Increases in drug and opioid overdose deaths—United States, 2000–2014. Am J Transplant 16: 1323-7

302. Rudd RA, Seth P, David F, Scholl L (2016) Increases in Drug and Opioid-Involved Overdose Deaths - United States, 2010-2015. MMWR Morbidity And Mortality Weekly Report 65: 1445-52. doi: 10.15585/mmwr.mm655051e1

303. Sakai M, Kanazawa A, Fujii A, Thseng F, Abe J, Shimamoto Y (2003) Phylogenetic

relationships of the chloroplast genomes in the genus Glycine inferred from four intergenic spacer sequences. Plant Syst Evol 239:29–54

304. Sakamoto K, Akiyama Y, Fukui K, Kamada H, Satoh S (1998) Characterization; genome sizes and morphology of sex chromosomes in hemp (Cannabis sativa L.). Cytologia 63: 459-64
305. Sanchez JJ, Børsting C, Hallenberg C, Buchard A, Hernandez A, Morling N (2003) Multiplex PCR and minisequencing of SNPs—a model with 35 Y chromosome SNPs. Forensic Sci Int 137:74–84
306. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci 74: 5463-7
307. Sansaloni CP, Petroli CD, Carling J et al (2010) A high-density Diversity Arrays Technology (DArT) microarray for genome-wide genotyping in Eucalyptus. Plant Methods 6: 16
308. Santos C, Pereira F (2017) Design and evaluation of PCR primers for amplification of four chloroplast DNA regions in plants. Conserv Genet Resour 9:9–12
309. Sarkar S (1994) Capillary permeability-increasing effect of Eucalyptus hybrid leaf and a Seseli indicum seed oils in rabbit. Indian J Pharmacol 26: 55
310. Saunders JA, Pedroni MJ, Penrose LD, Fist AJ (2001) AFLP analysis of opium poppy. Crop Sci 41: 1596-601
311. Schaal B, Hayworth D, Olsen KM, Rauscher J, Smith W (1998) Phylogeographic studies in plants: problems and prospects. Mol Ecol 7:465–474
312. Schmitz R (1985) Friedrich Wilhelm Sertürner and the discovery of morphine. Pharm Hist 27: 61-74

313. Schnitzler P, Schön K, Reichling J (2001) Antiviral activity of Australian tea tree oil and eucalyptus oil against herpes simplex virus in cell culture. Die Pharmazie 56: 343-7

314. Scott GN. (2017) Poppy Seed Tea: Beware This Beverage. Medscape

315. Şelale H, Çelik İ, Gültekin V, Allmer J, Doğanlar S, Frary A (2013) Development of EST-SSR markers for diversity and breeding studies in opium poppy. Plant Breed 132: 344-51

316. Seo SB, King JL, Warshauer DH, Davis CP, Ge J, Budowle B (2013) Single nucleotide polymorphism typing with massively parallel sequencing for human identification. Int J Legal Med 127: 1079-86

317. Sexton TR, Henry RJ, McManus LJ, Bowen S, Shepherd M (2010) Capture of assay template by multiplex PCR of long amplicons for genotyping SNPs and InDels with MALDI-TOF mass spectrometry. Molecular Breed 25: 471-80

318. Seyani C, Green P, Daniel L, Pegden A (2017) An interesting case of opium tea toxicity. BMJ Case Reports 2017: bcr-2016-218971

319. Shajitha P, Dhanesh N, Ebin P et al (2016) A combined chloroplast atpB–rbcL and trnL-F phylogeny unveils the ancestry of balsams (Impatiens spp.) in the Western Ghats of India. 3 Biotech 6:258

320. Shaw J, Lickey EB, Beck JT et al (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am J Bot 92: 142-66

321. Shaw J, Lickey EB, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in

angiosperms: the tortoise and the hare III. Am J Bot 94:275–288

322. Shaw J, Shafer HL, Leonard OR, Kovach MJ, Schorr M, Morris AB (2014) Chloroplast DNA sequence utility for the lowest phylogenetic and phylogeographic inferences in angiosperms: the tortoise and the hare IV. Am J Bot 101:1987–2004
323. Shepherd M, Kasem S, Lee D, Henry R (2006) Construction of microsatellite linkage maps for Corymbia. Silvae Genet 55: 228-38
324. Shirley N, Allgeier L, LaNier T, Coyle HM (2013) Analysis of the NMI01 marker for a population database of cannabis seeds. J Forensic Sci 58
325. Shirota O, Watanabe A, Yamazaki M et al (1998) Random amplified polymorphic DNA and restriction fragment length polymorphism analyses of Cannabis sativa. Natural medicines= 生薬學雜誌 52: 160-6
326. Shokralla S, Porter TM, Gibson JF et al (2015) Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform. Sci Rep 5: 9687
327. Sirikantaramas S, Morimoto S, Shoyama Y et al (2004) The gene controlling marijuana psychoactivity molecular cloning and heterologous expression of Δ1-tetrahydrocannabinolic acid synthase from Cannabis sativa L. J Biol Chem 279: 39767-74
328. Siver PA, Lord WD, McCarthy DJ (1994) Forensic limnology: the use of freshwater algal community ecology to link suspects to an aquatic crime scene in southern New England. J Forensic Sci 39: 847-53
329. Small RL, Ryburn JA, Cronn RC, Seelanan T, Wendel JF (1998) The tortoise and the hare: choosing between noncoding plastome and nuclear Adh sequences for

phylogeny reconstruction in a recently diverged plant group. Am J Bot 85:1301–1315

330. Smith C (2019) What Is the Opioid Epidemic? AddictionCenter
331. Smith ML, Hughes RO, Levine B, Dickerson S, Darwin WD, Cone EJ (1995) Forensic drug testing for opiates. VI. Urine testing for hydromorphone, hydrocodone, oxymorphone, and oxycodone with commercial opiate immunoassays and gas chromatography-mass spectrometry. J Anal Toxicol 19: 18-26
332. Smith SD, Cowan RS, Gregg KB, Chase MW, Maxted N, Fay MF (2004) Genetic discontinuities among populations of Cleistes (Orchidaceae, Vanilloideae) in North America. Bot J Linn Soc 145:87–95
333. Snyder B (2014) Revisiting old friends: update on opioid pharmacology. Issues: 1.
334. Soler S, Gramazio P, Figàs MR et al (2017) Genetic structure of Cannabis sativa var. indica cultivars based on genomic SSR (gSSR) markers: implications for breeding and germplasm management. Ind Crops Prod 104: 171-8
335. Staginnus C, Zörntlein S, de Meijer E (2014) A PCR marker Linked to a THCA synthase Polymorphism is a Reliable Tool to Discriminate Potentially THC-Rich Plants of C annabis sativa L. J Forensic Sci 59: 919-26
336. Steane D, Jones R, Vaillancourt R (2005) A set of chloroplast microsatellite primers for Eucalyptus (Myrtaceae). Mol Ecol Notes 5: 538-41
337. Steane D, Vaillancourt R, Russell J, Powell W, Marshall D, Potts B (2001) Development and characterisation of microsatellite loci in Eucalyptus globulus (Myrtaceae). Silvae Genet 50: 89-91

338. Steane D, Vaillancourt R, Russell J, Powell W, Marshall D, Potts B (2001) Development and characterisation of microsatellite loci in Eucalyptus globulus (Myrtaceae). Silvae Genet 50: 89-91

339. Steele KP, Vilgalys R (1994) Phylogenetic analyses of Polemoniaceae using nucleotide sequences of the plastid gene matK. Syst Bot 19:126–142

340. Stefanidou M, Athanaselis S, Spiliopoulou C, Dona A, Maravelias C (2010) Biomarkers of opiate use. Int J Clin Pract 64: 1712-8

341. Stewart CP, Stolman A. (2013) Toxicology: mechanisms and analytical methods. Elsevier

342. Stuyt E (2018) The problem with the current high potency THC marijuana from the perspective of an addiction psychiatrist. Missouri medicine 115: 482

343. Subashini V, Shanmugapriya A, Yasodha R (2014) Hybrid purity assessment in Eucalyptus F 1 hybrids using microsatellite markers. 3 Biotech 4: 367-73

344. Substance Abuse and Mental Health Services Administration (2018) Key substance use and mental health indicators in the United States: Results from the 2017 National Survey on Drug Use and Health (HHS Publication No. SMA 18-5068, NSDUH Series H-53). Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration Rockville, MD

345. Substance Abuse and Mental Health Services Administration. (2019) Key substance use and mental health indicators in the United States: Results from the 2018 National Survey on Drug Use and Health (HHS Publication No. PEP19-5068, NSDUH Series H-54). Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration Rockville, MD

346. Sucher NJ, Hennell JR, Carles MC. (2012) DNA fingerprinting, DNA barcoding, and next generation sequencing technology in plants. Plant DNA Fingerprinting and Barcoding. Springer. pp. 13-22

347. Sugimoto K, Suzuki J, Nakagawa K et al (2005) Eucalyptus leaf extract inhibits intestinal fructose absorption, and suppresses adiposity due to dietary sucrose in rats. Br J Nutr 93: 957-63

348. Sumathi M, Yasodha R (2014) Microsatellite resources of Eucalyptus: current status and future perspectives. Bot Stud 55: 73

349. SWGDAM. (2012) Validation guidelines for DNA analysis methods. Scientific working group on DNA analysis methods

350. Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol 17:1105–1109

351. Tacconi L. (2012) Illegal Logging: Law Enforcement, Livelihoods and the Timber Trade. Routledge

352. Takasaki M, Konoshima T, Fujitani K et al (1990) Inhibitors of Skin-Tumor Promotion. VIII.: Inhibitory Effects of Euglobals and Their Related Compounds on Epstein-Barr Virus Activation.(1). Chem Pharm Bull 38: 2737-9

353. Takayama K, OHI-TOMA T, Kudoh H, Kato H (2005) Origin and diversification of Hibiscus glaber, species endemic to the oceanic Bonin Islands, revealed by chloroplast DNA polymorphism. Mol Ecol 14:1059–1071

354. Tanda G, Loddo P, Di Chiara G (1999) Dependence of mesolimbic dopamine transmission on Δ9-tetrahydrocannabinol. Eur J Pharmacol 376: 23-6

355. Tanksley S, Young N, Paterson A, Bonierbale M (1989) RFLP mapping in plant breeding: new tools for an old science. Biotechnol 7: 257-64

356. Taura F, Morimoto S, Shoyama Y, Mechoulam R (1995) First direct evidence for the mechanism of. DELTA. 1-tetrahydrocannabinolic acid biosynthesis. J Am Chem Soc 117: 9766-7

357. Tétényi P (1997) Opium poppy (Papaver somniferum): botany and horticulture. Hortic Rev 19: 373-408

358. Thamarus K, Groom K, Murrell J, Byrne M, Moran G (2002) A genetic linkage map for Eucalyptus globulus with candidate loci for wood, fibre, and floral traits. Theor Appl Genet 104: 379-87

359. Thermo Fisher Scientific (2010) ABI PRISM® SNaPshot™ Multiplex Kit. South San Francisco, CA

360. Thermo Fisher Scientific (2011) BigDye® Direct Cycle Sequencing Kit. South San Francisco, CA

361. Thermo Fisher Scientific (2015) Qubit® dsDNA HS Assay Kits Protocol

362. Thermo Fisher Scientific (2016) BigDye™ Terminator v3.1 Cycle Sequencing Kit User Guide. South San Francisco, CA

363. Thevis M, Opfermann G, Schänzer W (2003) Urinary concentrations of morphine and codeine after consumption of poppy seeds. J Anal Toxicol 27: 53-6

364. Thiel T, Michalek W, Varshney R, Graner A (2003) Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (Hordeum vulgare L.). Theor Appl Genet 106: 411-22

365. Thornton JI, Nakamura GR (1972) The identification of marijuana. J Forensic Sci Soc 12: 461-519

366. Timme RE, Kuehl JV, Boore JL, Jansen RK (2007) A comparative analysis of the Lactuca and Helianthus (Asteraceae) plastid genomes: identification of divergent regions and categorization of shared repeats. Am J Bot 94:302–312

367. Tnah L, Lee S, Ng K, Faridah Q, Faridah-Hanum I (2010) Highly variable STR markers of Neobalanocarpus heimii (Dipterocarpaceae) for forensic DNA profiling. J Tropic For Sci: 214-26

368. Toth JA, Stack GM, Cala AR et al (2020) Development and validation of genetic markers for sex and cannabinoid chemotype in Cannabis sativa L. GCB Bioenergy: 213-22

369. Trafkowski J, Madea B, Musshoff F (2006) The significance of putative urinary markers of illicit heroin use after consumption of poppy seed products. Ther Drug Monit 28: 552-8

370. Tsai L-C, Hsieh H-M, Wang J-C, Huang L-H, Linacre A, Lee JC-I (2006) Cannabis seed identification by chloroplast and nuclear DNA. Forensic Sci Int 158: 250-2

371. United Nations (1961) Single Convention on Narcotic Drugs, 1961

372. United States Department of Justice (2010) DrugAlert Watch: Opium Tea. National Drug Intelligence Center, U.S. Department of Justice Washington, D.C.

