This is to certify that the

thesis entitled

DEVELOPMENT OF A MICROSATELLITE MARKER PANEL FOR GENOTYPING MICHIGAN WHITE-TAILED DEER

presented by

Laurie Ann Molitor

has been accepted towards fulfillment
of the requirements for

M.S. degree in Criminal Justice

Major professor

Date 12/14/99

DATE DUE	DATE DUE	DATE DUE

DEVELOPMENT OF A MICROSATELLITE MARKER PANEL FOR GENOTYPING MICHIGAN WHITE-TAILED DEER

By

Laurie Ann Molitor

A THESIS

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

MASTER OF SCIENCE

School of Criminal Justice

1999

D

impo

of ide

In for

chara

poten

to qua

docur

reacti

marke

of gen

distrib

indivi

tailed

one ele

allowi

have a

inbreed

ABSTRACT

DEVELOPMENT OF A MICROSATELLITE MARKER PANEL FOR GENOTYPING MICHIGAN WHITE-TAILED DEER

By

Laurie Ann Molitor

Deoxyribonucleic acid (DNA) microsatellite markers are becoming an increasingly important tool for uniquely identifying individuals. Forensic scientists face the challenge of identifying individuals to the exclusion of all others with a high degree of probability. In forensic wildlife cases, it is necessary to identify the animal involved in the crime by characterizing genetic variability among the species in order to obtain high exclusion potential. Wildlife animals are being poached and illegally imported at a rate too difficult to quantify because the suspect is rarely caught red-handed and there is very little documented data about the poaching problem. With the advent of polymerase chain reaction (PCR) and fluorescent detection methods, amplification of DNA microsatellite markers for identification purposes in forensic science is becoming a widely used method of genetic typing. The hypervariability of microsatellites with their widespread distribution and high abundance in the genome allows for genetically typing an individual. Three highly polymorphic loci were found in this study of Michigan white-tailed deer and were multiplexed together in one PCR assay and run simultaneously in one electrophoretic lane of an Applied Biosystems, Inc. (ABI) 377 DNA Sequencer, thus allowing for increased speed and low cost of analysis. The benefits of this study will have a positive impact on our ability to enforce wildlife laws, derive estimates of inbreeding within the population and assist in wildlife management strategies.

Mic

this

and

Chris

sugg

I

statis

vaiua

F

Oncol

for the

patient

ACKNOWLEDGEMENTS

I would like to thank Dr. Paul Coussens from the Department of Animal Science at Michigan State University for his guidance and assistance throughout the completion of this research project, and the Department of Natural Resources (DNR) and the U.S. Fish and Wildlife Service for funding this research project.

I would like to thank the committee members: Dr. Jay Siegel, Dr. Frank Horvath, Dr. Christina DeJong and Dr. Charles Barna for their time and effort reviewing and providing suggestions concerning this research project.

I would like to thank Dr. Kim Scribner and Dr. Robert Tempelman for their statistical expertise and Dr. Hsiao-Ching Liu and Dr. Christiane Hansen for sharing valuable knowledge and advice related to this research project.

Finally, I would like to thank Dr. Hans Cheng from the USDA Avian Disease and Oncology Laboratory in East Lansing for the use of his laboratory equipment, necessary for the completion of this research project, and my friends for supporting and being patient with me during the completion of this master's thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (CONT'D)

poss

illeg

shoo

thes

poac

relia

evid

typi

taile

mate

the n

posse

exclu

devel

by ma

wildli

T

indivi

suspec

NONTECHNICAL SUMMARY

Poaching, as defined by the Colorado Division of Wildlife, is the illegal taking or possession of any game, fish or non-game wildlife (Zumbo, 1999). Poaching entails the illegal hunting of wildlife on other people's property, taking wildlife out of season or shooting more than the amount allowed. With any violation of poaching, the cost of these poaching crimes is well into the billions (Zumbo, 1999). Wildlife animals are being poached and illegally imported at a high rate and presently there is no economical, reliable, sensitive or time-efficient method of providing critical evidence linking evidentiary samples to an individual unless apprehended "red-handed".

The objective of this study is to develop a practical, economical, time-efficient DNA typing system for wildlife forensic scientists to utilize for individualizing forensic white-tailed deer evidence. In forensic wildlife cases it is most usual to "match" evidentiary material to material in the possession of the suspect (i.e. the gut pile at the crime scene to the meat in the suspect's freezer or a drop of blood from the deer in the suspect's possession). Forensic scientists face the challenge of identifying individuals to the exclusion of all others with a high degree of probability or with utmost certainty. To develop a DNA typing system that can positively place a suspect at the scene of a crime by matching two separate deer samples provides a powerful law enforcement tool for wildlife officials.

Three regions of the deer's DNA were found to be highly variable between individuals and in combination could distinguish an innocent suspect from a guilty suspect with an extremely high degree of probability. That is, there is almost 100 percent

cer

sus

me

dev

wil

enf

(pre

crim

a su

mon

estab

is ve

criter

There

after t

the fu

accura

certainty that the evidence at the crime scene can be matched to the evidence in the suspect's possession. With the commission of any crime there needs to be a definitive method of placing the suspect at the scene of the crime and the DNA typing system developed in this project can do just that. This DNA typing system is a powerful tool for wildlife forensic scientists to utilize to match evidence samples and will assist law enforcement officers in enforcing wildlife laws and convicting poachers.

The benefits of having this DNA typing system are twofold in that of deterrence (prevention before the act of the crime) and of enforcement (conviction after the act of the crime). The enforcement of wildlife laws can immediately benefit by being able to place a suspect at the scene of the crime. The effect of deterrence is not so obvious because monitoring the effect of something not happening is hard to do unless a solid database is established to monitor and evaluate the numbers. According to Jim Zumbo (1999) there is very little data about poaching and the data is not being collected using universal criteria amongst organizations or agencies. The data can thus not be evaluated accurately. Therefore, the goal of this DNA typing system is to assist in the conviction of poachers after they commit the crimes and to prevent or deter criminal activity from occurring in the future. A future goal would be to establish a universal database from which an accurate evaluation of the deterrent effect can be achieved.

INTRODUCTION

Forensic scientists face the challenge of identifying individuals to the exclusion of all others with a high degree of probability. Evidence found at a crime scene is not always in the optimal condition for many traditional scientific assays. Many samples collected for analysis have been exposed to environmental insults, such as extreme temperatures, sunlight or moisture, which degrade the biological components. Typically proteins degrade rapidly while DNA may be more stable for use in several techniques.

Traditional serological techniques utilizing antigens, proteins or enzymes are limited to non-degraded samples and produce results with limited statistical probability of inclusion. Techniques using these biochemical markers are easy to perform and results can be scored unambiguously however, they are limited in their allelic variation which, in turn, lowers the statistical probabilities associated with matching the evidence to the suspect or victim.

Traditional restriction fragment length polymorphism (RFLP) technique utilizing DNA is also limited to non-degraded samples. This technique, while still being used extensively, is limited to the use of high quality DNA. Evidence samples containing non-degraded DNA are analyzed by this technique, which result in a unique "individual-specific" DNA profile. The RFLP technique has one of the highest discriminatory potentials and probability of inclusion between evidence and suspect amongst all the molecular biology DNA testing technologies.

PCR technology has superseded RFLP with its' ability to analyze very small and even partially degraded DNA isolated from environmentally challenged evidence

samples. This technique is performed utilizing minute quantities of DNA recovered from evidentiary material, amplified in repetitive cycles of denaturation, hybridization with specific primer pairs and extension to create an exponential amount of target DNA. PCR technology amplifying microsatellite markers with fluorescent detection results in a highly powerful discriminatory tool for identification purposes. Combining several microsatellite markers in a single analysis has allowed this technology to approach the discriminatory potential and probability of RFLP technology.

In forensic wildlife cases it is most usual to "match" evidentiary material to material obtained from the suspect (i.e. the gut pile at the crime scene to the meat in the suspect's freezer). Matching evidentiary and suspect derived samples requires characterizing genetic variability among the species in order to find DNA regions that offer high discriminatory potential. The use of PCR amplifying microsatellite DNA markers and fluorescent detection is a strong combination for such a matching process and thus, assisting law enforcement officials in enforcing wildlife laws. Wildlife animals are being poached and illegally imported at a high rate and presently there is no economical, reliable, sensitive or time-efficient method of providing critical evidence linking evidentiary samples to an individual unless apprehended "red-handed". The application of modern molecular biology techniques to DNA testing will help reduce the chance a criminal will evade conviction.

The objective of this study is to develop a practical, economical, time-efficient DNA typing system for wildlife forensic scientists to utilize for individualizing forensic white-tailed deer evidence. This involves the development of a panel of highly polymorphic microsatellite markers among the Michigan white-tailed deer population. Microsatellite

m

th

se

of

(ca

rep

the

each

poin

requ

order

proba

exhib

multip

ABI 3

markers must show genetic diversity within the population in order to uniquely identify the individual and have high exclusion potential with regards to the evidence. The use of several highly polymorphic markers is necessary to achieve a high degree of probability of a match between the individual and the evidence.