373. Vallone PM, Butler JM (2004) AutoDimer: a screening tool for primer-dimer and hairpin structures. Biotechniques 37: 226-31

374. Valverde L, Lischka C, Erlemann S et al (2014) Nomenclature proposal and SNPSTR haplotypes for 7 new Cannabis sativa L. STR loci. Forensic Sci Int Genet 13: 185-6

375. Van Bakel H, Stout JM, Cote AG et al (2011) The draft genome and transcriptome of Cannabis sativa. Genom Biol 12: R102

376. Van der Nest M, Steenkamp E, Wingfield B, Wingfield M (2000) Development of simple sequence repeat (SSR) markers in Eucalyptus from amplified inter-simple sequence repeats (ISSR). Plant Breed 119: 433-6

377. Vergara D, Baker H, Clancy K et al (2016) Genetic and genomic tools for Cannabis sativa. Crit Rev Plant Sci 35: 364-77

378. Vergara D, Huscher EL, Keepers KG et al (2019) Gene copy number is associated with phytochemistry in Cannabis sativa. AoB Plants 11: plz074

379. Vergara D, White KH, Keepers KG, Kane NC (2016) The complete chloroplast genomes of Cannabis sativa and Humulus lupulus. Mitochondrial DNA Part A 27: 3793-4

380. Vilela GR, de Almeida GS, D'Arce MABR et al (2009) Activity of essential oil and its major compound, 1, 8-cineole, from Eucalyptus globulus Labill., against the storage fungi Aspergillus flavus Link and Aspergillus parasiticus Speare. J Stored Prod Res 45: 108-11

381. Volpe DA, Tobin GAM, Mellon RD et al (2011) Uniform assessment and ranking of opioid Mu receptor binding constants for selected opioid drugs. Regul Toxicol Pharmacol 59: 385-90. doi: 10.1016/j.yrtph.2010.12.007

382. Vos P, Hogers R, Bleeker M et al (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23: 4407-14

383. Wang D-Y, Zhang Q, Liu Y, Lin Z-F, Zhang S-X, Sun M-X (2010) The levels of male gametic mitochondrial DNA are highly regulated in angiosperms with regard to mitochondrial inheritance. Plant Cell 22: 2402-16

384. Wang Z, Zhou D, Wang H et al (2017) Massively parallel sequencing of 32 forensic markers using the Precision ID GlobalFiler™ NGS STR Panel and the Ion PGM™ System. Forensic Sci Int Genet 31: 126-34

385. Watts G (2006) Cannabis confusions. Bmj 332: 175-6

386. Weiblen GD, Wenger JP, Craft KJ et al (2015) Gene duplication and divergence affecting drug content in Cannabis sativa. New Phytologist 208: 1241-50

387. Weising K, Gardner RC (1999) A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. Genome 42:9–19

388. Werf H, Mathussen E, Haverkort A (1996) The potential of hemp (Cannabis sativa L.) for sustainable fibre production: a crop physiological appraisal. Ann Appl Biol 129:109–123

389. White KH, Vergara D, Keepers KG, Kane NC (2016) The complete mitochondrial genome for Cannabis sativa. Mitochondrial DNA Part B 1: 715-6

390. Wiedenhoeft A (2016) Best practice guide for forensic timber identification. In: United Nations Office on Drugs and Crime: International Consortium on Combating Wildlife Crime Vienna, Austria: Laboratory and Scientific Section, Global Programme for Combating Wildlife and Forest Crime: 1-226: 1-226

391. World Wide Fund for Nature (2019) Illegal Logging

392. Xin Z, Chen J (2012) A high throughput DNA extraction method with high yield and quality. Plant Methods 8: 26

393. Xu D, Abe J, Sakai M, Kanazawa A, Shimamoto Y (2000) Sequence variation of non-coding regions of chloroplast DNA of soybean and related wild species and its implications for the evolution of different chloroplast haplotypes. Theor Appl Genet 101: 724–732

394. Yang J-B, Yang H-Q, Li D-Z, Wong K-M, Yang Y-M (2010) Phylogeny of Bambusa and its allies (Poaceae: Bambusoideae) inferred from nuclear GBSSI gene and plastid psbA-trnH, rpl32-trnL and rps16 intron DNA sequences. Taxon 59:1102–1110

395. Yang M-Q, van Velzen R, Bakker FT, Sattarian A, Li D-Z, Yi T-S (2013) Molecular phylogenetics and character evolution of Cannabaceae. Taxon 62: 473-85

396. Yasodha R, Sumathi R, Chezhian P, Kavitha S, Ghosh M (2008) Eucalyptus microsatellites mined in silico: survey and evaluation. J Genet 87: 21-5

397. Yoon CK (1993) Botanical witness for the prosecution. Science 260: 894-6

398. Young AM, Havens JR, Leukefeld CG (2010) Route of administration for illicit prescription opioids: a comparison of rural and urban drug users. Harm Reduct J 7: 24

399. Zhang Q, Liu Y (2003) Examination of the cytoplasmic DNA in male reproductive cells to determine the potential for cytoplasmic inheritance in 295 angiosperm species. Plant Cell Physiol 44:941– 951

400. Zhou C, He X, Li F et al (2014) Development of 240 novel EST-SSRs in Eucalyptus L'Hérit. Mol Breed 33: 221-5

401. Zhou C, Li F, Weng Q, Yu X, Li M, Gan S (2010) Comparison between direct sequencing and pool-cloning-based sequencing of PCR products in EST-SSR marker development in Eucalyptus. Mol Plant Breed 8: 1-10

402. Zhou J, Tang F, Mao G, Shao J, Wang Y, Bian R (2003) Effect of Eucalyptus globulus oil on activation of nuclear factor-kappaB in THP-1 cells. Zhejiang da xue xue bao Yi xue ban= Journal of Zhejiang University Medical Sciences 32: 315-8, 26

403. Ziegler J, Facchini PJ, Geißler R et al (2009) Evolution of morphine biosynthesis in opium poppy. Phytochem 70: 1696-707

404. Zuardi AW (2006) History of cannabis as a medicine: a review. Braz J Psychiatry 28: 153-7

405. ZyGEM. Plant DNA Extraction (for use with phytoGEM kits). Charlottesville, VA

APPENDIX

Certificates of Analysis for 15 Hemp Strains

CHEMHISTORY
CANNABIS TESTING LABORATORY

Certificate of Analysis

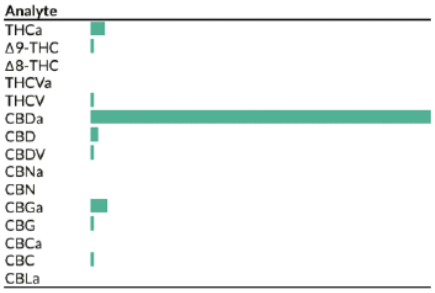
Industrial Hemp

Powered by Confident Cannabis
1 of 4

Sample: 1810CH0391.1720
Strain: Elektra
Batch#: ; Batch Size: - grams
Sample Received: 10/24/2018; Report Created: 11/07/2018; Expires: 11/07/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF - Elektra

Plant, Flower - Cured, Outdoor
Harvest Process Lot: ; METRC Batch: ; METRC Sample:



Cannabinoids

154 HPLC4 20181025-2
10/25/2018

Pass

0.06%	16.60%
Δ9-THC	Total CBD** (Calculated Decarboxylated Potential)
20.71%	**8.0%**
Total Cannabinoids Analyzed	Moisture

Analyte	LOQ	Mass	Mass
	mg/g	%	mg/g
THCa	0.2	0.69	6.9
Δ9-THC	0.2	0.06	0.6
Δ8-THC	0.2	<LOQ	<LOQ
THCVa	0.2	<LOQ	<LOQ
THCV	0.2	0.02	0.2
CBDa	0.2	18.44	184.4
CBD	0.2	0.43	4.3
CBDV	0.2	0.03	0.3
CBNa	0.2	<LOQ	<LOQ
CBN	0.2	<LOQ	<LOQ
CBGa	0.2	0.88	8.8
CBG	0.2	0.10	1.0
CBCa	0.2	<LOQ	<LOQ
CBC	0.2	0.07	0.7
CBLa	0.2	<LOQ	<LOQ
Total		**20.71**	**207.1**

*Total THC = THCa * 0.877 + d9-THC. **Total CBD = CBDa * 0.877 + CBD. LOQ = Limit of Quantification; NR = Not Reported; ND = Not Detected

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CA5E ORELAP 4057

Douglas Duncan
Laboratory Director

Confident Cannabis
All Rights Reserved
support@confidentcannabis.com
(866) 506-5866
www.confidentcannabis.com

Values reported relate only to the product tested. ChemHistory makes no claims as to the consumer safety or other risks associated with any detected or non-detected levels of any pesticides or residual solvents. This Certificate shall not be reproduced except in full, without the written approval of ChemHistory. ORELAP accredited ID#4057 and OLCC Licensed Cannabis Testing Laboratory #010-1002015CA5E.

CHEMHISTORY
CANNABIS TESTING LABORATORY

Certificate of Analysis

Powered by Confident Cannabis

Sample: 1810CH0391.1720

Strain: Elektra
Batch#: ; Batch Size: - grams
Sample Received: 10/24/2018; Report Created: 11/07/2018; Expires: 11/07/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF - Elektra

Plant, Flower - Cured, Outdoor
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Quality Control Data

Analytical Batch ID	QC Sample ID	Assay Name	QC Category Name
154 HPLC4 20181025-2	PMB102341802F	Cannabinoids	Blank

QC Notes
None

Compounds	Blank Result	Blank Max % (LOQ)	Blank Result %	Notes
THCa	0	0.05	<LOQ	-
Δ9-THC	0	0.05	<LOQ	-
CBDa	0	0.05	<LOQ	-
CBD	0	0.05	<LOQ	-

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CA5E ORELAP 4057

Confident Cannabis

support@confidentcannabis.com
(866) 506-5866
www.confidentcannabis.com

CHEMHISTORY
CANNABIS TESTING LABORATORY

Certificate of Analysis

Powered by Confident Cannabis

Sample: 1810CH0391.1720

Strain: Elektra
Batch#: ; Batch Size: - grams
Sample Received: 10/24/2018; Report Created: 11/07/2018; Expires: 11/07/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF - Elektra

Plant, Flower - Cured, Outdoor
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Quality Control Data

Analytical Batch ID	QC Sample ID	Assay Name	QC Category Name
154 HPLC4 20181025-2	PLCS102341802F	Cannabinoids	Lab Control Sample

QC Notes
None

Compounds	LCS Result (%)	LCS Spike	LCS % Rec	% Limits	Notes
THCa	4.68	5.23	89.48	70 - 130	-
Δ9-THC	3.04	3.11	97.75	70 - 130	-
CBDa	3.56	4	89	70 - 130	-
CBD	4.39	4.81	91.27	57 - 143	-

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CA5E ORELAP 4057

Confident Cannabis
support@confidentcannabis.com
(866) 506-5866
www.confidentcannabis.com

CHEMHISTORY
CANNABIS TESTING LABORATORY

Certificate of Analysis

Powered by Confident Cannabis
4 of 4

Sample: 1810CH0391.1720

Strain: Elektra
Batch#: ; Batch Size: - grams
Sample Received: 10/24/2018; Report Created: 11/07/2018; Expires: 11/07/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF - Elektra

Plant, Flower - Cured, Outdoor
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Quality Control Data

Analytical Batch ID	QC Sample ID	Assay Name	QC Category Name
154 HPLC4 20181025-2	PMS102341802F	Cannabinoids	Spike

QC Notes
None

Compounds	LOQ	Matrix Duplicate % RPD	Duplicate RPD Notes	Matrix Spike % Rec	% Limits	Notes
THCa	0.02	8	<20	100	70 - 130	-
Δ9-THC	0.02	0	<20	100	70 - 130	-
CBDa	0.02	0.88	<20	78	70 - 130	-
CBD	0.02	6.9	<20	95	70 - 130	-

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CA5E ORELAP 4057

CHEMHISTORY
CANNABIS TESTING LABORATORY

Certificate of Analysis

Industrial Hemp

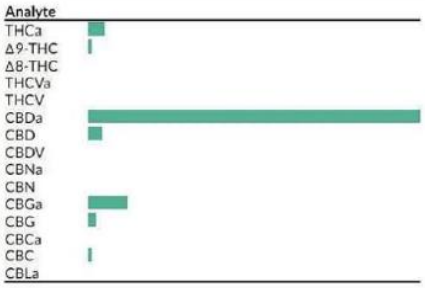
Powered by Confident Cannabis
1 of 5

Sample: 1809CH0213.0782

Strain: OF-Lifter
Batch#: 3; Batch Size: 5 - grams
Sample Received: 09/17/2018; Report Created: 09/26/2018; Expires: 09/26/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Lifter

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:



Cannabinoids

Complete

240 HPLC3 20180920-3
09/20/2018

0.02% Δ9-THC	**14.23%** Total CBD** (Calculated Decarboxylated Potential)
19.08% Total Cannabinoids Analyzed	**10.5%** Moisture

Analyte	LOQ mg/g	Mass %	Mass mg/g
THCa		0.75	7.5
Δ9-THC		0.02	0.2
Δ8-THC		<LOQ	<LOQ
THCVa		<LOQ	<LOQ
THCV		<LOQ	<LOQ
CBDa		15.55	155.5
CBD		0.59	5.9
CBDV		<LOQ	<LOQ
CBNa		<LOQ	<LOQ
CBN		<LOQ	<LOQ
CBGa		1.84	18.4
CBG		0.32	3.2
CBCa		<LOQ	<LOQ
CBC		0.02	0.2
CBLa		<LOQ	<LOQ
Total		**19.08**	**190.8**

*Total THC = THCa * 0.877 + d9-THC, **Total CBD = CBDa * 0.877 + CBD. LOQ = Limit of Quantification; NR = Not Reported; ND = Not Detected

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CA5E ORELAP 4057

Douglas Duncan
Laboratory Director

Confident Cannabis
All Rights Reserved
support@confidentcannabis.com
(866) 506-5866
www.confidentcannabis.com

CHEMHISTORY
CANNABIS TESTING LABORATORY

Certificate of Analysis

Informational Use Only - Not For Regulatory Use

Powered by Confident Cannabis
2 of 5

Sample: 1809CH0213.0782

Strain: OF-Lifter
Batch#: 3; Batch Size: 5 - grams
Sample Received: 09/17/2018; Report Created: 09/26/2018; Expires: 09/26/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Lifter

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Terpenes

338 GCFID1 20180918-2
09/18/2018

Analyte	Mass %	Mass mg/g	LOQ %
α-Pinene	0.04	0.4	0.02
Camphene	<LOQ	<LOQ	0.02
Sabinene	<LOQ	<LOQ	0.02
(-) -β-Pinene	0.03	0.3	0.02
β-Myrcene	0.51	5.1	0.02
α-Phellandrene	<LOQ	<LOQ	0.02
δ-3-Carene	<LOQ	<LOQ	0.02
α-Terpinene	<LOQ	<LOQ	0.02
p-Isopropyltoluene	<LOQ	<LOQ	0.02
δ-Limonene	0.05	0.5	0.02
Eucalyptol	<LOQ	<LOQ	0.02
Ocimene 1	<LOQ	<LOQ	0.02
Ocimene 2	0.09	0.9	0.02
γ-Terpinene	<LOQ	<LOQ	0.02
Sabinene Hydrate	<LOQ	<LOQ	0.02
Fenchone	<LOQ	<LOQ	0.02
Terpinolene	<LOQ	<LOQ	0.02
Linalool	0.04	0.4	0.02
Endo-Fenchyl Alcohol	<LOQ	<LOQ	0.02
Camphor	<LOQ	<LOQ	0.02
(-)-Isopulegol	<LOQ	<LOQ	0.02

Analyte	Mass %	Mass mg/g	LOQ %
Isoborneol	<LOQ	<LOQ	0.02
Borneol	<LOQ	<LOQ	0.02
Hexahydro Thymol	<LOQ	<LOQ	0.02
α-Terpineol	<LOQ	<LOQ	0.02
γ-Terpineol	<LOQ	<LOQ	0.02
Nerol	<LOQ	<LOQ	0.02
Pulegone	<LOQ	<LOQ	0.02
Geraniol	<LOQ	<LOQ	0.02
Farnesene	<LOQ	<LOQ	0.02
α-Cedrene	<LOQ	<LOQ	0.02
β-Caryophyllene	0.37	3.7	0.02
α-Humulene	0.13	1.3	0.02
Valencene	<LOQ	<LOQ	0.02
Geranyl Acetate	0.40	4.0	0.02
cis-Nerolidol	0.04	0.4	0.02
β-Nerolidol	0.09	0.9	0.02
trans-Nerolidol	0.04	0.4	0.02
Caryophyllene Oxide	<LOQ	<LOQ	0.02
(-)-Guaiol	0.04	0.4	0.02
Cedrol	<LOQ	<LOQ	0.02
(-)-α-Bisabolol	0.13	1.3	0.02