To arrive at this DNA typing system, eight microsatellite markers from bovine (cattle), ovine (sheep) and cervine (deer) origin were used to amplify DNA from a representative sampling of Michigan white-tailed deer to assess genetic diversity among the population. The sampling population consisted of five random deer samples from each of 83 counties and two islands in Michigan. The database thus contains 850 data points per marker providing a large database for probability estimates. These estimates require a determination of the frequency with which alleles occur in the population in order to individualize a sample and match it to evidentiary material with a high degree of probability. Of the eight markers tested, three were found to be highly polymorphic and exhibit high heterozygosity (i.e. > 70 %). In combination, these markers were able to be multiplexed together in one PCR and run simultaneously in one electrophoretic lane of an ABI 377 DNA Sequencer.

HISTORY OF "TYPING" TECHNOLOGY AND LITERATURE REVIEW

Serological Typing Techniques

The beginning of human identification was established by using red blood cell antigen typing systems in order to facilitate disputed paternity cases at the biochemical level, with the following chronology listed in Melvin *et al.* (1988). The ABO blood group system discovered by Karl Landsteiner in 1901 was the start to resolving questioned paternity cases. Rubin Ottenberg in 1921 performed family studies applying the ABO typing system to paternity problems with the knowledge of Mendelian inheritance from von Dungern and Hirschfeld in 1911. Then in 1927, Landsteiner and Levine discovered another blood group system MN and in 1937 Landsteiner and Wiener discovered the Rh system. Several other red cell antigen systems were discovered between 1945 and 1965 given the names Kell, Kidd and Duffy. These systems are performed by immunological testing of allozymes using standard hemagglutination reactions of the antigen on the surface of the red blood cell to an antibody to the antigen. Each of these systems identifies between three and nine different phenotypes.

Electrophoretic separations of red blood cell isoenzymes (e.g. phosphoglucomutase, acid phosphatase, adenylate kinase and adenosine deaminase) and serum proteins (e.g. transferrin, haptoglobin and properdin factor B) through a medium (cellulose acetate, agarose, starch or acrylamide) under a constant applied electric field were developed in 1955 (Melvin *et al.*, 1988). This technology has allowed for the separation of proteins based on their electrical charge properties, and are detected by an enzymatic reaction to the protein or by staining of the protein using colored dyes. Isoelectric focusing

electrophoresis using a continual and linear pH gradient was developed in the 1970's in order to separate proteins with increased resolution and decreased electrophoretic problems. Also developed in the 1970's was the human leukocyte antigen (HLA) system showing a higher degree of polymorphism compared to previously mentioned systems. Genes coding for the HLA antigens reside at four loci and thus, exhibit high variability and show high exclusion potential due to the low frequency of allelic variants in the population (Melvin *et al.*, 1988).

These biochemical methods of differentiating individuals using antigens, proteins and enzymes are very powerful tools in identifying differences between individuals and providing exclusionary information. These techniques are easy to perform and the results can be scored unambiguously. However, they are systems limited in both their allelic variation (which limits the statistical probability of inclusion) and their use of protein products, which invariably degrade over time. Antigen-antibody systems have to be concerned with the intensity of the reaction due to loss of antigen in the sample or weak antigenicity. Proteins and enzymes are not stable molecules and thus are denatured easily therefore null results can be misleading from the testing of samples using these biological components. Contamination from mixed samples and rare cross reactivity between different marker systems may also produce erroneous results. Immunologic assays can not distinguish between all variations in amino acids and are not as sensitive as DNA testing assays. DNA is highly stable, can be isolated from both living and dead cells and codes for the synthesis of these secondary by-products using an exact blueprint for all the biological components necessary for building and maintaining the life of the organism. In addition, every nucleated cell in the body contains the same genetic "blueprint".

Traditional serologic testing has been revolutionized by the use of DNA testing. The use of DNA is by far more preferred because the code itself is being tested and the results have greater discriminatory power.

A Brief History of DNA Advancements

While advances at the protein level were continuing to expand, so were the advances at the DNA (deoxyribonucleic acid) level according to the following chronology listed in Krawczak and Schmidtke (1994).

- In 1944 Avery and co-workers discovered that DNA is the genetic material, the "blueprint of life" and is in every nucleated cell of all living organisms.
- In 1953 Watson and Crick proposed a model for the structure of DNA.
- In 1961 Nirenberg and Matthaei deciphered the genetic code.
- In 1970 Arber, Nathans and Smith discovered restriction enzymes which cut DNA at specific sites.
- In 1972 Berg and co-workers developed molecular cloning of DNA.
- In 1977 Sanger and co-workers developed methods for sequencing DNA.
- In 1979 E.M. Southern developed a method of transferring single-stranded DNA fragments to a more permanent membrane called Southern blotting.
- In 1985 Kary Mullis and co-workers at Cetus discovered the amplification of DNA segments by PCR.

This is only a partial list of the contributions to science since the beginning of DNA discovery in 1944. The applications of these key contributions to the field of molecular biology will be discussed in this chapter, but for now a discussion of DNA is necessary.

DNA Structure

DNA is a long polymer comprised of four subunits, adenine (A), guanine (G), cytosine (C) and thymine (T) found in all nucleated cells. DNA encodes all information necessary for the primary structure of proteins and thus helps direct the necessary life processes of an organism. DNA subunits called nucleotides are classified into two groups, the purines are adenine and guanine and the pyrimidines are cytosine and thymine, with each nucleotide containing a phosphoric acid group, a deoxyribose sugar and a nitrogenous base. Both purines and pyrimidines are heterocyclic, flat, planar molecules with the purines having a double ring structure and the pyrimidines having a single ring structure which allows the bases to stack one on top of each other. Each nucleotide is joined together by their sugar and phosphate groups forming a repeating backbone of sugar-phosphate-sugar-phosphate, etc. along each strand of DNA. Two strands make up the double stranded DNA molecule which are antiparallel to each other and are hydrogen bonded together by adjacent nucleotides on opposite strands. The purine (A) on one strand will always bind to the pyrimidine (T) on the other strand and this is called a base-pair, as with (C) to (G) according to Chargaff's rule in order to maintain a stable DNA molecule. Therefore, each strand is complementary to the other but oriented in opposite directions along the duplex helical molecule. This specificity of base pairing allows for the storage and transfer of genetic information based on the order of nucleotides in each strand of DNA. The DNA sequence specifies the exact genetic instructions required to create a totally unique organism.

D

nu

ge

reg

nec

met

mo

disp

com

nonc

whic

the g

Repe

T

no fur

the rep

Intersp

tandem

of the g

rich reg

Alu I res

(SINE).

DNA Classification

DNA can be put into three classes based on the function and characteristics of the nucleotide sequence. The first and smallest class, comprising only about 3% of the total genome, contains the unique sequences that are rarely repeated, representing the coding regions for genes (Frank, 1997). These sequences carry the genetic information necessary for the synthesis of proteins required for building and maintaining the metabolic processes in an organism. The next class contains sequences, which are moderately repeated at least 1,000 times, are located adjacent to unique DNA and are dispersed throughout the genome. These gene-related sequences for the RNA components of ribosomes, such as tRNA and rRNA, along with histones are considered noncoding DNA (Lindquester, 1997). The last class contains highly repetitive sequences, which are repeated thousands to even millions of times and are found in certain regions of the genome.

Repetitive DNA

This repetitive or noncoding DNA, often referred to as 'junk' DNA because it serves no function in protein synthesis, can be further classified into groups based on whether the repeat sequences are situated in tandem or interspersed within the genome. Interspersed elements are single units dispersed throughout the genome, usually not in tandem, but occur hundreds to thousands of times within the untranslated intronic regions of the genome. The most abundant sequence in the human genome associated with GC rich regions is called an Alu repeat element because the sequence contains the enzyme *Alu* I restriction site (AGCT) and is considered a short interspersed nuclear element (SINE), less than 500 bp long (Kobilinsky, 1993). The L1 or Kpn repeat element is

associated with AT rich regions because the sequence contains the enzyme *Kpn* I restriction site (GGTACC) and is considered a long interspersed nuclear element (LINE), greater than 500 bp long (Kobilinsky, 1993).

Tandem Repetitive DNA

Tandemly repeated or clustered sequences contain core sequences repeated numerous times and arranged side by side thus, in tandem. These core sequences, usually between two and six bases long (Tautz, 1989), repeated between ten and a hundred times, scattered throughout the genome millions of times, but localized to the centromeric region of the heterochromatin of chromosomes are called microsatellites or short tandem repeats (STR's). Minisatellites, often associated with the term VNTR, have core sequences ranging from 10 to 50 bases long, repeated two to several hundred times and are mainly clustered in the proterminal or telomeric regions of chromosomes, but have also been found in interstitial regions (Royle *et al.*, 1988).

Minisatellite Description

Minisatellite sequences can vary in both the sequence of nucleotides in the repeat and the number of repeat units and thus, with variable numbers of tandem repeats are called VNTR's. These hypervariable regions are excellent genetic markers due to their high polymorphic nature and their discriminating power for characterizing DNA. Minisatellites are thought to arise from either unequal exchange during mitosis between sister chromatids or during meiosis between homologous chromosomes (Jeffreys *et al.*, 1985a). These minisatellite sequences may be responsible for promoting recombination events similar to the *chi* sequence (GGGCAGGAXG) in *Escherichia coli* (Jeffreys *et al.*, 1985a). The GC rich core sequence of minisatellites are considered to be a recombination

hotspot in human DNA (Jeffreys *et al.*, 1985a). VNTR's can be located at one locus or at several loci thus, given the names single-locus VNTR's and multi-locus VNTR's, respectively, and these genetic markers are identified by restriction fragment length polymorphisms (RFLP).

RFLP Description

RFLP's are variations in the length of DNA fragments as a result of a restriction enzyme cleaving the genomic DNA at a certain sequence recognition site, usually four, six or eight bases long. Variation is either due to a single nucleotide substitution creating or eliminating a cleavage site for a specific restriction endonuclease or a rearrangement of a DNA segment accounting for the variable length fragment. RFLP's are inherited dominantly in a Mendelian fashion, one allele from each parent, and are only dimorphic, that is, showing the presence or absence of a restriction site detected by length polymorphisms.

Single locus VNTR's were first analyzed by using RFLP's adjacent to the beta-globin locus in human DNA associated with sickle cell anemia (Kan and Dozy, 1978). Single locus VNTR's create only one or two bands depending on the individual's zygosity, which allows for unambiguous interpretation of the results. Single locus VNTR results, however, are not very informative because each individual will have only one or two different bands. The first highly polymorphic locus was identified in a nonspecific gene in human DNA showing high variability with at least eight bands (Wyman and White, 1980). Using a multi-locus VNTR, present at multiple loci in the genome, scattered throughout the telomeric ends of chromosomes, creates an individual-specific DNA 'fingerprint' resulting in multiple bands which are unique to an individual (Jeffreys

et al., 1985a). The work done by Alec Jeffreys initiated interest by forensic scientists to analyze forensic specimens using RFLP DNA typing. The use of several single-locus VNTR's (Nakamura *et al.*, 1987) simultaneously will also create a unique DNA 'fingerprint' for identification purposes however, interpretation of the resulting pattern may be problematic due to possible band sharing between loci. Band sharing occurs when the size of bands from one locus overlap at the same size as bands from another locus, thus, creating ambiguous interpretation of the results.

RFLP Procedure

The RFLP DNA typing technique is an extensive multi-step procedure which results in an autoradiograph or fluorograph containing a unique DNA fragment pattern (except in identical twins) when several DNA probes are used or a multi-locus probe is used. The beginning of this procedure requires high quality and quantity, at least 5-10 μg of genomic DNA. The isolation and purification of the DNA can be performed using standard organic procedures (Sambrook *et al.*, 1989) or by using commercially available nonhazardous kits resulting in high molecular weight DNA, relatively free of proteins and not sheared into small pieces. The DNA is then digested with restriction enzyme(s) (RE) selected based on the frequency with which they cut the DNA and the sequence of their recognition site, so as to not cut within the probing sequence.

The RE digested DNA is electrophoresed overnight, applying low voltage, through a low percentage agarose gel in order to separate the fragments based on size. The double-stranded DNA fragments in the gel are then denatured to become single-stranded by soaking the gel in an alkaline denaturing solution. It is necessary to transfer single-

strand

hybri

know

transf

ultrav

the m

single

probe

compl

condit

L

the siz

bands

determ

immig

cell an

v Andr

VNTR

Michig

DNA F

Fo

RFLP

descrit

stranded fragments to a sturdier membrane, usually nylon or nitrocellulose, so that hybridization of the single-stranded probe can be achieved. This transfer process is known as Southern blotting (Southern, 1975) and requires capillary action to perform the transfer. The membrane is then baked in an oven (nitrocellulose) or exposed to ultraviolet light (nylon) in order to cross-link or permanently fix the DNA fragments onto the membrane. Hybridization of the membrane-bound denatured DNA fragments to the single-stranded complementary sequence of the radioactive or chemiluminescent labeled probe is performed under optimal conditions for specific binding of the probe to the complementary sequence. Unbound probe is washed away under varying stringency conditions.

Lastly, the membrane bound with the labeled probe is exposed to x-ray film to detect the size of fragments showing complementarity to the probe. The resulting pattern of bands created by this RFLP DNA typing technique has proved useful for comparison in determination of paternity (Jeffreys *et al.*, 1985b; Wells *et al.*, 1989), in solving immigration cases (Jeffreys *et al.*, 1985c), in diagnosing medical diseases such as sickle cell anemia (Chang and Kan, 1982) and directly assisting in the first criminal case (State v Andrews, 1987) resulting in a conviction based on VNTR DNA testing evidence. VNTR was the first DNA evidence technique accepted by most courts, including those in Michigan.

DNA Fingerprinting in Forensic Casework

Forensic casework has gone through a major transition from traditional serology to RFLP DNA testing since the work done by Alec Jeffreys and coworkers in 1985 describing 'DNA fingerprinting' and its potential use in forensic science. Lifecodes

Co

cas

the

the

tech

Stat

area

Gui

the s

ques

guid

case

guide

court

the sc

the U

for for

assura

DNA

laborat

laborat

these gu

Company in 1986 and Cellmark Diagnostics in 1987 both took interest in accepting casework by performing RFLP technology in their laboratories (Weedn, 1993). In 1988 the Federal Bureau of Investigation (FBI) started using RFLP for casework, along with the first state crime laboratory in Quantico, Virginia in 1989 (Weedn, 1993). This new technology has sparked interest from crime laboratories in every state in the United States, along with other private companies investing into this technologically advanced area.

Guidelines for DNA Testing

With any new technology, the reviews and criticisms are always there to follow from the scientific community, criminal justice professionals and the average intelligent person questioning the reliability and accuracy of this technique and any other DNA testing. The guidelines for the admissibility of scientific evidence were established in 1923 with the case United States v. Frye, also called the "Frye" hearings (Baird, 1992). These guidelines are used in determining whether DNA evidence should be admissible in a court of law based on the premise that there is general acceptance of the evidence from the scientific community. In 1990, the Office of Technology Assessment which is part of the United States Congress concluded that DNA evidence is reliable and can be utilized for forensic casework only if appropriate technology, quality control and quality assurance procedures are implemented (Weedn, 1993). The Technical Working Group on DNA Analysis Methods (TWGDAM) represented by scientists from North American laboratories developed DNA methodology and quality assurance guidelines for laboratories to follow for forensic RFLP typing (TWGDAM, 1989). Included within these guidelines is the application of external proficiency testing by a reputable laboratory

or a

and

in o

Res

TW

Cou

influ

Simi

have

and P

These

appro

potent

(Natio

PCR

Th

Mullis

replacin

enzyma

DNA in

expone

to detec

or hybr

or accrediting agency (TWGDAM, 1990), such as the College of American Pathologists and the American Society of Crime Laboratory Directors Laboratory Accreditation Board in order to assure quality performance of the laboratory (Weedn, 1993). The National Research Council in 1992 issued a report stressing the guidelines developed by the TWGDAM regarding standardized laboratory procedures in 1990 (National Research Council, 1992). The American Society of Crime Laboratory Directors has also been influential in making recommendations regarding DNA technology in forensic science. Similar guidelines, quality assurance and accreditation programs for PCR technology have also been established by the TWGDAM. Revisions to these guidelines for RFLP and PCR DNA testing will be necessary as the technologies and experience advances. These methodologies are now widely accepted in the criminal justice community when appropriate quality control methods are followed however, current controversy focuses on potential human and technical errors and on the statistical interpretation of results (National Research Council, 1996; Weir, 1996).

PCR Description

The amplification of specific DNA segments by PCR discovered in 1985 by Kary Mullis and co-workers at Cetus has revolutionized molecular biology and is rapidly replacing RFLP DNA typing for identification of forensic evidence. This *in vitro* enzymatic amplification of a DNA segment is performed utilizing minute quantities of DNA in repetitive cycles of denaturation, hybridization and extension creating an exponential amount of target DNA. This amplified DNA can then be directly sequenced to detect single base polymorphisms, electrophoresed to detect sequence length variations or hybridized to allele-specific probes to detect sequence polymorphisms. Forensic

scientists have benefited tremendously by this new technology because the quantity and quality of recoverable DNA found at crime scenes is often less than optimal.

PCR technology does not require high molecular weight DNA or large quantities of DNA like RFLP technology. PCR can be performed with only a few copies of a target DNA sequence as long as the target sequence to be amplified is not degraded or exists as low molecular weight DNA. Degradation of DNA for RFLP analysis limits the availability of restriction sites for the enzyme being used which in turn affects the interpretation of the results. PCR can be performed with degraded DNA because 1 μg of genomic mammalian DNA corresponds to approximately 3×10^5 copies of autosomal genes (Cha and Thilly, 1993). The chance that each copy shows degradation in the same region to be amplified is very small. Thus, in any sample, there is likely to be several sections that are suitable for amplification. Most PCR applications only require nanogram quantities of DNA per reaction as compared to RFLP requiring several micrograms. PCR amplification is sensitive to interfering polymerase inhibitors found in DNA samples extracted from materials containing forensic evidence such as detergents, heme, melanin pigments, dyestuffs, sodium acetate, metal cations, urea or EDTA (ethylenedinitrilo tetraacetic acid) (Sensabaugh and Blake, 1993).