Primary Aromas

2.00% Total Terpenes	Pine	Hops	Orange	Earthy	Lavender

Method: GC-FID; based on dry weight; LOQ = Limit of Quantification; NR = Not Reported; ND = Not Detected

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CA5E ORELAP 4057

Douglas Duncan
Laboratory Director

CHEMHISTORY
CANNABIS TESTING LABORATORY

Certificate of Analysis

Powered by Confident Cannabis

Sample: 1809CH0213.0782

Strain: OF-Lifter
Batch#: 3; Batch Size: 5 - grams
Sample Received: 09/17/2018; Report Created: 09/26/2018; Expires: 09/26/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Lifter

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Quality Control Data

Analytical Batch ID	QC Sample ID	Assay Name	QC Category Name
240 HPLC3 20180920-3	PMB09191801H	Cannabinoids	Blank

QC Notes
None

Compounds	Blank Result	Blank Max % (LOQ)	Blank Result %	Notes
THCa	0	0.002	<LOQ	-
Δ9-THC	0	0.002	<LOQ	-
CBDa	0	0.002	<LOQ	-
CBD	0	0.002	<LOQ	-

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CA5E ORELAP 4057

CHEMHISTORY
CANNABIS TESTING LABORATORY

Certificate of Analysis

Powered by Confident Cannabis

Sample: 1809CH0213.0782

Strain: OF-Lifter
Batch#: 3; Batch Size: 5 - grams
Sample Received: 09/17/2018; Report Created: 09/26/2018; Expires: 09/26/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Lifter
Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Quality Control Data

Analytical Batch ID	QC Sample ID	Assay Name	QC Category Name
240 HPLC3 20180920-3	PLCS09191801H	Cannabinoids	Lab Control Sample

QC Notes
None

Compounds	LCS Result (%)	LCS Spike	LCS % Rec	% Limits	Notes
THCa	7.11	6.57	108.22	70 - 130	-
Δ9-THC	7.56	7.32	103.28	70 - 130	-
CBDa	2.34	2.19	106.85	70 - 130	-
CBD	2.45	2.3	106.52	57 - 143	-

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CA5E ORELAP 4057

CHEMHISTORY
CANNABIS TESTING LABORATORY

Certificate of Analysis

Powered by Confident Cannabis
5 of 5

Sample: 1809CH0213.0782

Strain: OF-Lifter
Batch#: 3; Batch Size: 5 - grams
Sample Received: 09/17/2018; Report Created: 09/26/2018; Expires: 09/26/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Lifter

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Quality Control Data

Analytical Batch ID	QC Sample ID	Assay Name	QC Category Name
240 HPLC3 20180920-3	PMS09191801H	Cannabinoids	Spike

QC Notes
None

Compounds	LOQ	Matrix Duplicate % RPD	Duplicate RPD Notes	Matrix Spike % Rec	% Limits	Notes
THCa	0.002	8.7	<20	105	70 - 130	-
Δ9-THC	0.002	0	<20	100	70 - 130	-
CBDa	0.002	7.57	<20	33	70 - 130	Below Limit Range
CBD	0.002	22.22	>20	92	70 - 130	-

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CA5E ORELAP 4057

CHEMHISTORY

Certificate of Analysis

Industrial Hemp

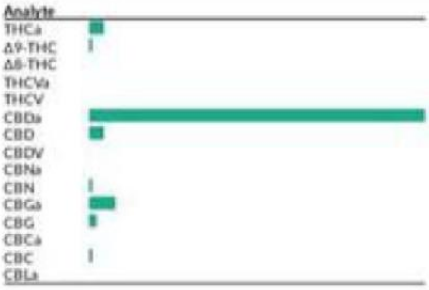
Powered by Confident Cannabis
1 of 5

Sample: 1809CH0213.0780

Strain: OF-Sour Space Candy
Batch#: 1; Batch Size: 5 - grams
Sample Received: 09/17/2018; Report Created: 09/26/2018; Expires: 09/26/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Sour Space Candy

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:



Cannabinoids

Complete

240HPLC32018092O-3
09/20/2018

0.03%	12.96%
Δ9-THC	Total CBD* (Calculated Decarboxylated Potential)
16.82%	12.2%
Total Cannabinoids Analyzed	Moisture

Analyte	LOQ	Mass %	Mass mg/g
	mg/g		
THCa		0.61	6.1
Δ9-THC		0.03	0.3
Δ8-THC		<LOQ	<LOQ
THCVa		<LOQ	<LOQ
THCV		<LOQ	<LOQ
CBDa		14.12	141.2
CBD		0.58	5.8
CBDV		<LOQ	<LOQ
CBNa		<LOQ	<LOQ
CBN		0.06	0.6
CBGa		1.13	11.3
CBG		0.28	2.8
CBCa		<LOQ	<LOQ
CBC		0.02	0.2
CBLa		<LOQ	<LOQ
Total		**16.82**	**168.2**

*Total THC = THCa * 0.877 + d9-THC, **Total CBD = CBDa * 0.877 + CBD. LOQ = Limit of Quantification, NR = Not Reported, ND = Not Detected

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CASE ORELAP 4057

Douglas Duncan
Laboratory Director

Values reported relate only to the product tested. ChemHistory makes no claims as to the consumer safety or other risks associated with any detected or non detected levels of any pesticides or residual solvents. This Certificate shall not be reproduced except in full, without the written approval of ChemHistory. ORELAP accredited ID#4057 and OLCC Licensed Cannabis Testing Laboratory #010-1002015CASE.

CHEMHISTORY

Certificate of Analysis

Informational Use Only - Not For Regulatory Use

Powered by Confident Cannabis
2 of 5

Sample: 1809CH0213.0780

Strain: OF-Sour Space Candy
Batch#: 1; Batch Size: 5 - grams
Sample Received: 09/17/2018; Report Created: 09/26/2018; Expires: 09/26/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Sour Space Candy

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Terpenes

338GCFID1 20180918-1
09/18/2018

Analyte	Mass %	Mass mg/g	LOQ %	Analyte	Mass %	Mass mg/g	LOQ %
α-Pinene	0.35	3.5	0.02	Isoborneol	<LOQ	<LOQ	0.02
Camphene	<LOQ	<LOQ	0.02	Borneol	<LOQ	<LOQ	0.02
Sabinene	<LOQ	<LOQ	0.02	Hexahydro Thymol	<LOQ	<LOQ	0.02
(-)-β-Pinene	0.14	1.4	0.02	α-Terpineol	<LOQ	<LOQ	0.02
β-Myrcene	0.97	9.7	0.02	γ-Terpineol	<LOQ	<LOQ	0.02
α-Phellandrene	<LOQ	<LOQ	0.02	Nerol	<LOQ	<LOQ	0.02
δ-3-Carene	<LOQ	<LOQ	0.02	Pulegone	<LOQ	<LOQ	0.02
α-Terpinene	<LOQ	<LOQ	0.02	Geraniol	<LOQ	<LOQ	0.02
p-Isopropyltoluene	<LOQ	<LOQ	0.02	Farnesene	<LOQ	<LOQ	0.02
δ-Limonene	0.09	0.9	0.02	α-Cedrene	<LOQ	<LOQ	0.02
Eucalyptol	<LOQ	<LOQ	0.02	β-Caryophyllene	0.34	3.4	0.02
Ocimene 1	<LOQ	<LOQ	0.02	α-Humulene	0.10	1.0	0.02
Ocimene 2	0.10	1.0	0.02	Valencene	0.03	0.3	0.02
γ-Terpinene	<LOQ	<LOQ	0.02	Geranyl Acetate	0.30	3.0	0.02
Sabinene Hydrate	<LOQ	<LOQ	0.02	cis-Nerolidol	0.12	1.2	0.02
Fenchone	<LOQ	<LOQ	0.02	β-Nerolidol	0.16	1.6	0.02
Terpinolene	0.33	3.3	0.02	trans-Nerolidol	0.03	0.3	0.02
Linalool	0.03	0.3	0.02	Caryophyllene Oxide	<LOQ	<LOQ	0.02
Endo-Fenchyl Alcohol	<LOQ	<LOQ	0.02	(-)-Guaiol	0.04	0.4	0.02
Camphor	<LOQ	<LOQ	0.02	Cedrol	<LOQ	<LOQ	0.02
(-)-Isopulegol	<LOQ	<LOQ	0.02	(-)-α-Bisabolol	0.06	0.6	0.02

Primary Aromas

3.19%					
Total Terpenes	Pine	Hops	Orange	Earthy	Turpentine

Method: GC-FID; based on dry weight; LOQ = Limit of Quantification; NR = Not Reported; ND = Not Detected

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CASE ORELAP 4057

Douglas Duncan
Laboratory Director

CHEMHISTORY

Certificate of Analysis

Powered by Confident Cannabis
3 of 5

Sample: 1809CH0213.0780

Strain: OF-Sour Space Candy
Batch#: 1; Batch Size: 5 - grams
Sample Received: 09/17/2018; Report Created: 09/26/2018; Expires: 09/26/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Sour Space Candy

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Quality Control Data

Analytical Batch ID	QC Sample ID	Assay Name	QC Category Name
240 HPLC3 20180920-3	PMB0919181801H	Cannabinoids	Blank

QC Notes
None

Compounds	Blank Result	Blank Max % (LOQ)	Blank Result %	Notes
THCa	0	0.002	<LOQ	-
Δ9-THC	0	0.002	<LOQ	-
CBDa	0	0.002	<LOQ	-
CBD	0	0.002	<LOQ	-

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CASE ORELAP 4057

CHEMHISTORY

Certificate of Analysis

Powered by Confident Cannabis

Sample: 1809CH0213.0780

Strain: OF-Sour Space Candy
Batch#: 1; Batch Size: 5 - grams
Sample Received: 09/17/2018; Report Created: 09/26/2018; Expires: 09/26/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Sour Space Candy

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Quality Control Data

Analytical Batch ID	QC Sample ID	Assay Name	QC Category Name
240 HPLC 3 20180920-3	PLCS0919180 1H	Cannabinoids	Lab Control Sample

QC Notes
None

Compounds	LCS Result (%)	LCS Spike	LCS % Rec	% Limits	Notes
THCa	7.11	6.57	108.22	70 - 130	-
Δ9-THC	7.56	7.32	103.28	70 - 130	-
CBDa	2.34	2.19	106.85	70 - 130	-
CBD	2.45	2.3	106.52	57 - 143	-

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CASE ORELAP 4057

CHEMHISTORY

Certificate of Analysis

Powered by Confident Cannabis
5 of 5

Sample: 1809CH0213.0780

Strain: OF-Sour Space Candy
Batch#: 1; Batch Size: 5 - grams
Sample Received: 09/17/2018; Report Created: 09/26/2018; Expires: 09/26/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Sour Space Candy

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Quality Control Data

Analytical Batch ID	QC Sample ID	Assay Name	QC Category Name
240 HPLC 3 20180920-3	PMS09191801H	Cannabinoids	Spike

QC Notes
None

Compounds	LOQ	Matrix Duplicate % RPD	Duplicate RPD Notes	Matrix Spike % Rec	% Limits	Notes
THCa	0.002	8.7	<20	105	70 - 130	-
Δ9-THC	0.002	0	<20	100	70 - 130	-
CBDa	0.002	7.57	<20	33	70 - 130	Below Limit Range
CBD	0.002	22.22	<20	92	70 - 130	-

5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CASE ORELAP 4057

CHEMHISTORY
CANNABIS TESTING LABORATORY

Certificate of Analysis

Industrial Hemp

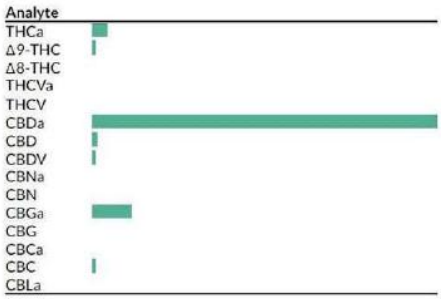
Powered by Confident Cannabis
1 of 5

Sample: 1810CH0054.0177

Strain: None
Batch#: ; Batch Size: - grams
Sample Received: 10/02/2018; Report Created: 10/05/2018; Expires: 10/05/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Special Sauce-BD-Bud

Plant, Flower - Cured, Outdoor
Harvest Process Lot: ; METRC Batch: ; METRC Sample:



Cannabinoids

Pass

136 HPLC4 20181003-4
10/03/2018

0.02% Δ9-THC	**16.61%** Total CBD** (Calculated Decarboxylated Potential)
21.88% Total Cannabinoids Analyzed	**10.2%** Moisture

Analyte	LOQ	Mass	Mass
	mg/g	%	mg/g
THCa	0.2	0.77	7.7
Δ9-THC	0.2	0.02	0.2
Δ8-THC	0.2	<LOQ	<LOQ
THCVa	0.2	<LOQ	<LOQ
THCV	0.2	<LOQ	<LOQ
CBDa	0.2	18.65	186.5
CBD	0.2	0.25	2.5
CBDV	0.2	0.03	0.3
CBNa	0.2	<LOQ	<LOQ
CBN	0.2	<LOQ	<LOQ
CBGa	0.2	2.14	21.4
CBG	0.2	<LOQ	<LOQ
CBCa	0.2	<LOQ	<LOQ
CBC	0.2	0.02	0.2
CBLa	0.2	<LOQ	<LOQ
Total		**21.88**	**218.8**

*Total THC = THCa * 0.877 + d9-THC. **Total CBD = CBDa * 0.877 + CBD. LOQ = Limit of Quantification; NR = Not Reported; ND = Not Detected

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CA5E ORELAP 4057

Douglas Duncan
Laboratory Director

CHEMHISTORY
CANNABIS TESTING LABORATORY

Certificate of Analysis

Informational Use Only - Not For Regulatory Use

Powered by Confident Cannabis
2 of 5

Sample: 1810CH0054.0177

Strain: None
Batch#: ; Batch Size: - grams
Sample Received: 10/02/2018; Report Created: 10/05/2018; Expires: 10/05/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Special Sauce-BD-Bud

Plant, Flower - Cured, Outdoor
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Terpenes

348 GCFID1 20181005-1
10/05/2018

Analyte	Mass %	Mass mg/g	LOQ %
α-Pinene	0.02	0.2	
Camphene	<LOQ	<LOQ	
Sabinene	<LOQ	<LOQ	
(-) -β-Pinene	0.02	0.2	
β-Myrcene	0.69	6.9	
α-Phellandrene	<LOQ	<LOQ	
δ-3-Carene	<LOQ	<LOQ	
α-Terpinene	<LOQ	<LOQ	
p-Isopropyltoluene	<LOQ	<LOQ	
δ-Limonene	0.09	0.9	
Eucalyptol	<LOQ	<LOQ	
Ocimene 1	<LOQ	<LOQ	
Ocimene 2	0.12	1.2	
γ-Terpinene	<LOQ	<LOQ	
Sabinene Hydrate	<LOQ	<LOQ	
Fenchone	<LOQ	<LOQ	
Terpinolene	<LOQ	<LOQ	
Linalool	0.04	0.4	
Endo-Fenchyl Alcohol	0.02	0.2	
Camphor	<LOQ	<LOQ	
(-)-Isopulegol	0.01	0.1	

Analyte	Mass %	Mass mg/g	LOQ %
Isoborneol	<LOQ	<LOQ	
Borneol	<LOQ	<LOQ	
Hexahydro Thymol	<LOQ	<LOQ	
α-Terpineol	0.03	0.3	
γ-Terpineol	<LOQ	<LOQ	
Nerol	<LOQ	<LOQ	
Pulegone	<LOQ	<LOQ	
Geraniol	<LOQ	<LOQ	
Farnesene	<LOQ	<LOQ	
α-Cedrene	<LOQ	<LOQ	
β-Caryophyllene	0.33	3.3	
α-Humulene	0.11	1.1	
Valencene	0.02	0.2	
Geranyl Acetate	0.67	6.7	
cis-Nerolidol	0.08	0.8	
β-Nerolidol	0.10	1.0	
trans-Nerolidol	0.04	0.4	
Caryophyllene Oxide	0.04	0.4	
(-)-Guaiol	0.03	0.3	
Cedrol	0.01	0.1	
(-)-α-Bisabolol	0.20	2.0	

Primary Aromas

2.67%					
Total Terpenes	Pine	Hops	Orange	Earthy	Lavender

Method: GC-FID; based on dry weight; LOQ = Limit of Quantification; NR = Not Reported; ND = Not Detected

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CA5E ORELAP 4057

Douglas Duncan
Laboratory Director

Confident Cannabis

support@confidentcannabis.com
(866) 506-5866
www.confidentcannabis.com

CHEMHISTORY
CANNABIS TESTING LABORATORY

Certificate of Analysis

Powered by Confident Cannabis
3 of 5

Sample: 1810CH0054.0177

Strain: None
Batch#: ; Batch Size: - grams
Sample Received: 10/02/2018; Report Created: 10/05/2018; Expires: 10/05/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Special Sauce-BD-Bud
Plant, Flower - Cured, Outdoor
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Quality Control Data

Analytical Batch ID	QC Sample ID	Assay Name	QC Category Name
136 HPLC4 20181003-4	PMB10021802F	Cannabinoids	Blank

QC Notes
None

Compounds	Blank Result	Blank Max % (LOQ)	Blank Result %	Notes
THCa	0	0.05	<LOQ	-
Δ9-THC	0	0.05	<LOQ	-
CBDa	0	0.05	<LOQ	-
CBD	0	0.05	<LOQ	-

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CA5E ORELAP 4057

Confident Cannabis

support@confidentcannabis.com
(866) 506-5866
www.confidentcannabis.com

CHEMHISTORY
CANNABIS TESTING LABORATORY

Certificate of Analysis

Powered by Confident Cannabis

Sample: 1810CH0054.0177

Strain: None
Batch#: ; Batch Size: - grams
Sample Received: 10/02/2018; Report Created: 10/05/2018; Expires: 10/05/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Special Sauce-BD-Bud
Plant, Flower - Cured, Outdoor
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Quality Control Data

Analytical Batch ID	QC Sample ID	Assay Name	QC Category Name
136 HPLC4 20181003-4	PLCS10021802F	Cannabinoids	Lab Control Sample

QC Notes
None

Compounds	LCS Result (%)	LCS Spike	LCS % Rec	% Limits	Notes
THCa	6.79	6.57	103.35	70 - 130	-
Δ9-THC	7.29	7.32	99.59	70 - 130	-
CBDa	2.12	2.19	96.8	70 - 130	-
CBD	2.27	2.3	98.7	57 - 143	-

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CA5E ORELAP 4057

CHEMHISTORY
CANNABIS TESTING LABORATORY

Certificate of Analysis

Powered by Confident Cannabis
5 of 5

Sample: 1810CH0054.0177

Strain: None
Batch#: ; Batch Size: - grams
Sample Received: 10/02/2018; Report Created: 10/05/2018; Expires: 10/05/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Special Sauce-BD-Bud

Plant, Flower - Cured, Outdoor
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Quality Control Data

Analytical Batch ID	QC Sample ID	Assay Name	QC Category Name
136 HPLC4 20181003-4	PMS10021802F	Cannabinoids	Spike

QC Notes
None

Compounds	LOQ	Matrix Duplicate % RPD	Duplicate RPD Notes	Matrix Spike % Rec	% Limits	Notes
THCa	0.02	1.8	<20	92	70 - 130	-
Δ9-THC	0.02	9.9	<20	98	70 - 130	-
CBDa	0.02	15.38	<20	76	70 - 130	-
CBD	0.02	<LOQ	<LOQ	99	70 - 130	-

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CA5E ORELAP 4057

Confident Cannabis

support@confidentcannabis.com
(866) 506-5866
www.confidentcannabis.com

CHEMHISTORY
CANNABIS TESTING LABORATORY

Certificate of Analysis

Industrial Hemp

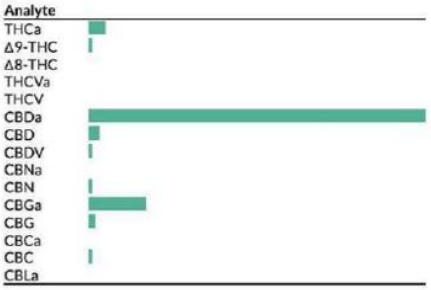
Powered by Confident Cannabis
1 of 5

Sample: 1809CH0213.0781

Strain: OF-Cascade
Batch#: 2; Batch Size: 0 - grams
Sample Received: 09/17/2018; Report Created: 09/26/2018; Expires: 09/26/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Cascade

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:



Cannabinoids

240 HPLC3 20180920-3
09/20/2018

Complete

0.04%	13.63%
Δ9-THC	Total CBD** (Calculated Decarboxylated Potential)
19.12%	**10.5%**
Total Cannabinoids Analyzed	Moisture

Analyte	LOQ	Mass	Mass
	mg/g	%	mg/g
THCa		0.72	7.2
Δ9-THC		0.04	0.4
Δ8-THC		<LOQ	<LOQ
THCVa		<LOQ	<LOQ
THCV		<LOQ	<LOQ
CBDa		15.04	150.4
CBD		0.44	4.4
CBDV		0.04	0.4
CBNa		<LOQ	<LOQ
CBN		0.01	0.1
CBGa		2.52	25.2
CBG		0.27	2.7
CBCa		<LOQ	<LOQ
CBC		0.04	0.4
CBLa		<LOQ	<LOQ
Total		**19.12**	**191.2**

*Total THC = THCa * 0.877 + d9-THC. **Total CBD = CBDa * 0.877 + CBD. LOQ = Limit of Quantification; NR = Not Reported; ND = Not Detected

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CA5E ORELAP 4057

Douglas Duncan
Laboratory Director

Confident Cannabis

support@confidentcannabis.com
(866) 506-5866
www.confidentcannabis.com

CHEMHISTORY
CANNABIS TESTING LABORATORY

Certificate of Analysis

Informational Use Only - Not For Regulatory Use

Powered by Confident Cannabis
2 of 5

Sample: 1809CH0213.0781

Strain: OF-Cascade
Batch#: 2; Batch Size: 0 - grams
Sample Received: 09/17/2018; Report Created: 09/26/2018; Expires: 09/26/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Cascade

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Terpenes

338 GCFID1 20180918-1
09/18/2018

Analyte	Mass %	Mass mg/g	LOQ %
α-Pinene	0.39	3.9	0.02
Camphene	<LOQ	<LOQ	0.02
Sabinene	<LOQ	<LOQ	0.02
(-) -β-Pinene	0.14	1.4	0.02
β-Myrcene	<LOQ	<LOQ	0.02
α-Phellandrene	<LOQ	<LOQ	0.02
δ-3-Carene	<LOQ	<LOQ	0.02
α-Terpinene	<LOQ	<LOQ	0.02
p-Isopropyltoluene	<LOQ	<LOQ	0.02
δ-Limonene	0.16	1.6	0.02
Eucalyptol	<LOQ	<LOQ	0.02
Ocimene 1	<LOQ	<LOQ	0.02
Ocimene 2	<LOQ	<LOQ	0.02
γ-Terpinene	<LOQ	<LOQ	0.02
Sabinene Hydrate	<LOQ	<LOQ	0.02
Fenchone	<LOQ	<LOQ	0.02
Terpinolene	<LOQ	<LOQ	0.02
Linalool	0.05	0.5	0.02
Endo-Fenchyl Alcohol	<LOQ	<LOQ	0.02
Camphor	<LOQ	<LOQ	0.02
(-)-Isopulegol	<LOQ	<LOQ	0.02

Analyte	Mass %	Mass mg/g	LOQ %
Isoborneol	<LOQ	<LOQ	0.02
Borneol	<LOQ	<LOQ	0.02
Hexahydro Thymol	<LOQ	<LOQ	0.02
α-Terpineol	<LOQ	<LOQ	0.02
γ-Terpineol	<LOQ	<LOQ	0.02
Nerol	<LOQ	<LOQ	0.02
Pulegone	<LOQ	<LOQ	0.02
Geraniol	<LOQ	<LOQ	0.02
Farnesene	<LOQ	<LOQ	0.02
α-Cedrene	<LOQ	<LOQ	0.02
β-Caryophyllene	0.34	3.4	0.02
α-Humulene	0.08	0.8	0.02
Valencene	<LOQ	<LOQ	0.02
Geranyl Acetate	0.30	3.0	0.02
cis-Nerolidol	<LOQ	<LOQ	0.02
β-Nerolidol	0.04	0.4	0.02
trans-Nerolidol	<LOQ	<LOQ	0.02
Caryophyllene Oxide	<LOQ	<LOQ	0.02
(-)-Guaiol	<LOQ	<LOQ	0.02
Cedrol	<LOQ	<LOQ	0.02
(-)-α-Bisabolol	0.09	0.9	0.02

Primary Aromas

1.59%					
Total Terpenes	Pine	Orange	Lavender	Cinnamon	Hops

Method: GC-FID; based on dry weight; LOQ = Limit of Quantification; NR = Not Reported; ND = Not Detected

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CA5E ORELAP 4057

Douglas Duncan
Laboratory Director

CHEMHISTORY
CANNABIS TESTING LABORATORY

Certificate of Analysis

Powered by Confident Cannabis

Sample: 1809CH0213.0781

Strain: OF-Cascade
Batch#: 2; Batch Size: 0 - grams
Sample Received: 09/17/2018; Report Created: 09/26/2018; Expires: 09/26/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Cascade

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Quality Control Data

Analytical Batch ID	QC Sample ID	Assay Name	QC Category Name
240 HPLC3 20180920-3	PMB09191801H	Cannabinoids	Blank

QC Notes
None

Compounds	Blank Result	Blank Max % (LOQ)	Blank Result %	Notes
THCa	0	0.002	<LOQ	-
Δ9-THC	0	0.002	<LOQ	-
CBDa	0	0.002	<LOQ	-
CBD	0	0.002	<LOQ	-

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CA5E ORELAP 4057

CHEMHISTORY
CANNABIS TESTING LABORATORY

Certificate of Analysis

Powered by Confident Cannabis

Sample: 1809CH0213.0781

Strain: OF-Cascade
Batch#: 2; Batch Size: 0 - grams
Sample Received: 09/17/2018; Report Created: 09/26/2018; Expires: 09/26/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Cascade

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Quality Control Data

Analytical Batch ID	QC Sample ID	Assay Name	QC Category Name
240 HPLC3 20180920-3	PLCS09191801H	Cannabinoids	Lab Control Sample

QC Notes
None

Compounds	LCS Result (%)	LCS Spike	LCS % Rec	% Limits	Notes
THCa	7.11	6.57	108.22	70 - 130	-
Δ9-THC	7.56	7.32	103.28	70 - 130	-
CBDa	2.34	2.19	106.85	70 - 130	-
CBD	2.45	2.3	106.52	57 - 143	-

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CA5E ORELAP 4057

CHEMHISTORY
CANNABIS TESTING LABORATORY

Certificate of Analysis

Powered by Confident Cannabis
5 of 5

Sample: 1809CH0213.0781

Strain: OF-Cascade
Batch#: 2; Batch Size: 0 - grams
Sample Received: 09/17/2018; Report Created: 09/26/2018; Expires: 09/26/2019
Harvest/Production Date:
Sampling: Random; Environment: Room Temp

OF-Cascade

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Quality Control Data

Analytical Batch ID	QC Sample ID	Assay Name	QC Category Name
240 HPLC3 20180920-3	PMS09191801H	Cannabinoids	Spike

QC Notes
None

Compounds	LOQ	Matrix Duplicate % RPD	Duplicate RPD Notes	Matrix Spike % Rec	% Limits	Notes
THCa	0.002	8.7	<20	105	70 - 130	-
Δ9-THC	0.002	0	<20	100	70 - 130	-
CBDa	0.002	7.57	<20	33	70 - 130	Below Limit Range
CBD	0.002	22.22	>20	92	70 - 130	-

CHEMHISTORY
5691 SE International Way C-2
Portland, OR
(503) 305-5252
http://chemhistory.com
Lic# OLCC 010-1002015CA5E ORELAP 4057

Confident Cannabis

support@confidentcannabis.com
(866) 506-5866
www.confidentcannabis.com

NV CANNLABS
FROM ART TO SCIENCE

PJLA Testing Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis
1 of 2

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1909NVC1403-7044

Strain: Hemp World Haze

Sample Received: 09/09/2019; Report Created: 09/12/2019

Hemp World Haze

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

<LOQ	<LOQ	20.537%
THCa	Total THC	Total CBD

Cannabinoids

Analyte	LOQ	Mass	Mass
	%	%	mg/g
THCa	0.115	<LOQ	<LOQ
Δ9-THC	0.115	<LOQ	<LOQ
CBDa	0.115	17.882	178.82
CBD	0.115	4.854	48.54
CBC	0.058	0.371	3.71
CBG	0.058	0.160	1.60
CBN	0.115	<LOQ	<LOQ
THCV	0.058	0.316	3.16
Δ8-THC	0.058	<LOQ	<LOQ
CBGa	0.058	0.399	3.99
CBDV	0.058	<LOQ	<LOQ
Total		**23.982**	**239.82**

Notes:

Total THC = THCa * 0.877 + Δ9-THC + Δ8-THC
Total CBD = CBDa * 0.877 + CBD
LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Cannabinoids analyzed by SOP-021.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Scientific Operations Director

All pass limits are as specified in NAC 453.A and DPBH Policies. Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. This product has been tested by NV Cann Labs using valid testing methodologies and a quality system as required by Nevada state law. Values reported relate only to the product tested. NV Cann Labs makes no claims as to the efficacy, safety or other risks associated with any detected or non-detected levels of any compounds reported herein. This Certificate shall not be reproduced except in full, without the written approval of NV Cann Labs. Uncertainty information is available upon request. All analysis were performed at NV Cann Labs unless otherwise stated. Sampling Plan SOP-001 and Sampling Method SOP-027 were used to collect samples. If sample(s) are NOT collected by NV Cann Labs, result(s) apply to the samples as received.