The sensitivity of PCR to contamination is of major concern because only minute amounts of extraneous human DNA may cause erroneous results. Contamination often comes from mixed samples at the crime scene, which is a reality for forensic scientists to deal with. Another source is from laboratory personnel or crime scene technicians and investigators introducing their own DNA into the evidence. This can be a problem if the

evide

conta

prior

requi

proce

conta

PCR

T

the sp

the su

Other

PCR (

amplif

thermo

has the

amoun

include

primers

(dNTP

Th

double

to thre

the sy

evidence sample contains a very small amount of DNA. The last major source of contamination arises when other amplified samples are accidentally mixed with evidence prior to amplification. This can be a serious problem because only an aerosol droplet is required to cause preferential amplification of the contaminant. Careful laboratory procedures need to be adopted to minimize these sources of contamination. Many contamination errors are detectable and can be corrected by reanalyzing the evidence.

PCR Procedure

The PCR procedure, first developed as a technique by Saiki *et al.* in 1985 to amplify the specific beta-globin sequence, is a cyclic process that does nothing more than increase the subanalytical quantities of DNA to a level that can be detected by routine methods. Other studies quickly followed using enzymatic amplification of DNA *in vitro* by the PCR (Mullis *et al.*, 1986; Saiki *et al.*, 1986; Mullis and Faloona, 1987). The amplification procedure is a relatively simple laboratory technique to perform with a thermocycler doing most the work. The thermocycler is a programmable machine, which has the ability to cycle through various temperatures (0-100° C) in a relatively short amount of time with temperature homogeneity and accuracy. The sample preparation includes double-stranded DNA to be amplified, two single-stranded oligonucleotide primers flanking this DNA segment, a DNA polymerase, deoxynucleotide triphosphates (dNTP's), a buffer, magnesium chloride ($MgCl_2$), salts and deionized water.

The three step cyclic procedure begins with the denaturation step to separate the double-stranded DNA molecules into single strands by heating them to 94-95° C for one to three minutes. This creates two strands of DNA, which are both used as templates for the synthesis of complementary strands of DNA. The temperature is then lowered to

allo

and

sequ

of th

sequ

temp

temp

doub

nucle

optim

range

T

order

cycle.

and af

molecu

Figure

becaus

the len

produc

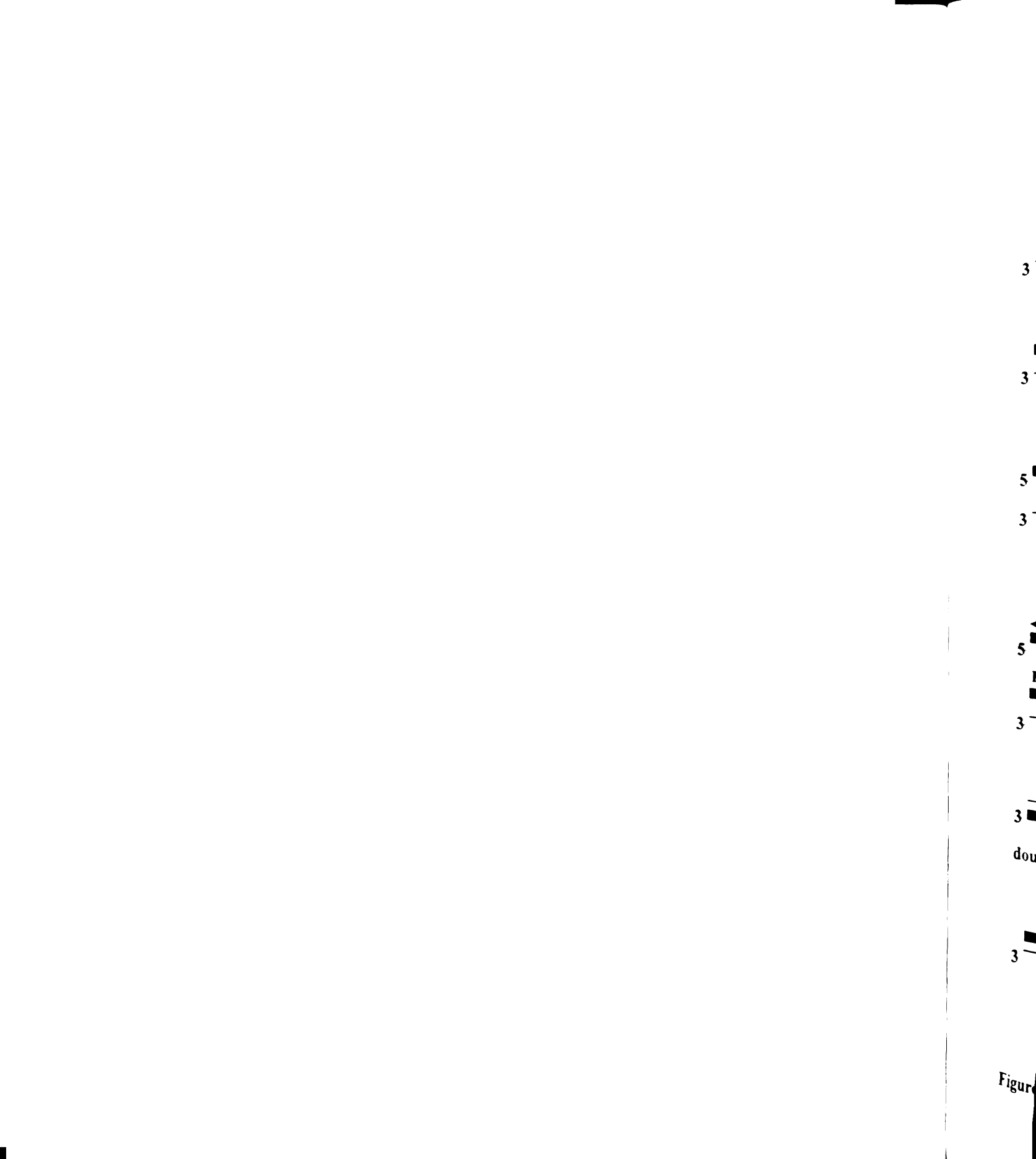
polyme

produc

both p

allow the oligonucleotide primers (stretches of up to twenty nucleotides in length), to find and hybridize to their complementary sequences on opposite strands flanking the sequence to be amplified, called the annealing step. The last step involves the extension of the primers at the 3′ end by the addition of nucleotides complementary to the target sequence mediated by a DNA polymerase, thus called the extension step. The temperature of the annealing step will vary between 45-65° C depending on the melting temperature T_m (the temperature at which half of the primer is single-stranded and half is double-stranded) of the oligonucleotide primers, which is directly related to their nucleotide length and content. The temperature of the extension step is dependent on the optimal temperature for the activity of the DNA polymerase being used, usually in the range of 70-75° C.

This three-step amplification process is cycled several times, usually 25 to 30 in order to generate millions of copies of the target sequence. After the first round of a cycle, one double-stranded molecule has been doubled to two double-stranded molecules and after the second round of a cycle, the two molecules have been doubled to four molecules and so on, an exponential (2^n) fold increase, for each additional cycle (see Figure 1). The number of target sequence has in effect been doubled after each cycle because each strand serves as a template for replication in subsequent cycles. However, the length of the products generated after the first cycle are longer than the desired product size because the polymerase extends until the denaturation step forces the polymerase to separate from the DNA template. In subsequent cycles the short or desired product size increases exponentially because the ends of the product have been defined by both primers, whereas the original double-stranded molecule only increases linearly.