NV CANNLABS
FROM ART TO SCIENCE

PJLA Testing Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis
2 of 2

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1909NVC1403-7044

Strain: Hemp World Haze

Sample Received: 09/09/2019; Report Created: 09/12/2019

Hemp World Haze

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Terpenes

Analyte	LOQ	Mass	Mass
	mg/g	mg/g	%
β-Caryophyllene	0.092	2.559	0.2559
β-Myrcene	0.092	2.444	0.2444
α-Humulene	0.092	1.434	0.1434
α-Bisabolol	0.092	1.009	0.1009
(-)-Guaiol	0.092	0.685	0.0685
α-Pinene	0.092	0.599	0.0599
δ-Limonene	0.092	0.510	0.0510
Linalool	0.092	0.496	0.0496
Caryophyllene Oxide	0.092	0.456	0.0456
(-) -β-Pinene	0.092	0.394	0.0394
Nerolidol	0.092	0.236	0.0236
Ocimene	0.092	0.154	0.0154
α-Terpinene	0.092	<0.092	<0.0092
Camphene	0.092	<0.092	<0.0092
δ-3-Carene	0.092	<0.092	<0.0092
γ-Terpinene	0.092	<0.092	<0.0092
Geraniol	0.092	<0.092	<0.0092
(-)-Isopulegol	0.092	<0.092	<0.0092
p-Cymene	0.092	<0.092	<0.0092
Terpinolene	0.092	<0.092	<0.0092

Primary Aromas

- Cinnamon
- Hops
- Chamomile
- Wood
- Pine

LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Terpenes analyzed by SOP-022.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Brenda Shalloo
Scientific Operations Director

All pass limits are as specified in NAC 453.A and DPBH Policies. Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. This product has been tested by NV Cann Labs using valid testing methodologies and a quality system as required by Nevada state law. Values reported relate only to the product tested. NV Cann Labs makes no claims as to the efficacy, safety or other risks associated with any detected or non-detected levels of any compounds reported herein. This Certificate shall not be reproduced except in full, without the written approval of NV Cann Labs. Uncertainty information is available upon request. All analysis were performed at NV Cann Labs unless otherwise stated. Sampling Plan SOP-001 and Sampling Method SOP-027 were used to collect samples. If sample(s) are NOT collected by NV Cann Labs, result(s) apply to the samples as received.

NV CANNLABS

PJLA Testing Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis
1 of 2

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1910NVC1598-8480

Strain: Sunset Sherbert #2

Sample Received: 10/07/2019; Report Created: 10/10/2019

Sunset Sherbert #2

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

<LOQ	<LOQ	19.489%
THCa	Total THC	Total CBD

Cannabinoids

Analyte	LOQ	Mass	Mass
	%	%	mg/g
THCa	0.253	<LOQ	<LOQ
Δ9-THC	0.253	<LOQ	<LOQ
CBDa	0.253	21.956	219.56
CBD	0.025	0.234	2.34
CBC	0.126	0.163	1.63
CBG	0.013	0.097	0.97
CBN	0.253	<LOQ	<LOQ
THCV	0.126	0.198	1.98
Δ8-THC	0.126	<LOQ	<LOQ
CBGa	0.126	0.611	6.11
CBDV	0.126	<LOQ	<LOQ
Total		**24.202**	**242.02**

Notes:

Total THC = THCa * 0.877 + Δ9-THC + Δ8-THC
Total CBD = CBDa * 0.877 + CBD
LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Cannabinoids analyzed by SOP-021.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Scientific Operations Director

All pass limits are as specified in NAC 453.A and Taxation Department Second Policies. Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. This product has been tested by NV Cann Labs using valid testing methodologies and a quality system as required by Nevada state law. Values reported relate only to the product tested. NV Cann Labs makes no claims as to the efficacy, safety or other risks associated with any detected or non-detected levels of any compounds reported herein. This Certificate shall not be reproduced except in full, without the written approval of NV Cann Labs. Uncertainty and statement of conformity are available upon request. All analysis were performed at NV Cann Labs unless otherwise stated. Sampling Plan SOP-001 and Sampling Method SOP-027 were used to collect samples. If sample(s) are NOT collected by NV Cann Labs, result(s) apply to the samples as received.

NV CANNLABS

PJLA Testing Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis
2 of 2

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1910NVC1598-8480

Strain: Sunset Sherbet #2

Sample Received: 10/07/2019; Report Created: 10/10/2019

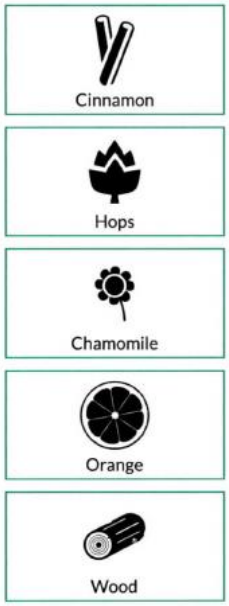
Sunset Sherbert #2

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Terpenes

Analyte	LOQ	Mass	Mass
	mg/g	mg/g	%
β-Caryophyllene	0.101	4.549	0.4549
β-Myrcene	0.101	2.325	0.2325
α-Humulene	0.101	1.497	0.1497
α-Bisabolol	0.101	0.991	0.0991
δ-Limonene	0.101	0.503	0.0503
Caryophyllene Oxide	0.101	0.447	0.0447
(-) -β-Pinene	0.101	0.270	0.0270
Linalool	0.101	0.223	0.0223
α-Pinene	0.101	<0.101	<0.0101
α-Terpinene	0.101	<0.101	<0.0101
Camphene	0.101	<0.101	<0.0101
δ-3-Carene	0.101	<0.101	<0.0101
γ-Terpinene	0.101	<0.101	<0.0101
Geraniol	0.101	<0.101	<0.0101
Nerolidol	0.101	<0.101	<0.0101
Ocimene	0.101	<0.101	<0.0101
(-)-Guaiol	0.101	<0.101	<0.0101
(-)-Isopulegol	0.101	<0.101	<0.0101
p-Cymene	0.101	<0.101	<0.0101
Terpinolene	0.101	<0.101	<0.0101

Primary Aromas



LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Terpenes analyzed by SOP-022.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Scientific Operations Director

All pass limits are as specified in NAC 453.A and Taxation Department Second Policies. Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. This product has been tested by NV Cann Labs using valid testing methodologies and a quality system as required by Nevada state law. Values reported relate only to the product tested. NV Cann Labs makes no claims as to the efficacy, safety or other risks associated with any detected or non-detected levels of any compounds reported herein. This Certificate shall not be reproduced except in full, without the written approval of NV Cann Labs. Uncertainty and statement of conformity are available upon request. All analysis were performed at NV Cann Labs unless otherwise stated. Sampling Plan SOP-001 and Sampling Method SOP-027 were used to collect samples. If sample(s) are NOT collected by NV Cann Labs, result(s) apply to the samples as received.

NV CANNLABS
FROM ART TO SCIENCE

PJLA Testing Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis
1 of 2

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1909NVC1403-7041

Strain: Paradise OG

Sample Received: 09/09/2019; Report Created: 09/12/2019

Paradise OG

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

<LOQ	<LOQ	21.814%
THCa	Total THC	Total CBD

Cannabinoids

Analyte	LOQ	Mass	Mass
	%	%	mg/g
THCa	0.116	<LOQ	<LOQ
Δ9-THC	0.116	<LOQ	<LOQ
CBDa	0.116	18.760	187.60
CBD	0.116	5.361	53.61
CBC	0.058	0.389	3.89
CBG	0.058	0.227	2.27
CBN	0.116	<LOQ	<LOQ
THCV	0.058	0.129	1.29
Δ8-THC	0.058	<LOQ	<LOQ
CBGa	0.058	0.509	5.09
CBDV	0.058	<LOQ	<LOQ
Total		**25.375**	**253.75**

Notes:

Total THC = THCa * 0.877 + Δ9-THC + Δ8-THC
Total CBD = CBDa * 0.877 + CBD
LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Cannabinoids analyzed by SOP-021.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Scientific Operations Director

All pass limits are as specified in NAC 453.A and DPBH Policies. Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. This product has been tested by NV Cann Labs using valid testing methodologies and a quality system as required by Nevada state law. Values reported relate only to the product tested. NV Cann Labs makes no claims as to the efficacy, safety or other risks associated with any detected or non-detected levels of any compounds reported herein. This Certificate shall not be reproduced except in full, without the written approval of NV Cann Labs. Uncertainty information is available upon request. All analysis were performed at NV Cann Labs unless otherwise stated. Sampling Plan SOP-001 and Sampling Method SOP-027 were used to collect samples. If sample(s) are NOT collected by NV Cann Labs, result(s) apply to the samples as received.

NV CANNLABS
FROM ART TO SCIENCE

PJLA Testing Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis
2 of 2

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1909NVC1403-7041

Strain: Paradise OG

Sample Received: 09/09/2019; Report Created: 09/12/2019

Paradise OG

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Terpenes

Analyte	LOQ	Mass	Mass
	mg/g	mg/g	%
β-Caryophyllene	0.093	2.248	0.2248
α-Bisabolol	0.093	1.483	0.1483
α-Humulene	0.093	1.032	0.1032
β-Myrcene	0.093	0.763	0.0763
(-)-Guaiol	0.093	0.493	0.0493
δ-Limonene	0.093	0.321	0.0321
Caryophyllene Oxide	0.093	0.318	0.0318
α-Pinene	0.093	0.307	0.0307
Linalool	0.093	0.304	0.0304
(-) -β-Pinene	0.093	0.196	0.0196
Nerolidol	0.093	0.194	0.0194
α-Terpinene	0.093	<0.093	<0.0093
Camphene	0.093	<0.093	<0.0093
δ-3-Carene	0.093	<0.093	<0.0093
γ-Terpinene	0.093	<0.093	<0.0093
Geraniol	0.093	<0.093	<0.0093
Ocimene	0.093	<0.093	<0.0093
(-)-Isopulegol	0.093	<0.093	<0.0093
p-Cymene	0.093	<0.093	<0.0093
Terpinolene	0.093	<0.093	<0.0093

Primary Aromas


Cinnamon


Chamomile


Hops


Wood


Orange

LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Terpenes analyzed by SOP-022.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Brenda Shalloo
Scientific Operations Director

NV CANNLABS

PJLA Testing Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis
1 of 2

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1910NVC1598-8477

Strain: Casino Cookies #2

Sample Received: 10/07/2019; Report Created: 10/10/2019

Casino Cookies #2

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

<LOQ	<LOQ	18.269%
THCa	Total THC	Total CBD

Cannabinoids

Analyte	LOQ	Mass	Mass
	%	%	mg/g
THCa	0.259	<LOQ	<LOQ
Δ9-THC	0.259	<LOQ	<LOQ
CBDa	0.259	20.663	206.63
CBD	0.026	0.148	1.48
CBC	0.130	0.144	1.44
CBG	0.013	0.064	0.64
CBN	0.259	<LOQ	<LOQ
THCV	0.130	0.212	2.12
Δ8-THC	0.130	<LOQ	<LOQ
CBGa	0.130	0.641	6.41
CBDV	0.130	<LOQ	<LOQ
Total		**22.816**	**228.16**

Notes:

Total THC = THCa * 0.877 + Δ9-THC + Δ8-THC
Total CBD = CBDa * 0.877 + CBD
LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Cannabinoids analyzed by SOP-021.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Brenda Shalloo
Scientific Operations Director

All pass limits are as specified in NAC 453.A and Taxation Department Second Policies. Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. This product has been tested by NV Cann Labs using valid testing methodologies and a quality system as required by Nevada state law. Values reported relate only to the product tested. NV Cann Labs makes no claims as to the efficacy, safety or other risks associated with any detected or non-detected levels of any compounds reported herein. This Certificate shall not be reproduced except in full, without the written approval of NV Cann Labs. Uncertainty and statement of conformity are available upon request. All analysis were performed at NV Cann Labs unless otherwise stated. Sampling Plan SOP-001 and Sampling Method SOP-027 were used to collect samples. If sample(s) are NOT collected by NV Cann Labs, result(s) apply to the samples as received.

NV CANNLABS

PJLA Testing Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis
2 of 2

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1910NVC1598-8477

Strain: Casino Cookies #2

Sample Received: 10/07/2019; Report Created: 10/10/2019

Casino Cookies #2

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Terpenes

Analyte	LOQ	Mass	Mass
	mg/g	mg/g	%
β-Caryophyllene	0.104	6.221	0.6221
β-Myrcene	0.104	3.274	0.3274
α-Humulene	0.104	1.917	0.1917
α-Bisabolol	0.104	1.198	0.1198
δ-Limonene	0.104	0.704	0.0704
Caryophyllene Oxide	0.104	0.493	0.0493
(-) -β-Pinene	0.104	0.310	0.0310
Linalool	0.104	0.293	0.0293
Ocimene	0.104	0.133	0.0133
α-Pinene	0.104	0.112	0.0112
α-Terpinene	0.104	<0.104	<0.0104
Camphene	0.104	<0.104	<0.0104
δ-3-Carene	0.104	<0.104	<0.0104
γ-Terpinene	0.104	<0.104	<0.0104
Geraniol	0.104	<0.104	<0.0104
Nerolidol	0.104	<0.104	<0.0104
(-)-Guaiol	0.104	<0.104	<0.0104
(-)-Isopulegol	0.104	<0.104	<0.0104
p-Cymene	0.104	<0.104	<0.0104
Terpinolene	0.104	<0.104	<0.0104

Primary Aromas

- Cinnamon
- Hops
- Chamomile
- Orange
- Wood

LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Terpenes analyzed by SOP-022.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Scientific Operations Director

NV CANNLABS
FROM ART TO SCIENCE

PJLA Testing Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1909NVC1403-7042

Strain: Durban Potion #2

Sample Received: 09/09/2019; Report Created: 09/12/2019

Durban Potion #2

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

<LOQ	<LOQ	14.802%
THCa	Total THC	Total CBD

Cannabinoids

Analyte	LOQ	Mass	Mass
	%	%	mg/g
THCa	0.119	<LOQ	<LOQ
Δ9-THC	0.119	<LOQ	<LOQ
CBDa	0.119	14.622	146.22
CBD	0.119	1.979	19.79
CBC	0.060	0.171	1.71
CBG	0.060	0.108	1.08
CBN	0.119	<LOQ	<LOQ
THCV	0.060	0.096	0.96
Δ8-THC	0.060	<LOQ	<LOQ
CBGa	0.060	0.545	5.45
CBDV	0.060	<LOQ	<LOQ
Total		**17.521**	**175.21**

Notes:

Total THC = THCa * 0.877 + Δ9-THC + Δ8-THC
Total CBD = CBDa * 0.877 + CBD
LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Cannabinoids analyzed by SOP-021.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Scientific Operations Director

All pass limits are as specified in NAC 453.A and DPBH Policies. Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. This product has been tested by NV Cann Labs using valid testing methodologies and a quality system as required by Nevada state law. Values reported relate only to the product tested. NV Cann Labs makes no claims as to the efficacy, safety or other risks associated with any detected or non-detected levels of any compounds reported herein. This Certificate shall not be reproduced except in full, without the written approval of NV Cann Labs. Uncertainty information is available upon request. All analysis were performed at NV Cann Labs unless otherwise stated. Sampling Plan SOP-001 and Sampling Method SOP-027 were used to collect samples. If sample(s) are NOT collected by NV Cann Labs, result(s) apply to the samples as received.

NV CANNLABS
FROM ART TO SCIENCE

PJLA Testing Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis
2 of 2

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1909NVC1403-7042

Strain: Durban Potion #2

Sample Received: 09/09/2019; Report Created: 09/12/2019

Durban Potion #2

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Terpenes

Analyte	LOQ	Mass	Mass
	mg/g	mg/g	%
β-Caryophyllene	0.095	2.102	0.2102
α-Bisabolol	0.095	1.270	0.1270
α-Humulene	0.095	1.032	0.1032
β-Myrcene	0.095	0.760	0.0760
(-)-Guaiol	0.095	0.696	0.0696
δ-Limonene	0.095	0.365	0.0365
Ocimene	0.095	0.322	0.0322
Terpinolene	0.095	0.316	0.0316
Caryophyllene Oxide	0.095	0.310	0.0310
Nerolidol	0.095	0.254	0.0254
Linalool	0.095	0.217	0.0217
α-Pinene	0.095	0.214	0.0214
(-) -β-Pinene	0.095	0.181	0.0181
α-Terpinene	0.095	<0.095	<0.0095
Camphene	0.095	<0.095	<0.0095
δ-3-Carene	0.095	<0.095	<0.0095
γ-Terpinene	0.095	<0.095	<0.0095
Geraniol	0.095	<0.095	<0.0095
(-)-Isopulegol	0.095	<0.095	<0.0095
p-Cymene	0.095	<0.095	<0.0095

Primary Aromas


Cinnamon


Chamomile


Hops


Wood


Orange

LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Terpenes analyzed by SOP-022.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Brenda Shalloo
Scientific Operations Director

NV CANNLABS
FROM ART TO SCIENCE

PJLA Testing Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1908NVC1326-6575

Strain: Tangie

Sample Received: 08/26/2019; Report Created: 08/29/2019

Tangie

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

<LOQ	<LOQ	17.121%
THCa	Total THC	Total CBD

Cannabinoids

Analyte	LOQ	Mass	Mass
	%	%	mg/g
THCa	0.110	<LOQ	<LOQ
Δ9-THC	0.110	<LOQ	<LOQ
CBDa	0.110	15.027	150.27
CBD	0.110	3.942	39.42
CBC	0.055	0.258	2.58
CBG	0.055	<LOQ	<LOQ
CBN	0.110	<LOQ	<LOQ
THCV	0.055	0.098	0.98
Δ8-THC	0.055	<LOQ	<LOQ
CBGa	0.055	0.192	1.92
CBDV	0.055	<LOQ	<LOQ
Total		**19.517**	**195.17**

Notes:

Total THC = THCa * 0.877 + Δ9-THC + Δ8-THC
Total CBD = CBDa * 0.877 + CBD
LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Cannabinoids analyzed by SOP-021.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Scientific Operations Director

All pass limits are as specified in NAC 453.A and DPBH Policies. Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. This product has been tested by NV Cann Labs using valid testing methodologies and a quality system as required by Nevada state law. Values reported relate only to the product tested. NV Cann Labs makes no claims as to the efficacy, safety or other risks associated with any detected or non-detected levels of any compounds reported herein. This Certificate shall not be reproduced except in full, without the written approval of NV Cann Labs. Uncertainty information is available upon request. All analysis were performed at NV Cann Labs unless otherwise stated. Sampling Plan SOP-001 and Sampling Method SOP-027 were used to collect samples. If sample(s) are NOT collected by NV Cann Labs, result(s) apply to the samples as received.

NV CANNLABS
FROM ART TO SCIENCE

PJLA Testing Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis
2 of 2

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1908NVC1326-6575

Strain: Tangie

Sample Received: 08/26/2019; Report Created: 08/29/2019

Tangie

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Terpenes

Analyte	LOQ	Mass	Mass
	mg/g	mg/g	%
β-Caryophyllene	0.088	2.278	0.2278
β-Myrcene	0.088	1.665	0.1665
α-Bisabolol	0.088	1.526	0.1526
α-Humulene	0.088	1.194	0.1194
(-)-Guaiol	0.088	0.931	0.0931
Linalool	0.088	0.468	0.0468
δ-Limonene	0.088	0.431	0.0431
Caryophyllene Oxide	0.088	0.427	0.0427
Nerolidol	0.088	0.296	0.0296
α-Pinene	0.088	0.234	0.0234
(-) -β-Pinene	0.088	0.207	0.0207
Ocimene	0.088	0.155	0.0155
α-Terpinene	0.088	<0.088	<0.0088
Camphene	0.088	<0.088	<0.0088
δ-3-Carene	0.088	<0.088	<0.0088
γ-Terpinene	0.088	<0.088	<0.0088
Geraniol	0.088	<0.088	<0.0088
(-)-Isopulegol	0.088	<0.088	<0.0088
p-Cymene	0.088	<0.088	<0.0088
Terpinolene	0.088	<0.088	<0.0088

Primary Aromas


Cinnamon


Hops


Chamomile


Wood


Lavender

LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Terpenes analyzed by SOP-022.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Brenda Shalloo
Scientific Operations Director

NV CANNLABS

PJLA Testing Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis
1 of 2

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1910NVC1598-8478

Strain: Juicy Fruit #2

Sample Received: 10/07/2019; Report Created: 10/10/2019

Juicy Fruits #2

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

<LOQ	<LOQ	18.284%
THCa	Total THC	Total CBD

Cannabinoids

Analyte	LOQ	Mass	Mass
	%	%	mg/g
THCa	0.255	<LOQ	<LOQ
Δ9-THC	0.255	<LOQ	<LOQ
CBDa	0.255	20.597	205.97
CBD	0.026	0.220	2.20
CBC	0.013	0.104	1.04
CBG	0.128	<LOQ	<LOQ
CBN	0.255	<LOQ	<LOQ
THCV	0.128	0.346	3.46
Δ8-THC	0.128	<LOQ	<LOQ
CBGa	0.128	0.588	5.88
CBDV	0.128	<LOQ	<LOQ
Total		**22.676**	**226.76**

Notes:

Total THC = THCa * 0.877 + Δ9-THC + Δ8-THC
Total CBD = CBDa * 0.877 + CBD
LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Cannabinoids analyzed by SOP-021.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Brenda Shalloo
Scientific Operations Director

All pass limits are as specified in NAC 453.A and Taxation Department Second Policies. Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. This product has been tested by NV Cann Labs using valid testing methodologies and a quality system as required by Nevada state law. Values reported relate only to the product tested. NV Cann Labs makes no claims as to the efficacy, safety or other risks associated with any detected or non-detected levels of any compounds reported herein. This Certificate shall not be reproduced except in full, without the written approval of NV Cann Labs. Uncertainty and statement of conformity are available upon request. All analysis were performed at NV Cann Labs unless otherwise stated. Sampling Plan SOP-001 and Sampling Method SOP-027 were used to collect samples. If sample(s) are NOT collected by NV Cann Labs, result(s) apply to the samples as received.

NV CANN LABS

PJLA Testing Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis
2 of 2

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1910NVC1598-8478

Strain: Juicy Fruit #2

Sample Received: 10/07/2019; Report Created: 10/10/2019

Juicy Fruits #2

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Terpenes

Analyte	LOQ	Mass	Mass
	mg/g	mg/g	%
β-Myrcene	0.102	6.156	0.6156
α-Humulene	0.102	2.869	0.2869
Terpinolene	0.102	2.148	0.2148
Nerolidol	0.102	1.422	0.1422
α-Pinene	0.102	1.228	0.1228
Ocimene	0.102	1.181	0.1181
(-) -β-Pinene	0.102	0.798	0.0798
δ-Limonene	0.102	0.758	0.0758
α-Bisabolol	0.102	0.406	0.0406
(-)-Guaiol	0.102	0.381	0.0381
Linalool	0.102	0.277	0.0277
α-Terpinene	0.102	<0.102	<0.0102
β-Caryophyllene	0.102	<0.102	<0.0102
Camphene	0.102	<0.102	<0.0102
Caryophyllene Oxide	0.102	<0.102	<0.0102
δ-3-Carene	0.102	<0.102	<0.0102
γ-Terpinene	0.102	<0.102	<0.0102
Geraniol	0.102	<0.102	<0.0102
(-)-Isopulegol	0.102	<0.102	<0.0102
p-Cymene	0.102	<0.102	<0.0102

Primary Aromas

Hops

Turpentine

Orange

Pine

Earthy

LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Terpenes analyzed by SOP-022.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Scientific Operations Director

All pass limits are as specified in NAC 453.A and Taxation Department Second Policies. Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. This product has been tested by NV Cann Labs using valid testing methodologies and a quality system as required by Nevada state law. Values reported relate only to the product tested. NV Cann Labs makes no claims as to the efficacy, safety or other risks associated with any detected or non-detected levels of any compounds reported herein. This Certificate shall not be reproduced except in full, without the written approval of NV Cann Labs. Uncertainty and statement of conformity are available upon request. All analysis were performed at NV Cann Labs unless otherwise stated. Sampling Plan SOP-001 and Sampling Method SOP-027 were used to collect samples. If sample(s) are NOT collected by NV Cann Labs, result(s) apply to the samples as received.

NV CANNLABS
FROM ART TO SCIENCE

PJLA Testing Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis
1 of 2

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1909NVC1403-7043

Strain: Trophy Wife

Sample Received: 09/09/2019; Report Created: 09/12/2019

Trophy Wife

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

<LOQ	<LOQ	17.264%
THCa	Total THC	Total CBD

Cannabinoids

Analyte	LOQ %	Mass %	Mass mg/g
THCa	0.117	<LOQ	<LOQ
Δ9-THC	0.117	<LOQ	<LOQ
CBDa	0.117	12.839	128.39
CBD	0.117	6.004	60.04
CBC	0.059	0.379	3.79
CBG	0.059	0.138	1.38
CBN	0.117	<LOQ	<LOQ
THCV	0.059	0.160	1.60
Δ8-THC	0.059	<LOQ	<LOQ
CBGa	0.059	0.338	3.38
CBDV	0.059	<LOQ	<LOQ
Total		**19.858**	**198.58**

Notes:

Total THC = THCa * 0.877 + Δ9-THC + Δ8-THC
Total CBD = CBDa * 0.877 + CBD
LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Cannabinoids analyzed by SOP-021.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Scientific Operations Director

NV CANNLABS
FROM ART TO SCIENCE

PJLA Testing Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis
2 of 2

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1909NVC1403-7043

Strain: Trophy Wife

Sample Received: 09/09/2019; Report Created: 09/12/2019

Trophy Wife

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Terpenes

Analyte	LOQ	Mass	Mass
	mg/g	mg/g	%
β-Caryophyllene	0.094	1.993	0.1993
α-Humulene	0.094	1.140	0.1140
β-Myrcene	0.094	1.061	0.1061
α-Bisabolol	0.094	0.983	0.0983
(-)-Guaiol	0.094	0.847	0.0847
Caryophyllene Oxide	0.094	0.372	0.0372
δ-Limonene	0.094	0.366	0.0366
α-Pinene	0.094	0.297	0.0297
Linalool	0.094	0.250	0.0250
Terpinolene	0.094	0.223	0.0223
(-) -β-Pinene	0.094	0.198	0.0198
Ocimene	0.094	0.191	0.0191
α-Terpinene	0.094	<0.094	<0.0094
Camphene	0.094	<0.094	<0.0094
δ-3-Carene	0.094	<0.094	<0.0094
γ-Terpinene	0.094	<0.094	<0.0094
Geraniol	0.094	<0.094	<0.0094
Nerolidol	0.094	<0.094	<0.0094
(-)-Isopulegol	0.094	<0.094	<0.0094
p-Cymene	0.094	<0.094	<0.0094

Primary Aromas


Cinnamon


Hops


Chamomile


Wood


Orange

LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Terpenes analyzed by SOP-022.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Scientific Operations Director

NV CANNLABS
FROM ART TO SCIENCE

PJLA Testing Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis
1 of 2

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1910NVC1598-8476

Strain: CBG Flower

Sample Received: 10/07/2019; Report Created: 10/10/2019

CBG Flower

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

0.241%	0.211%	<LOQ
THCa	Total THC	Total CBD

Cannabinoids

Analyte	LOQ	Mass	Mass
	%	%	mg/g
THCa	0.025	0.241	2.41
Δ9-THC	0.253	<LOQ	<LOQ
CBDa	0.253	<LOQ	<LOQ
CBD	0.253	<LOQ	<LOQ
CBC	0.126	0.176	1.76
CBG	0.126	<LOQ	<LOQ
CBN	0.253	<LOQ	<LOQ
THCV	0.126	<LOQ	<LOQ
Δ8-THC	0.126	<LOQ	<LOQ
CBGa	0.126	21.825	218.25
CBDV	0.126	<LOQ	<LOQ
Total		**22.242**	**222.42**

Notes:

Total THC = THCa * 0.877 + Δ9-THC + Δ8-THC
Total CBD = CBDa * 0.877 + CBD
LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Cannabinoids analyzed by SOP-021.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Scientific Operations Director

NV CANNLABS
FROM ART TO SCIENCE

PJLA Testing Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis
2 of 2

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1910NVC1598-8476

Strain: CBG Flower

Sample Received: 10/07/2019; Report Created: 10/10/2019

CBG Flower

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Terpenes

Analyte	LOQ	Mass	Mass
	mg/g	mg/g	%
α-Bisabolol	0.101	2.981	0.2981
β-Caryophyllene	0.101	2.750	0.2750
(-)-Guaiol	0.101	2.678	0.2678
α-Humulene	0.101	0.831	0.0831
β-Myrcene	0.101	0.579	0.0579
Caryophyllene Oxide	0.101	0.277	0.0277
δ-Limonene	0.101	0.252	0.0252
Linalool	0.101	0.170	0.0170
α-Pinene	0.101	<0.101	<0.0101
α-Terpinene	0.101	<0.101	<0.0101
Camphene	0.101	<0.101	<0.0101
δ-3-Carene	0.101	<0.101	<0.0101
γ-Terpinene	0.101	<0.101	<0.0101
Geraniol	0.101	<0.101	<0.0101
Nerolidol	0.101	<0.101	<0.0101
Ocimene	0.101	<0.101	<0.0101
(-) -β-Pinene	0.101	<0.101	<0.0101
(-)-Isopulegol	0.101	<0.101	<0.0101
p-Cymene	0.101	<0.101	<0.0101
Terpinolene	0.101	<0.101	<0.0101

Primary Aromas

- Chamomile
- Cinnamon
- Wood
- Hops
- Orange

LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Terpenes analyzed by SOP-022.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Scientific Operations Director

All pass limits are as specified in NAC 453.A and Taxation Department Second Policies. Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. This product has been tested by NV Cann Labs using valid testing methodologies and a quality system as required by Nevada state law. Values reported relate only to the product tested. NV Cann Labs makes no claims as to the efficacy, safety or other risks associated with any detected or non-detected levels of any compounds reported herein. This Certificate shall not be reproduced except in full, without the written approval of NV Cann Labs. Uncertainty and statement of conformity are available upon request. All analysis were performed at NV Cann Labs unless otherwise stated. Sampling Plan SOP-001 and Sampling Method SOP-027 were used to collect samples. If sample(s) are NOT collected by NV Cann Labs, result(s) apply to the samples as received.