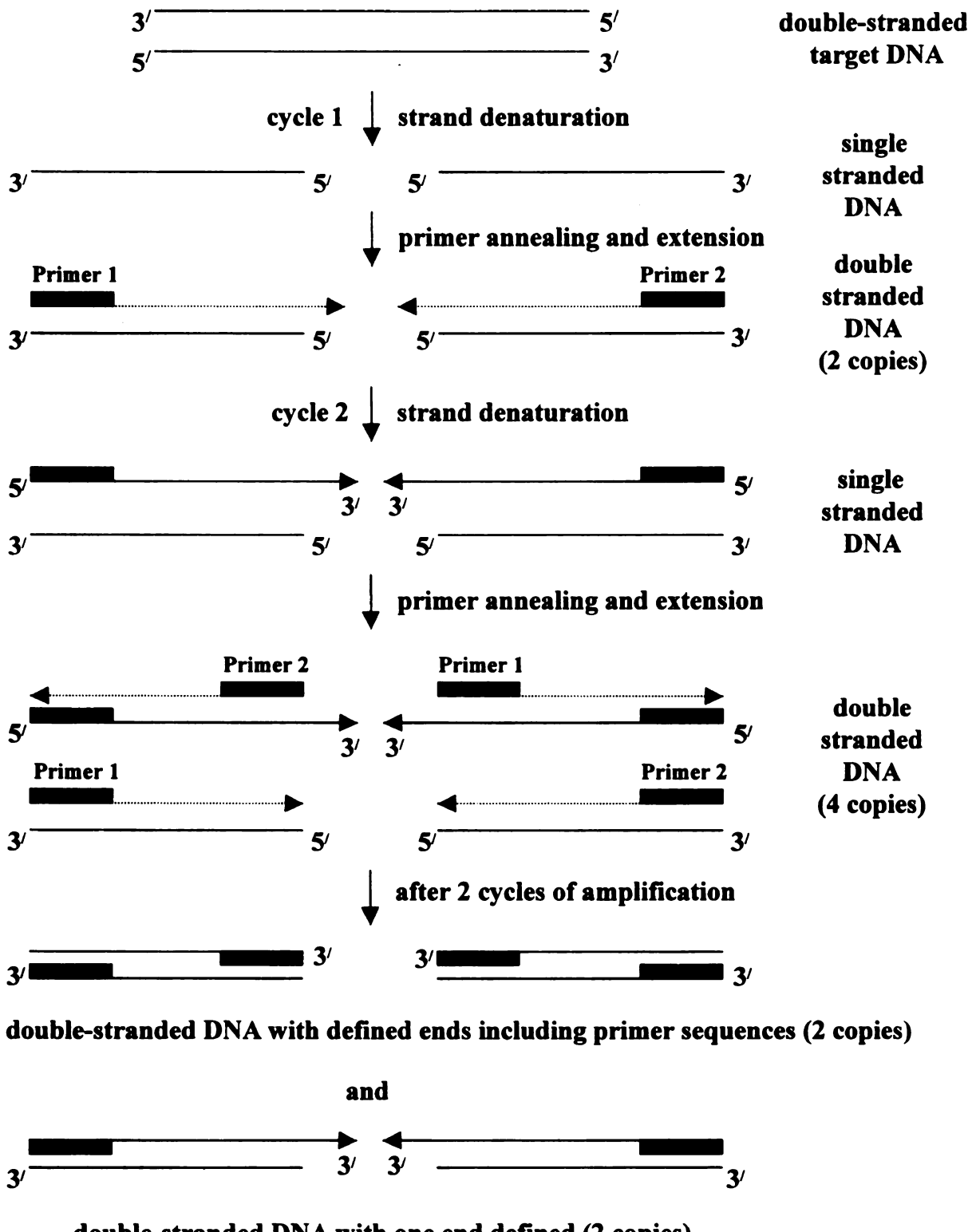


Figure 1. Schematic diagram of polymerase chain reaction (PCR) amplification.

P

minute

optimi

amplif

thermo

the am

every

of MgC

amplif

individ

PCR

Th

with th

temper

initial t

content

primer

used. f

of time

amplif

extens

time f

PCR, with its specificity to the desired target sequence and extreme sensitivity to minute quantities of target DNA, is a technique that requires many components to be optimized in order to obtain the desired product and only the desired product. The amplification conditions, containing both the components in the reaction and the thermocycling temperatures, times and number of cycles, are variables that directly affect the amplification process. These variables need to be optimized for almost every locus studied, unless similarity exists between marker type, primer T_m and amount of $MgCl_2$ required for each reaction. Therefore, much time is spent optimizing the amplification conditions prior to validating a marker system for use in population studies, individual identification or other PCR applications.

PCR Thermocycler Programming

The temperature and amount of time selected for the thermocycling conditions, along with the number of cycles are dependent on the following criteria. The denaturation temperature should be close to boiling in order to completely separate the strands with an initial time of a few minutes. The annealing temperature depends on the length and GC content of the primers and, in general, should be around five degrees below the T_m of the primer pair. The extension temperature should be selected based on the polymerase being used, for *Thermus aquaticus* (*Taq*) the optimal temperature would be 72° C. The amount of time for this step is based on the length of the target DNA being amplified. That is, amplifying a fragment greater than 1Kb (kilobase) in length would require more extension time for the polymerase to complete the full sequence, otherwise a standard time for smaller fragments would be 30 seconds to 1 minute for each step. The number

of cycles depends on the amount of product generated based on the amount of starting template, which can vary from as little as 20 cycles to as many as 35 or even 40 cycles.

PCR Components

The components of a standard PCR are contained in a reaction volume of 25, 50 or 100 µl with 50 to 100 ng of genomic DNA mixed with 50 mM KCl, 10 mM Tris HCl pH 8.4, 0.25 µM of each primer, 200 µM of each dNTP (deoxynucleoside triphosphate), 1.5 mM $MgCl_2$, 2.5 units of *Taq* polymerase, deionized water and overlayed with a few drops of mineral oil (Saiki, 1989). The mineral oil is added to reactions when the thermocycler being used does not have the heated lid feature to prevent evaporation of the reaction mixture and to increase the rate of thermal equilibration. These are standard conditions that may not provide adequate results in all situations. Modifications may be necessary to obtain the desired specificity and yield of PCR product.

The component with the most dramatic effect on specificity and yield is the amount of $MgCl_2$ in the reaction, that is, the amount of free Mg++ available for *Taq* polymerase. The effect of dNTP's chelating magnesium will lower the optimal concentration of magnesium. If all the Mg++ is bound by the dNTP's, then none will be present to facilitate *Taq* in amplifying the target DNA. The absence of a PCR product can often be corrected by adding a higher concentration of $MgCl_2$ to the reaction, while the presence of too many products is the result of too high of $MgCl_2$. Optimal levels of free Mg++ should be around 1.0 mM for most PCR's. Concentrations of Mg++ approaching 10 mM $MgCl_2$ inhibit the activity of *Taq* (Gelfand, 1989).

Because they act as building blocks for the amplification product, the amount of dNTP's should always be in excess, 200 μM, of DNA target sequences for efficient amplification but, it is not recommended to increase this amount because it will chelate Mg++ as mentioned above. However, if amplifying a large target sequence, greater than 1 Kb, a higher concentration of dNTP's is necessary. *Taq* polymerase is actually inhibited in the presence of millimolar concentrations of dNTP's (Gelfand, 1989) and these high dNTP concentrations may actually promote misincorporations.

The selection of oligonucleotide primers is of paramount importance to the success of an amplification reaction. The criteria for primer design entails creating both primers which are between 20 and 30 bases in length with a GC content between 40 and 50 percent and with a random base distribution not containing stretches of similar bases. A 20mer has a specificity of $1/4^{20}$ in locating the exact complementary sequence in the template DNA and with a 50 percent GC content has a T_m range of 56-62° C (Cha and Thilly, 1993). Each additional nucleotide will increase the specificity of the primer by a factor of four and increase the T_m by two or four degrees depending on which base is added. Primers should not be able to form secondary structures within themselves because then they will not be able to hybridize to the template DNA. They also should not have complementarity to each other in order to avoid the formation of primer dimers. This is of major concern when using multiple primer pairs in one reaction (multiplex). Perfect complementarity should exist at the 3′ end of the primer, usually the presence of a G or C at the very end will assure primer extension by the polymerase because these bases form stronger bonds and will anchor the primer to the template. Lastly, each

primer should be in a ten-fold excess of the target sequence during PCR (Cha and Thilly, 1993).

The specificity, sensitivity, yield and length of PCR products being amplified has been substantially improved with the replacement of the Klenow fragment of *E. coli* DNA polymerase I with an enzyme isolated from the thermophilic bacterium *Thermus aquaticus* (*Taq*) (Saiki *et al.*, 1988). Prior to this discovery, fresh Klenow enzyme was added to the reaction mixture after each cycle of denaturation because this enzyme is heat-sensitive and would have no activity after one round of amplification. This sequential addition of polymerase to the reaction mixture has been eliminated with the discovery of *Taq*, which is very thermostable with repeated exposure to temperatures of 95° C. This allows the PCR cycling to be fully automated using thermal cycling machines and increases the yield of product because the enzyme still maintains activity after numerous cycles. The processivity of *Taq* polymerase during DNA synthesis allows for the amplification of longer DNA segments, up to 10 Kb (Jeffreys *et al.*, 1988), which adds to the list of PCR applications. The specificity and sensitivity has been greatly enhanced by the use of *Taq* polymerase because at higher temperatures, the annealing of complementary primers to the template DNA will result in only amplification of the desired target sequence with an increase in the yield of that product, eliminating nonspecific amplification.

The fidelity of *Taq* polymerase in PCR amplification is not a strong feature with a relatively high error rate of 10^{-5} nucleotides/cycle (Gelfand and White, 1990) compared to other polymerases. This misincorporation of nucleotides is due to the lack of the 'proofreading' 3′ to 5′ exonuclease activity of *Taq* polymerase. Misincorporation of a

nucle

temp

PCR

errors

becau

detec

fewer

probl

mutat

polym

nucle

specif

gener

Sourc

P

preser

additi

produ

separa

Both

micro

deal w

nucleotide would have to be during the early stages of PCR, with very limited amount of template, affecting many copies, in order to have a significant effect on the amount of PCR products detectable by electrophoretically visible bands (Saiki *et al.*, 1988). The errors made by *Taq* polymerase should not be a problem for many PCR applications because the analysis of several samples to establish a consensus sequence is performed to detect any misincorporated nucleotides. Using more initial template DNA molecules, fewer cycles of amplification and modifying the reaction conditions can reduce this problem exhibited by *Taq* polymerase. Other applications, such as cloning or point mutation analysis where fidelity is of concern, may require the use of the *Pfu* DNA polymerase isolated from *Pyrococcus furiosus* with an error rate of 6.5×10^{-7} nucleotide/cycle (Andre *et al.*, 1997). However, the thermostability, processivity and specificity of *Taq* polymerase has made it the most utilized polymerase in PCR to generate large quantities of a specific target sequence.

Sources of Error During Amplification of Microsatellites

PCR amplification of dinucleotide microsatellite markers using *Taq* polymerase presents two sources or error in genotyping studies. The catalysis of a nontemplated addition of a nucleotide, predominantly adenosine, to the 3′ end of the amplified PCR product (Clark, 1988) and the appearance of shadow bands with the PCR products separated by intervals of two nucleotides (Hauge and Litt, 1993; Murray *et al.*, 1993). Both of these can cause problems when genotyping the sizes of dinucleotide microsatellite fragments generated by PCR. However, strategies have been proposed to deal with these biological occurrences.

The addition of the nontemplated nucleotide to the PCR product creates bands which differ in size by a single base and is primer specific (Smith *et al.*, 1995). The degree of addition varies from marker to marker with some markers not affected at all and some only partially affected with both product sizes being present in similar quantities. The known fact that this situation occurs with some markers, PCR based strategies (Smith *et al.*, 1995) and primer modifications (Brownstein *et al.*, 1996) strategies to minimize this source of error have been addressed. Changing the PCR thermal cycling conditions can significantly reduce the error rate of inconsistent allele calling. To minimize the production of allele + A (adenine) products, a two step cycling protocol can be used which eliminates the extension step that causes the polymerase to add the extra A. To favor production of the allele + A product, a three step cycling protocol can be used with an additional final extension step facilitating the nontemplated nucleotide addition. Modifying the reverse primer on the 5′ end by adding the sequence GTTTCTT can facilitate adenylation of the 3′ end of the forward primer, thus generating products which are all allele + A (Brownstein *et al.*, 1996). Other methods employing an additional enzymatic step after amplification by adding T4 DNA polymerase with a 3′ to 5′ exonuclease activity to remove the unpaired base have been proposed (Kimpton *et al.*, 1993).

The characteristic feature of *Taq* polymerase is to frequently skip or occasionally add repeat units during the extension step of PCR amplification causing great difficulty in analyzing dinucleotide microsatellite markers. Other short tandem repeats (tri-, tetra- or pentanucleotides) exhibit this skipping by *Taq* polymerase very infrequently. This slipped strand mispairing by *Taq* is the major mechanism for the generation of shadow

bands.

(Haug

itself f

shorter

et al. i

band a

the act

amplif

unit. I

thermo

Multi

T

multip

PCR,

et al.,

comb

loci o

seque

react

The p

overl

locus

bands, which are less intense and intervals of two nucleotides short of the correct allele (Hauge and Litt, 1993). This slippage by *Taq* results from the temporary dissociation of itself from the template strand allowing repeats in the template to form a loop, which shortens the PCR product by the number of repeat units in the loop. The study by Murray *et al.* in 1993 also supports this mechanism because they sequenced the actual shadow band and found it to contain the flanking sequences on both sides of the repeat, but that the actual repeat sequence was ambiguous. This suggests that the shadow band is the amplified product with the repeat sequence being scrambled, but still short by one repeat unit. Presently, the only correction to the occurrence of these shadow bands is to use a thermostable DNA polymerase with a higher processivity than *Taq*.

Multiplex PCR

The ability of PCR to be extended even further to the simultaneous amplification of multiple loci using the same DNA template in a single PCR reaction, called multiplex PCR, has greatly increased the throughput of data in considerably less time (Chamberlain *et al.*, 1988). However, more effort is required to develop optimal conditions for the combination of primer pairs being amplified simultaneously. Multiplexing two or more loci often can be performed under the same reaction conditions used to amplify the sequences separately. The main consideration is that the components of the amplification reactions and the temperature cycling conditions are the same for all the loci involved. The primers must not show homology to each other and the product sizes should not overlap each other. All the factors discussed previously for a single amplification of one locus must also be considered when optimizing a multiplex PCR. The adjustment of

these

availa

Mult

T

syste

(Kim

Walla

micro

with n

many

semi-a

O

fluoro

respec

amplif

of a D

specif

length

T

every

elimi-

across

preser

these variables is specific to the unique loci being amplified and general guidelines are available to assist the scientist in optimizing the multiplex PCR (Henegariu *et al.*, 1997).

Multiplex PCR and Fluorescent Detection

The combination of multiplex PCR and semi-automated fluorescent detection systems has become a rapid and powerful technique for individual identification (Kimpton *et al.*, 1993). The amplification of loci consisting of two alleles (Skolnick and Wallace, 1988) has been extended to the amplification of highly polymorphic microsatellite loci and genotyping these loci using automated DNA sizing technology with multi-color fluorescent detection (Ziegle *et al.*, 1992). This ability to coamplify many loci in one PCR has been facilitated by the use of fluorescently labeled primers and semi-automated fluorescent DNA sequencers.

Oligonucleotide primers are labeled at their 5′ end with one of three possible fluorochrome dye molecules: Fam, Hex or Tet, fluorescing blue, yellow and green, respectively. The labeling of the primers at the 5′ end does not effect the PCR amplification, but it does allow for the simultaneous analysis of multiple loci in one lane of a DNA sequencing gel. The fluorescent dye is only on one primer of each locus-specific primer pair, since both primers are amplifying the same target sequence the length will be the same.

The use of an internal size standard, labeled with the red fluorescent dye TAMRA, in every gel lane is referenced when calling the fragment sizes. The benefit of this is to eliminate interpretation errors due to electrophoretic mobility differences between lanes across the gel and thus, increasing the size calling accuracy of the alleles. Fragments present in each lane are automatically sized against the internal size standard comigrating

in each lane, which compensates for any gel distortions in that lane. Size standard TAMRA 350 from ABI, for example, contains known fragments with sizes: 50, 75, 100, 139, 150, 160, 200, 250, 300, 340 and 350, representing fragment lengths in nucleotide base pairs.

Fluorescence-based detection of microsatellite PCR products compared with conventional autoradiographic methods (Schwengel *et al.*, 1994) and silver staining (Lins *et al.*, 1996; Budowle *et al.*, 1997) for identification of individuals has resulted in many advantages. Fluorescent compounds are safer to use, easier and cheaper to dispose of and have a longer shelf life than radioisotopes. They are extremely more sensitive with the ability to detect picomole quantities of fluorescently labeled primers and picogram quantities of initial template DNA. Fluorescent signals are linear over a wider range of intensities, where signal strengths of multiple loci greatly vary in magnitude (Schwengel *et al.*, 1994). Multiple loci can be simultaneously analyzed in a single electrophoretic lane, whereas with radioisotopes and silver staining this variation in signal intensity cannot be tolerated. The signal intensity of the peaks can be used to estimate the amount of PCR product present. Scoring alleles of a heterozygous individual being genotyped for a dinucleotide microsatellite marker with alleles differing by one repeat unit is much easier with fluorescence. The intensity of the smaller allele overlapping the stutter band of the larger allele will be greater than the intensity of the larger allele (see Figure 2).

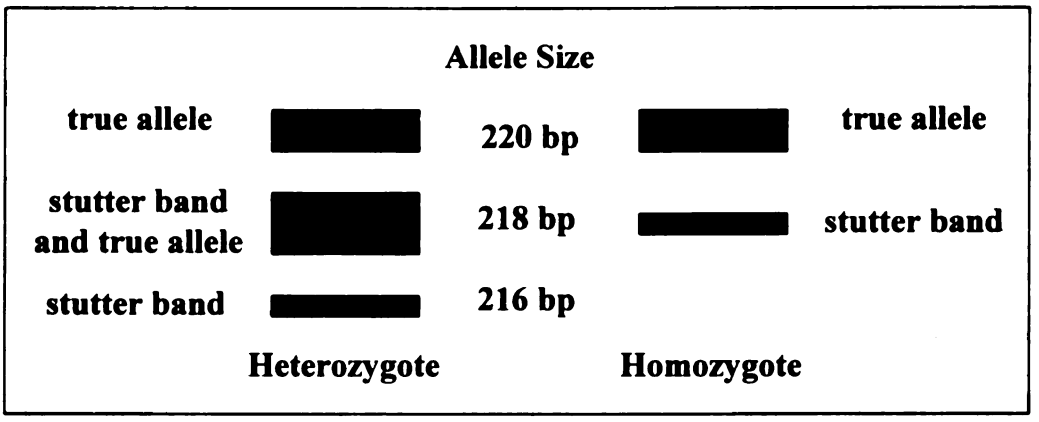


Figure 2. Schematic diagram representing the fluorescent intensity of alleles differing by one repeat unit of a dinucleotide microsatellite marker.

The time required to score alleles and the number of genotyping errors drop considerably due to the combination of automated fluorescence-based electrophoretic systems and genotyping analysis software. Genotyping analysis is being facilitated by the use of internal size standards and analysis software for the accurate assignment of fragment sizes and the unambiguous scoring of alleles. Ghosh *et al.* in 1997 describe methods for accurate sizing of alleles in order to reduce genotyping error rates. Manual methods of scoring autoradiographs or gels stained with silver lead to more typing and sizing errors because both DNA strands are being detected. These methods are extremely time-consuming and are not compatible with high throughput genotyping.

Several validation studies, for use in forensic applications, have been performed utilizing multiplex PCR of short tandem repeats with fluorescent detection (Fregeau *et al.*, 1999; Budowle *et al.*, 1997). These studies support the use and benefit of this genetic typing system for forensic identification and according to Fregeau *et al.*, could provide discriminatory power of approximately 0.9999. This discriminatory potential of fluorescent multiplex STR marker systems in human identification is the basis for applying this technology to individualize evidence samples in forensic wildlife cases.