NV CANNLABS
FROM ART TO SCIENCE

PJLA Testing Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis
1 of 2

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1910NVC1598-8479

Strain: Jazzy CBG

Sample Received: 10/07/2019; Report Created: 10/10/2019

Jazzy CBG

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

0.240%	0.210%	<LOQ
THCa	Total THC	Total CBD

Cannabinoids

Analyte	LOQ %	Mass %	Mass mg/g
THCa	0.025	0.240	2.40
Δ9-THC	0.251	<LOQ	<LOQ
CBDa	0.251	<LOQ	<LOQ
CBD	0.251	<LOQ	<LOQ
CBC	0.126	0.182	1.82
CBG	0.126	<LOQ	<LOQ
CBN	0.251	<LOQ	<LOQ
THCV	0.126	<LOQ	<LOQ
Δ8-THC	0.126	<LOQ	<LOQ
CBGa	0.126	19.721	197.21
CBDV	0.126	<LOQ	<LOQ
Total		**20.143**	**201.43**

Notes:

Total THC = THCa * 0.877 + Δ9-THC + Δ8-THC
Total CBD = CBDa * 0.877 + CBD
LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Cannabinoids analyzed by SOP-021.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Brenda Shalloo
Scientific Operations Director

All pass limits are as specified in NAC 453.A and Taxation Department Second Policies. Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. This product has been tested by NV Cann Labs using valid testing methodologies and a quality system as required by Nevada state law. Values reported relate only to the product tested. NV Cann Labs makes no claims as to the efficacy, safety or other risks associated with any detected or non-detected levels of any compounds reported herein. This Certificate shall not be reproduced except in full, without the written approval of NV Cann Labs. Uncertainty and statement of conformity are available upon request. All analysis were performed at NV Cann Labs unless otherwise stated. Sampling Plan SOP-001 and Sampling Method SOP-027 were used to collect samples. If sample(s) are NOT collected by NV Cann Labs, result(s) apply to the samples as received.

NV CANNLABS
FROM ART TO SCIENCE

PJLA Testing
Accreditation #97453

Certificate of Analysis

Powered by Confident Cannabis
2 of 2

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1910NVC1598-8479

Strain: Jazzy CBG

Sample Received: 10/07/2019; Report Created: 10/10/2019

Jazzy CBG

Plant, Flower - Cured
Harvest Process Lot: ; METRC Batch: ; METRC Sample:

Terpenes

Analyte	LOQ	Mass	Mass
	mg/g	mg/g	%
β-Caryophyllene	0.101	2.608	0.2608
α-Bisabolol	0.101	2.574	0.2574
(-)-Guaiol	0.101	2.092	0.2092
α-Humulene	0.101	0.790	0.0790
β-Myrcene	0.101	0.584	0.0584
Caryophyllene Oxide	0.101	0.339	0.0339
δ-Limonene	0.101	0.238	0.0238
Linalool	0.101	0.214	0.0214
α-Pinene	0.101	<0.101	<0.0101
α-Terpinene	0.101	<0.101	<0.0101
Camphene	0.101	<0.101	<0.0101
δ-3-Carene	0.101	<0.101	<0.0101
γ-Terpinene	0.101	<0.101	<0.0101
Geraniol	0.101	<0.101	<0.0101
Nerolidol	0.101	<0.101	<0.0101
Ocimene	0.101	<0.101	<0.0101
(-) -β-Pinene	0.101	<0.101	<0.0101
(-)-Isopulegol	0.101	<0.101	<0.0101
p-Cymene	0.101	<0.101	<0.0101
Terpinolene	0.101	<0.101	<0.0101

Primary Aromas

- Cinnamon
- Chamomile
- Wood
- Hops
- Orange

LOQ = Limit of Quantitation; The reported result is based on a sample weight with the applicable moisture content for that sample; Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Terpenes analyzed by SOP-022.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com

Brenda Shalloo
Scientific Operations Director

All pass limits are as specified in NAC 453.A and Taxation Department Second Policies. Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. This product has been tested by NV Cann Labs using valid testing methodologies and a quality system as required by Nevada state law. Values reported relate only to the product tested. NV Cann Labs makes no claims as to the efficacy, safety or other risks associated with any detected or non-detected levels of any compounds reported herein. This Certificate shall not be reproduced except in full, without the written approval of NV Cann Labs. Uncertainty and statement of conformity are available upon request. All analysis were performed at NV Cann Labs unless otherwise stated. Sampling Plan SOP-001 and Sampling Method SOP-027 were used to collect samples. If sample(s) are NOT collected by NV Cann Labs, result(s) apply to the samples as received.

VITA

Madeline G. Roman

EDUCATION

Ph.D. in Forensic Science, Sam Houston State University, Huntsville, Texas, **May 2020 (anticipated)**

- GPA: 4.0

B.S. in Biology, Texas A&M University, College Station, Texas, **2012**

RESEARCH EXPERIENCE

Dissertation Research (2017-present)
Department of Forensic Science, Sam Houston State University, Huntsville, Texas (Mentor: Rachel Houston, Ph.D.)

- Title of research: "Applications of forensic plant science in drug trafficking and environmental crimes"
- Research design, collected and interpreted data, prepared manuscripts, and submitted grant proposals
- Supervised and trained graduate students in forensic genetic techniques and instrument use and maintenance
- Other projects: comparison of the ForenSeq™ DNA Signature Prep Kit and GlobalFiler PCR Amplification Kit, evaluation of Investigator® 26plex QS Kit, comparison of swab types and evaluation of the SwabSaver® from Gentueri for collection and preservation of "touch" DNA, and development and validation of a qPCR assay for canine DNA

Capstone Research (2015-2017)
Department of Forensic Science, Sam Houston State University, Huntsville, Texas (Mentors: Sheree Hughes-Stamm, Ph.D. and Bobby LaRue, Ph.D.)

- Title of research: "DNA degradation in decomposing cadavers and gene expression in blood deposits as indicators of postmortem interval"
- Experience collecting human tissue samples at the Applied Anatomical Research Center (body farm)
- Other projects: evaluation of collection and preservation methods for decomposing tissues and comparison of swab types for collection of DNA from skin and muscle of decomposing cadavers

Forensic Genetics Laboratory Intern (2016)
Forensic Genetics Laboratory, Harris County Institute of Forensic Sciences, Houston, Texas (Mentor: Michael Donley, M.S.)

- Student trainee with experience in serology and DNA processing
- Performed internal validation of the Crime-Lite® ML2 for presumptively visualizing biological fluids and gunshot residue on fabric
- Assisted with internal validation of mixture interpretation guidelines for the GlobalFiler PCR Amplification Kit
- Performed comparison and internal validation of confirmatory blood tests

Senior Research Assistant (2013-2015)
Barshop Institute for Longevity and Aging Studies, The University of Texas Health Science Center at San Antonio, San Antonio, Texas (Mentor: Yuji Ikeno, M.D., Ph.D.)

- Designed and coordinated the development of new transgenic rat models
- Supervised and trained research assistants in pathological analysis of animals, preparation of tissues for histology, animal handling, genotyping, immunoblot analysis, immunohistochemical analysis, and other experimental techniques

Research Assistant (2012-2013)
Barshop Institute for Longevity and Aging Studies, The University of Texas Health Science Center at San Antonio, San Antonio, Texas (Mentor: Yuji Ikeno, M.D., Ph.D.)

- Performed molecular cloning experiments, rodent necropsy and harvesting of tissues, maintenance of animal colonies, and PCR genotyping
- Assisted with preparation and submission of grant proposals and manuscripts

Student Associate II (2009-2011)
Department of Biochemistry, The University of Texas Health Science Center at San Antonio, San Antonio, Texas (Mentor: Bettie Sue Masters, Ph.D.)

- Performed DNA isolation and purification, PCR mutagenesis, molecular cloning, extraction and purification of proteins, spectrophotometric enzyme assays, cell culture, and Western blot experiments

TEACHING EXPERIENCE

Graduate Laboratory Assistant (2019)
Department of Forensic Science, Sam Houston State University, Huntsville, Texas (Supervisors: Rachel Houston, Ph.D. and Bobby LaRue, Ph.D.)

- Teaching assistant for graduate Forensic Biology and Advanced Forensic DNA laboratory classes
- Laboratory maintenance, inventory, preparation of materials and reagents, administrative duties
- Teaching assistant trainings at Sam Houston State University include classroom management, Title IX, FERPA, managing conflict, community building, and classroom communication and feedback

Doctoral Teaching Fellow (2018)
Department of Forensic Science, Sam Houston State University, Huntsville, Texas (Supervisor: Jorn Yu, Ph.D.)

- Taught an online section of Introduction to Forensic Science (undergraduate FORS3366) on Blackboard Learn
- Faculty certification in Blackboard Learn at Sam Houston State University, 2019-2020; trainings include course building, communication, and assessment

PEER-REVIEWED PUBLICATIONS

Roman, M.G., Gutierrez, R., and Houston, R. "Massively parallel sequencing of *Cannabis sativa* chloroplast hotspots for determination of biogeographical origin and crop type." In preparation.

Roman, M.G., Cheng, Y.C., and Houston, R. "Evaluation of THCA synthase polymorphisms for distinguishing between marijuana and hemp." In preparation.

Roman, M.G., Gutierrez, R., Mayes, C., LaRue, B., and Houston, R. "Evaluation of the Investigator® 26plex QS STR kit and comparison with two commercially available STR kits." In preparation.

Roman, M.G., Flores, L.C., Cunningham, G.M., Cheng, C., Dube, S., Allen, C., Van Remmen, H., Bai, Y., Hubbard, G.B., Saunders, T.L., and Ikeno, Y. "Thioredoxin overexpression in mitochondria showed minimum effects on aging and age-related disease in male C57BL/6 mice." *Aging Pathology and Therapeutics.* In press.

Roman, M.G. and Houston, R. (2020). "Investigation of chloroplast regions rps16 and clpP for determination of *Cannabis sativa* crop type and biogeographical origin." *Legal Medicine*. Submitted.

Roman, M.G., Gangitano, D., Figueroa, A., Solano, J., Anabalon, L., and Houston, R. (2020). "Use of eucalyptus DNA profiling in a case of illegal logging." *Science & Justice*. Submitted.

Tasker, E., **Roman, M.G.**, Akosile, M., Mayes, C., Hughes-Stamm, S., and LaRue, B. (2020). "Efficacy of 'touch' DNA recovery and room-temperature storage from assault rifle magazines." *Legal Medicine*. In press.

Young, B.*, **Roman, M.G.***, LaRue, B., Gangitano, D., and Houston, R. (2019). "Evaluation of 19 short tandem repeat markers for individualization of *Papaver somniferum*." *Science & Justice*. In press.
*** Indicates equal contribution by authors.**

Di Nunzio, M., Houston, R., **Roman, M.G.**, Di Nunzio, C., Gangitano, D., and Barrot, C. (2019). "European validation of a *Cannabis sativa* 13-locus STR multiplex kit for genetic identification: A preliminary study." *Forensic Science International: Genetics Supplement Series*, 7(1), 224-226.

Roman, M.G., Gangitano, D., and Houston, R. (2019). "Characterization of new chloroplast markers to determine biogeographical origin and crop type of *Cannabis sativa*." *International Journal of Legal Medicine*, 133(6), 1721-1732.

Cunningham, G.M, Flores, L.C., **Roman, M.G.**, Cheng, C., Dube, S., Allen, C., Valentine, J.M., Hubbard, G.B., Bai, Y., Saunders, T.L., and Ikeno, Y. (2018). "Thioredoxin overexpression in both the cytosol and mitochondria accelerates age-related disease and shortens lifespan in male C57BL/6 mice." *GeroScience*, 40(5-6), 453-468.

Flores, L.C., **Roman, M.G.**, Cunningham, G.M., Cheng, C., Dube, S., Allen, C., Van Remmen, H., Hubbard, G.B, Saunders, T.L., and Ikeno, I. (2018). "Continuous overexpression of thioredoxin 1 enhances cancer development and does not extend maximum lifespan in male C57BL/6 mice." *Pathobiology of Aging and Age-related Diseases*, 8(1), 1533754.

Holmes, A., **Roman, M.G.**, and Hughes-Stamm, S. (2018). "In-field collection and preservation of decomposing human tissues to facilitate rapid purification and STR typing." *Forensic Science International: Genetics*, 36, 124-129.

Cunningham, G.M.*, **Roman, M.G.***, Flores, L.C., Hubbard, G.B., Salmon, A., Zhang, Y., Gelfond, J., and Ikeno, Y. (2015). "The paradoxical role of thioredoxin on cancer and aging." *Archives of Biochemistry and Biophysics, 576*, 32-38.
*** Indicates equal contribution by authors.**

Ikeno, Y., Hubbard, G.B., Lee, S., Dube, S., Flores, L.C., **Roman, M.**, and Bartke, A. (2013). "Do Ames dwarf and calorie-restricted mice share common effects on age-related pathology?" *Pathobiology of Aging & Age-related Diseases, 3*, 20833.

Hinchee-Rodriquez, K., Garg, N., Venkatakrishnan, P., **Roman, M.G.**, Adamo, M.L., Masters, B.S., and Roman, L.J. (2013) "Neuronal nitric oxide synthase is phosphorylated in response to insulin stimulation in skeletal muscle." *Biochemical and Biophysical Research Communications, 435*, 501-505.