Multiplex PCR using microsatellite markers for characterizing genetic variability among Michigan white-tailed deer, for the purpose of individualizing evidence samples, should be both cost-effective and beneficial. The homology between microsatellite loci is often conserved among related species (Moore *et al.*, 1991), saving time and effort in primer development. Bovine microsatellites are highly conserved among Red deer (Kuhn *et al.*, 1996) and Sika deer (Slate *et al.*, 1998) of the cervine family. White-tailed deer microsatellites are conserved in bovids (DeWoody *et al.*, 1995), microsatellites are conserved in bovine, ovine and caprine (Kemp *et al.*, 1995) and also conserved between other species of artiodactyls (Engel *et al.*, 1996). The selection of primer pairs for this study then focused on known primers that amplify bovine, ovine and cervine DNA.

Dot Blot Procedure

The first PCR based genetic typing kit of the HLA-DQ alpha locus using dot blot methodology for detection of sequence polymorphisms was in 1986 (Saiki *et al.*, 1986), with the first use in a criminal case in Pennsylvania in 1986, *Pennsylvania v Pestinikis* (Blake *et al.*, 1992). PCR amplification of the HLA-DQ alpha locus and detection of point mutations in the different allelic forms using dot blot procedures and allele-specific oligonucleotide (ASO) probes (Saiki *et al.*, 1986; Erlich *et al.*, 1986) has made the transition from RFLP to PCR a simpler technology for detecting polymorphisms. Hybridization, under highly stringent conditions, of a complementary DNA probe to the amplified sequence bound to the membrane is indicated by a color reaction as a spot on the membrane for that particular allele. The presence of a spot is indicative of perfect complementarity in DNA sequence between probe and target DNA thus, identifying the allele present. Reverse dot blot procedure was developed to simultaneously hybridize the

DNA

proce

al.. 19

preser

RAPI

A

ampli

This t

oligor

70 pe

to opp

regio

be am

are ur

separ

polyn

(Wel

and s

seque

al., 1

for u

dinuc

DNA to the immobilized ASO probes on the nylon membranes (Saiki *et al.*, 1989). This procedure identifies six different alleles and twenty-one possible genotypes (Helmuth *et al.*, 1990). This method is simple and very time efficient because it will detect the allele present in a single hybridization step by a simple colorimetric reaction.

RAPD Technique

Another segregating dominant genetic marker amplified by PCR is called random amplified polymorphic DNA (RAPD) and was developed by Williams *et al.* in 1990. This technique involves the amplification of genomic DNA using an arbitrary oligonucleotide primer, usually nine or ten bases long with a GC content between 50 and 70 percent. This primer binds to homologous sites in genomic DNA and when they bind to opposite DNA strands, which are relatively close to each other, amplification of the region between them will occur. Several sites, randomly distributed in the genome, will be amplified by the oligonucleotide primers thus, creating discrete DNA products which are unable to distinguish a heterozygote from a homozygote. PCR fragments are separated by gel electrophoresis and detected by one of several methods to identify polymorphisms between individuals and generate a genomic profile of PCR products (Welsh and McClelland, 1990). RAPD's were used to detect genetic diversity in cattle and sheep (Kantanen *et al.*, 1995) and these species show homology with cervids. DNA sequences of RAPD fragments showed high sequence similarity in artiodactyls (Kostia *et al.*, 1996). This allows for characterization of the fragments from one species to another for use in genome mapping of these markers in closely related species, similar to dinucleotide microsatellite homology among closely related species (Moore *et al.*, 1991).

SSC

and

dena

subs

sam

elec

stan

prod

elect

diffe

confo

(Orit

Micr

by PC

genor

simpl

genor

mamr

tande

$(CA)_n$

Webe

SSCP Technique

Single-stranded conformational polymorphisms (SSCP) are also amplified by PCR and are variations in the mobility of single-stranded DNA fragments under non-denaturing conditions. This method of detecting sequence changes, including single base substitutions, is based on the fact that DNA fragments with different sequences but the same length will migrate differently during non-denaturing polyacrylamide gel electrophoresis (Orita *et al.*, 1989a). DNA, in an area of interest, is amplified under standard PCR conditions using 5′ end labeled oligonucleotide primers synthesizing products in the range of 150-250 base pairs (Tuggle, 1994). Fragments generated are electrophoresed and polymorphisms are detected between individuals based on differences in band migration. Single-stranded DNA fragments form folded conformations that are sequence specific and are stabilized by intrastrand interactions (Orita *et al.*, 1989b).

Microsatellite Description

Microsatellites are excellent codominant genetic markers, amenable to amplification by PCR, due to their high variability, high abundance and widespread distribution in the genome (Weber and May, 1989). Microsatellites are regions of DNA consisting of simple sequence motifs repeated in tandem and abundantly scattered throughout the genome. They are found in many eukaryotic genomes (Tautz and Renz, 1984) and all mammalian genomes so far examined (Stallings *et al.*, 1991). These motifs can be tandem arrays of di-, tri-, tetra-, or pentanucleotides, with dinucleotide motifs such as the $(CA)_n$ repeat being the most abundant in humans (Hamada *et al.*, 1982; Beckmann and Weber, 1992) and are also found to be in high abundance in most mammals such as

white-tailed deer (DeWoody *et al.*, 1995). The first locus to be tested for microsatellite polymorphisms using PCR was in an intron in the human cardiac muscle actin gene containing variable numbers of the tandem dinucleotide $(TG)_n$ repeats, resulting in twelve different allelic fragments (Litt and Luty, 1989).

Microsatellite markers exhibit variability based on the number of repeats that are in tandem at a locus and in the repeat sequence type (Tautz, 1989). Microsatellites can be classified into one of three categories: perfect repeat sequences, imperfect repeat sequences or compound repeat sequences (Weber, 1990). Perfect repeat sequences have no interruptions in the repeat sequence, imperfect repeat sequences have a few scattered single repeat interruptions of a different sequence in the repeat sequence and compound repeat sequences have a tandemly repeated sequence of a different sequence within the repeat sequence. Microsatellite length variations are thought to arise from either unequal exchange during mitosis between sister chromatids or during meiosis between homologous chromosomes or by DNA slippage during replication of the lagging strand (Schlotterer and Tautz, 1992).

Microsatellites are presently considered functionless because many of them lie outside genes in 3′ or 5′ untranslated regions and roughly ten percent are located in introns between genes. Research suggests that they may facilitate the production rate of proteins according to David G. King of Southern Illinois University, may play a role in forming left-handed conformation (Z-DNA) (Hamada *et al.*, 1982), may regulate transcription (Hamada *et al.*, 1984), may play a role in recombination (Pardue *et al.*, 1987) or may facilitate the pairing of chromatids during mitosis. The positioning of microsatellites in noncoding regions of heterochromatin is forgiving of a high mutation

rate.

varia

from

anem

1999

mark

genet

Mark

A

allozy

that is

not pa

not ob

protei

inheri

genet

diagn

and p

these

states

that t

princ

other

rate, estimated to be approximately between 10^{-4} and 10^{-5}, leading to extensive allelic variation (Saiki *et al.*, 1988). " . . . it is 10,000 times more likely to gain or lose a repeat from one generation to the next than a gene such as the one responsible for sickle cell anemia is to undergo the single-base mutation leading to that disease" (Moxon and Wills, 1999). This variability of microsatellites, along with their existence as codominant markers following Mendelian inheritance of alleles gives an excellent reason to use these genetic markers for individual identification.

Marker Requirements

A marker is a genetic unit (gene, microsatellite, minisatellite, restriction site, allozyme, etc.) within the genome that can be followed from generation to generation, that is, from both parents to their offspring. If a region of interest within the genome does not pass from each parent to their offspring, then Mendelian inheritance of that region is not observed and can not be used as a marker to identify individuals. All DNA, whether protein coding or having no sequence dependent function, follows the same rules of inheritance. Mendelian inheritance allows scientists to perform numerous applications: genetic mapping and linkage studies, mapping disease loci, paternity testing, medical diagnostics, pedigree analysis, individual identification, anthropological, evolutionary and population studies and the list goes on. In order for a genetic marker to be useful for these purposes it must follow Mendel's laws of inheritance. The principle of segregation states that each of the two alleles, one from each homologous chromosome segregate so that the offspring has an equal chance of obtaining either allele from both parents. The principle of independent assortment states that these alleles segregate independently of other alleles at different loci. Individual identification requires population allele

frequencies of unlinked polymorphic genetic markers following Mendelian inheritance so that the power of discrimination can be achieved and utilized in forensic applications.

The use of genetic markers for characterizing genetic variability between individuals for identification purposes relies on several factors. The nature of the marker must be informative by showing measurable polymorphisms either in DNA sequence or DNA length between individuals. The polymorphism information content (PIC) of a marker is calculated from allele frequencies in the population. Markers must exhibit genetic diversity among the population, that is, high allelic variation in order to distinguish individuals from each other. The marker must show Mendelian inheritance of the alleles, that is, one allele from each parent must be present in the offspring and should be inherited independently of any other markers being analyzed simultaneously. Lastly, markers that can be easily genotyped will be the marker of choice for individual identification.

Selection of Marker, Technique and Method of Detection

The selection of the most appropriate molecular marker, molecular technique and detection system for answering a particular question, for example in this study individual identification, is based on several important considerations. The informativeness of the marker, the speed, ease and reliability of the method, costs of the materials, the quality and quantity of DNA, the amount of sequence information required to perform the technique, the dominance of the marker, the detection system and the ease of analysis and interpretation of results, will all be compared between various markers and methodologies. They all have their strong points and weaknesses depending on the results desired and the informativeness needed to answer the desired question.