Panda, S.P., Polusani, S.R., Kellogg, D.L., III, Venkatakrishnan, P., **Roman, M.G.**, Demeler, B., Masters, B.S., and Roman, L.J. (2013). "Intra- and inter-molecular effects of a conserved arginine residue of neuronal and inducible nitric oxide synthases on FMN and calmodulin binding." *The Archives of Biochemistry and Biophysics, 533*, 88-94.

PRESENTATIONS

Roman, M.G.*, Gutierrez, R., Mayes, C., LaRue, B., and Houston, R. (February 2020). "Evaluation of the Investigator® 26plex QS STR kit and comparison with two commercially available STR kits." Proceedings of the American Academy of Forensic Sciences Meeting, Anaheim, California.

Roman, M., Gangitano, D., and Houston, R.* (October 2019). "Development of *Cannabis sativa* autosomal and organelle genotyping methods for forensic and intelligence purposes." Proceedings of the 71st Southeastern Regional Meeting of the American Chemical Society, Savannah, Georgia.

Gutierrez, R., **Roman, M.**, Harrel, M., Hughes-Stamm, S., Houston, R., and LaRue, B.* "Improved mitochondrial region analysis of degraded bone and hair samples using a new small overlapping amplicon library preparation method for MiSeq FGx." Proceedings of the International Symposium on Human Identification Meeting, Palm Springs, California.

Roman, M.G.*, Gutierrez, R., Mayes, C., LaRue, B., and Houston, R. (September 2019). "Evaluation of the Investigator® 26plex QS STR kit and comparison with two commercially available STR kits." Proceedings of the International Symposium on Human Identification Meeting, Palm Springs, California.

Di Nunzio, M.*, Houston, R., **Roman, M.G.**, Di Nunzio, C., Gangitano, D., and Feixat, C.B. (September 2019). "European validation of a *Cannabis sativa* 13-locus STR multiplex kit for genetic identification: A preliminary study." Proceedings of the International Society for Forensic Genetics Meeting, Prague, Czech Republic.

Roman, M.G.*, Di Nunzio, M., Houston, R., Larue, B., and Gangitano, D. (September 2019). "Characterization of new chloroplast polymorphisms to determine biogeographical origin and crop type of *Cannabis sativa* samples." Proceedings of the International Society for Forensic Genetics Meeting, Prague, Czech Republic.

Gutierrez, R., **Roman, M.**, Houston, R, and LaRue, B*. (August 2019). "ForenSeq™ DNA Signature Prep and ForenSeq™ mtDNA control region solution." Proceedings of the Association of Forensic DNA Analysts and Administrators Meeting, Houston, Texas.

Roman, M.G.*, Houston, R., LaRue, B., and Gangitano, D. (February 2019). "Characterization of new chloroplast markers to determine biogeographical origin and crop type of *Cannabis sativa* samples." Proceedings of the American Academy of Forensic Sciences Meeting, Baltimore, Maryland.

Roman, M.G.*, Houston, R., LaRue, B., and Gangitano, D. (August 2018). "Investigation of ancestry markers in the *Cannabis sativa* chloroplast genome for determination of biogeographical origin and crop type." Proceedings of the Association of Forensic DNA Analysts and Administrators Meeting, Houston, Texas.

Holmes, A.S.*, **Roman, M.**, Gangitano, D., and Hughes-Stamm, S. (February-March 2018). "Alternate methods for collection, preservation & processing of DNA from decomposing human remains." Proceedings of Pittcon Conference & Expo, Orlando, Florida.

Roman, M.G.*, Holmes, A., Gangitano, D., and Hughes-Stamm, S. (February 2018). "Effect of body mass and cadaveric bloat on DNA quantity and downstream STR success." Proceedings of the American Academy of Forensic Sciences Meeting, Seattle, Washington.

Flores, L.C.*, Callahan, B.C., Cunningham, G.M., **Roman, M.G.**, Zhang, Y., Salmon, A., Liu, Y., Hubbard, G.B., and Ikeno, Y. (October 2015). "The effects of downregulating thioredoxin in the cytosol and mitochondria on aging." Proceedings of the 6th Annual Alliance for Healthy Aging Meeting, Newcastle Upon Tyne, United Kingdom.

Cunningham, G.M.*, **Roman, M.G.**, Flores, L.C., Zhang, Y., Salmon, A., Liu, Y., Hubbard, G.B., and Ikeno, Y. (May-June, 2015). "Overexpression of thioredoxin in mitochondria combined with downregulation in the cytosol alters aging in mice." Proceedings of the American Aging Association 43rd Annual Meeting: Aging and Geroscience: New Approaches to Old Problems, Marina Del Rey, California.

Roman, M.G.*, Flores, L.C., Cunningham, G.M., Zhang, Y., Salmon, A., Liu, Y., Hubbard, G.B., Kirkland, J., Pirtskhalava, T., Tchkonia, T., and Ikeno, Y. (May-June, 2015). "Mechanisms that extend lifespan in Sprague-Dawley rats overexpressing Cu/ZnSOD." Proceedings of the American Aging Association 43rd Annual Meeting: Aging and Geroscience: New Approaches to Old Problems, Marina Del Rey, California.

Cunningham, G.M., **Roman, M.G.**, Flores, L.C., Zhang, Y., Salmon, A., Liu, Y., Hubbard, G.B., and Ikeno, Y.* (November, 2014). "Thioredoxin down-regulation in the cytosol and mitochondria attenuates aging in mice." Proceedings of the Society for Free Radical Biology and Medicine 21st Annual Meeting, Seattle, Washington.

Roman, M.G., Flores, L.C., Cunningham, G.M., Zhang, Y., Salmon, A., Liu, Y., Hubbard, G.B., and Ikeno, Y.* (November, 2014). "Mechanisms that extend lifespan in Sprague-Dawley rats overexpressing Cu/ZnSOD." Proceedings of the Society for Free Radical Biology and Medicine 21st Annual Meeting, Seattle, Washington.

Invited symposium participant. Ikeno, Y.* (March, 2014). New insights and current concepts of aging. 17th Biennial Meeting of the Society for Free Radical Research International, Kyoto, Japan.

Roman, M.G.*, Flores, L.C., Cunningham, G.M., Zhang, Y., Salmon, A., Qi, W., Liu, Y., Hubbard, G.B., and Ikeno, Y. (May-June, 2014). "Thioredoxin overexpression in the mitochondria or down-regulation in the cytosol attenuates aging in mice." Proceedings of the American Aging Association 43rd Annual Meeting: Slowing Aging: Signaling, Stress

Resistance, Stem Cells, and Small Molecules, Insights from Old and New Models, San Antonio, Texas.

Flores, L.C.*, **Roman, M.G.**, Salmon, A.B., Musi, N., Lee, S., Hubbard, G.B., Van Remmen, H., Kirkland, J., Pirtskhalava, T., Tchkonia, T., Richardson, A., and Ikeno, Y. (November, 2013). "Cu/ZnSOD overexpression extends lifespan and increases protection against oxidative stress in obese rodents." Proceedings of the 4th Annual Mayo Clinic Robert and Arlene Kogod Center on Aging Conference: Molecular Mechanisms of Age-related Multi Morbidity, Groningen, Netherlands.

Roman, M.G.*, Flores, L.C., Bhattacharya, A., Zhang, Y., Salmon, A., Qi, W., Liu, Y., Lee, S., Van Remmen, H., Richardson, A., Hubbard, G.B., and Ikeno, Y. (November, 2013). "Thioredoxin overexpression in the mitochondria or down-regulation in the cytosol increases lifespan in mice." Proceedings of the 4th Annual Mayo Clinic Robert and Arlene Kogod Center on Aging Conference: Molecular Mechanisms of Age-related Multi Morbidity, Groningen, Netherlands.

Hinchee-Rodriquez, K.*, Garg, N., Venkatakrishnan, P., **Roman, M.G.**, Adamo, M.L., Masters, B.S., and Roman, L.J. (June, 2013). "A muscle-specific isoform of neuronal nitric oxide synthase is phosphorylated in response to insulin." Proceedings of the Endocrine Society Annual Meeting and Expo, San Francisco, California.

Dube, S., Flores, L.C., **Roman, M.**, Zhang, Y., Salmon, A., Qi, W., Liu, Y., Lee, S., Van Remmen, H., Richardson, A., Hubbard, G.B., and Ikeno, Y.* (May-June, 2013) "Thioredoxin 2 overexpression and thioredoxin 1 down-regulation attenuate aging through independent and common mechanisms." Proceedings of the American Aging Association 42nd Annual Meeting, Baltimore, Maryland.

Roman, M., Flores, L.C., Salmon, A.B., Dube, S., Musi, N., Lee, S., Hubbard, G.B., Van Remmen, H., Kirkland, J., Pirtskhalava, T., Tchkonia, T., Richardson, A., and Ikeno, Y.* (May-June, 2013). "Mechanisms that extend lifespan in Sprague-Dawley rats by Cu/ZnSOD overexpression." Proceedings of the American Aging Association 42nd Annual Meeting, Baltimore, Maryland.

Hinchee-Rodriquez, K.*, Garg, N., Venkatakrishnan, P., **Roman, M.G.**, Adamo, M.L., Masters, B.S., and Roman, L.J. (March-April, 2013). "A muscle-specific isoform of neuronal nitric oxide synthase is phosphorylated in response to insulin." Proceedings of the American Society for Biochemistry and Molecular Biology Annual Meeting, Boston, Massachusetts.

Dube, S.*, Ortiz, M., Flores, L.C., **Roman, M.**, Bhattacharya, A., Zhang, Y., Salmon, A., Qi, W., Liu, Y., Lee, S., Van Remmen, H., Richardson, A., Hubbard, G.B., and Ikeno, Y. (November, 2012). "Do thioredoxin 2 overexpression and thioredoxin 1 down-regulation attenuate aging through different mechanisms?" Proceedings of the 3rd Annual Mayo Clinic Robert and Arlene Kogod Center on Aging Conference: Senescence and Healthspan, Rochester, Minnesota.

Flores, L.C.*, Salmon, A.B., Ortiz, M., Dube, S., **Roman, M.**, Musi, N., Qi, W., Lee, S., Hubbard, G.B., Van Remmen, H., Battacharya, A., Liu, Y., Kirkland, J., Pirtskhalava, T., Tchkonia, T., Richardson, A., and Ikeno, Y. (November, 2012). "Cu/ZnSOD overexpression extended the lifespan of Sprague-Dawley rats but had little effect in F344 rats." Proceedings of the 3rd Annual Mayo Clinic Robert and Arlene Kogod Center on Aging Conference: Senescence and Healthspan, Rochester, Minnesota.

Roman, M.*, Flores, L.C., Ortiz, M., Dube, S., Zhang, Y., Lee, S., Hubbard, G.B., and Ikeno, Y. (November, 2012). "Potential mechanism for anti-tumor activity of calorie restriction on ethylnitrosourea-induced glioma in rats." Proceedings of the 3rd Annual Mayo Clinic Robert and Arlene Kogod Center on Aging Conference: Senescence and Healthspan, Rochester, Minnesota.

Dube, S.*, Flores, L.C., Salmon, A.B., Ortiz, M., **Roman, M.**, Musi, N., Qi, W., Lee, S., Hubbard, G.B., Van Remmen, H., Bhattacharya, A., Liu, Y., Kirkland, J., Pirtskhalava, T., Tchkonia, T., Richardson, A., and Ikeno, Y. (October, 2012). "Effects of Cu/ZnSOD overexpression on aging under obese and non-obese conditions in rats." Proceedings of the San Antonio Nathan Shock Center Conference on Aging: Mouse Healthspan: Why Lifespan is No Longer Enough, Bandera, Texas.

Ortiz, M.*, Cortez, L.A., Webb, C.R., Mahlke, M., Dube, S., **Roman, M.**, Bhattacharya, A., Zhang, Y., Salmon, A., Qi, W., Liu, Y., Lee, S., Van Remmen, H., Richardson, A., Hubbard, G.B., and Ikeno, Y. (October, 2012). "The effects of overexpressing or down-regulating thioredoxin 1 and 2 on aging and cancer." Proceedings of the San Antonio Nathan Shock Center Conference on Aging: Mouse Healthspan: Why Lifespan is No Longer Enough, Bandera, Texas.

Roman, M.*, Flores, L.C., Ortiz, M., Dube, S., Zhang, Y., Lee, S., Hubbard, G.B., and Ikeno, Y. (October, 2012). "Calorie restriction retards growth of ethylnitrosourea-induced glioma in rats." Proceedings of the San Antonio Nathan Shock Center Conference on Aging: Mouse Healthspan: Why Lifespan is No Longer Enough, Bandera, Texas.

AWARDS/ACTIVITIES

American Academy of Forensic Sciences – Student Affiliate Member (2016-Present)

Sam Houston State University Society of Forensic Science – Secretary (2016-2019)

Sam Houston State University Criminal Justice Summer Camp – Forensic Science Speaker (2017)

Academic community engagement – Forensic trace evidence presentation at Conroe High School, Conroe, Texas (2016)

Academic community engagement – DNA evidence presentation at Huntsville High School, Huntsville, Texas (2016)

Travel award recipient for the 3rd Annual Mayo Clinic Robert and Arlene Kogod Center on Aging Conference: *Senescence and Healthspan*, November 2012

LABORATORY SKILLS

Forensic DNA

- Serological analysis (ALS, presumptive tests, and confirmatory tests)
- DNA extraction (manual, QIAcube, QIAsymphony, EZ1 Advanced XL)
- DNA quantitation (7500 and Step-One Real-Time PCR Systems; Quantifiler® Duo and Trio DNA Quantification Kits, Investigator® Quantiplex Pro Kit, and SYBR™ Green PCR Master Mix)
- STR amplification (GlobalFiler and GlobalFiler Express PCR Amplification Kits, Investigator® 24plex QS and 24plex GO! Kits, Investigator® 26plex QS Kit, and AmpFLSTR Yfiler PCR Amplification Kit)
- Capillary electrophoresis (3500 and 3130xl Genetic Analyzers, GeneMapper™ ID and ID-X)
- Mixture interpretation and probabilistic genotyping (STRmix™)

DNA Sequencing

- Sanger sequencing (including mtDNA)
- Massively parallel sequencing (MiSeq)
- SNaPshot™ minisequencing

Forensic Chemistry

- Gas chromatography-mass spectrometry
- Liquid chromatography-mass spectrometry
- Fourier transform infrared spectroscopy
- Ion-mobility spectrometry
- Thin layer chromatography

Other Laboratory Skills

- Polymerase chain reaction, qPCR, and reverse-transcriptase PCR
- Agarose and polyacrylamide gel electrophoresis
- Immunoblot analysis
- Cell culture (bacterial and mammalian)
- Site-directed mutagenesis
- Protein expression and purification
- UV-vis spectroscopy and enzyme activity assays
- DNA cloning techniques, including restriction endonuclease digestion, ligation, and bacterial transformation
- FPLC

- Immunological, histological, and immunohistochemical analysis
- Tissue collection from human and rodent cadavers
- Rodent necropsy and gross pathological analysis