The first consideration that might limit the choice of method is how much and what quality of DNA can be obtained from the type of samples at hand. Forensic samples tend to result in a limited quantity of DNA due to the often minimal amount of evidence left at the crime scene. The quality of DNA is usually poor, as samples are partially degraded due to environmental insults. Based on these initial limitations, one may rule out a marker system based on RFLP and Southern hybridization because this technique requires microgram quantities of DNA of very high quality. Using RFLP on low quantity degraded DNA will lead to ambiguous results which will be very difficult to interpret due to the presence of extra fragments in the already complex band pattern. In paternity cases where non-degraded samples are readily available, RFLP is an excellent choice using single or multi-locus minisatellite probes where Mendelian inheritance is maintained (Pena and Chakraborty, 1994). PCR amplification of small fragments using any of the suitable systems (RAPD, SSCP, STR) would be the favored choice to resolve questions using minute quantities (nanogram amounts) of degraded DNA (Lorente *et al.*, 1997).

The ease and amount of time required to perform a technique and the length of time before obtaining the results are major concerns in this fast-paced society where results are wanted yesterday and for a competitive price. Traditional serologic methods and dot blot assays are the easiest to perform in a short amount of time, usually a few hours, with relatively low cost considering identification kits are manufactured by companies, such as, Perkin-Elmer and Cetus Corporation which do not require expensive detection apparatus. RFLP using Southern hybridization, on the other hand, is very labor intensive requiring restriction digests, gel electrophoresis, Southern transfer, probe hybridization, several washes and film exposure, with results generated in about a week or more. The

cost of the materials is high for enzymes, membranes, x-ray film and the radioactive or chemiluminescent detection of polymorphisms, to say nothing of the highly skilled labor required to perform the technique. PCR based methods are easy to perform with no hybridizations and results are obtained in a few hours utilizing minimal technologist time. The major cost is in the initial purchase of a PCR machine and a detection system, which can be quite high. The continual purchase of PCR reagents and oligonucleotide primers, if radioactively or fluorescently labeled, is expensive but spreading the cost over several hundred assays reduces the per assay cost.

Detection systems can range from simple and inexpensive ethidium bromide, silver staining or colorimetric to the more expensive chemiluminescent, radioactive or fluorescent detection of polymorphisms. The advantages of using ethidium bromide, silver staining or colorimetric are the cost of materials and the ease of use. The disadvantage of using radioactivity is the isotopic component, while the benefit of using chemiluminescence or fluorescence is that they are non-isotopic, safer to use and have a longer shelf life than radioisotopes. Fluorescent detection allows for simultaneous amplification and detection of three or more loci, which greatly increases the speed and throughput of samples with an increase in sensitivity compared to other systems (Lins *et al.*, 1996).

The informativeness of the marker along with the dominance or codominance characteristic are very important considerations when choosing the most appropriate marker for the job. Allozyme markers are not very informative because they represent too few loci and exhibit too little variation. RFLP's probed with minisatellites are at the other end of the marker spectrum, with hypervariability across many loci. Multi-locus

minisatellite RFLP's are extremely informative because they result in numerous bands representing a unique individual-specific DNA profile, except in identical twins. The disadvantage of multi-locus markers is that they are dominant markers only, with some genetic information possibly being lost or hidden due to the absence of a restriction site. Single-locus minisatellite or cDNA probe is not very informative because a single locus will only result in one or two alleles depending on the individual's zygosity. They are codominant markers with both alleles being represented, one band for homozygous individuals and two bands for heterozygous individuals. The use of several unlinked single-locus probes will increase the informativeness of the technique. In fact this was the first accepted use of DNA testing in forensics. RAPD's are dominant markers with polymorphic bands scored as present or absent. The informativeness of this marker is limited because a heterozygous individual cannot be distinguished from a homozygous individual. SSCP is considered a codominant marker, with the ability to detect heterozygotes and with limited informativeness because only a single base pair mutation is being detected and only if the substitution changes the mobility of the DNA fragment. Microsatellites are also codominant markers and are extremely informative based on the high degree of polymorphism present among individuals.

The amount of sequence information required prior to performing the technique is a matter of concern if money and time are an issue when choosing the method of analysis. RFLP using Southern hybridization does not require actual sequence information, but it does require known gene (cDNA) probe information, though this information may not have to be regenerated for each species tested. The use of a previously identified gene as a probe eliminates the time and money required to obtain species-specific sequence

information. Multi-locus probes for DNA typing are VNTR's which are not necessarily species-specific, but require sequence information, with no effort and cost to attain. PCR derived systems require exact sequence information for the development of oligonucleotide primers with high cost and effort to obtain. However, comparative mapping utilizing primer pairs from closely related species can be both economical and time efficient when there is conservation of heterologous primer pairs between species (Moore *et al.*, 1991). RAPD's, on the other hand, do not require any sequence information due to the arbitrary generation of oligonucleotide primer pairs used in the technique.

Interpretation and analysis of results can be ambiguous or straightforward depending on which marker and detection method are used. Traditional serologic procedures using agglutination of antigens and staining of allozymes with colored dyes, along with the colorimetric detection in dot blot assays are very straightforward analyses. The presence of a reaction taking place, detected by agglutination, staining or color development, is very unambiguous, that is, present or absent. Analysis of RFLP's when a single-locus probe is used is straightforward, depending only on the presence of one or two bands of a specific size. Use of a multi-locus probe can lead to a very problematic analysis. The banding pattern can be complex due to the number of fragments present, the possibility of extra fragments from slightly degraded samples or the sharing of bands between loci can make interpretation of the data challenging. The use of size standards in adjacent lanes of an RFLP gel can result in estimates of allele sizes, usually with errors of 2-5 bp. RAPD's can also be difficult to interpret because of the presence of many bands, some of which could be similarly sized alleles from different loci and some of which are artifactual

based on the sensitivity of the assay to the PCR conditions. SSCP analysis can be difficult if the running conditions are not optimized because slight changes can effect the mobilities of single-stranded DNA molecules leading to unpredictable results. Lastly, microsatellites can be easy to score if the tandem repeat is not a dinucleotide motif because these alleles show a characteristic stutter band due to the slippage of the polymerase during extension of the PCR product (Hauge and Litt, 1993; Murray *et al.*, 1993). The stutter usually creates a band two and four base pairs smaller in size than the original fragment size, creating ambiguities in calling a true heterozygote with alleles differing by a single repeat unit from a homozygote with a strong stutter band. However, the presence of these characteristic stutter bands is beneficial in separating true alleles from background noise. Fluorescent detection of microsatellite alleles makes calling the sizes of fragments more accurate because of the use of an internal size standard in each lane of the gel, thus, correcting for lane to lane and gel to gel variability.

The information presented here, summarized below in Table 1, and in this literature review is the basis for the selection of using ubiquitous, microsatellite markers, fluorescent detection and PCR technology to achieve the goal of this research project. The objective of this study is to develop a practical, economical, time-efficient DNA typing system for wildlife forensic scientists to utilize for individualizing forensic evidence related to Michigan white-tailed deer.

Table 1. Comparison of markers, techniques and methods of detection.

Characteristic	Marker/Technique		
	serologic/dot blot/ dot blot-PCR	**RAPD,SSCP, STR/PCR**	**RFLP/Southern hybridization**
quality of sample	relatively fresh samples dot blot-PCR: low molecular weight DNA	low molecular weight DNA	high molecular weight DNA
quantity of sample	minimal amounts	nanogram amounts	microgram amounts
turn around time for results	few hours	one or two days	several days to a week or more
technologist time required to perform technique	minimal time easiest to perform	minimal time easy to perform	labor intensive requires a skilled technologist
cost of materials	commercial kits can be expensive overall: very cheap	initial cost: expensive continual cost: inexpensive	cost: very high
Informativeness of marker	least informative dot blot-PCR: very informative	most informative when using STR's, otherwise very limited	most informative when using multi-locus probes
marker characteristic	dominant	RAPD's: dominant SSCP: codominant STR's: codominant	single-locus: codominant multi-locus: dominant
sequence information requirement	none dot blot-PCR: complete	RAPD's: none SSCP: complete STR's: complete	single-locus: none multi-locus: complete
interpretation of results	straightforward	RAPD's and SSCP: ambiguous STR's : unambiguous	single-locus: straightforward multi-locus: ambiguous
	Detection System		
	silver staining ethidium bromide colorimetric	**radioisotope**	**chemiluminescent fluorescent**
advantage of detection system	simple to use inexpensive	high sensitivity clear interpretation of results	safer to use easier to dispose of longer shelf-life high sensitivity
disadvantage of detection system	limited sensitivity ethidium bromide is a carcinogen	use of radioactivity	expensive

METHODOLOGY

This study involved using random samples of the population of Michigan white-tailed deer to establish a database. Tissue samples were collected from deer in all 83 counties in Michigan by Tim Tesmer, Jessie Marcus, Ron Southwick and Dr. Paul Coussens. The Department of Natural Resources (DNR) and the U.S. Fish and Wildlife Service jointly funded the project in 1994. Hunters were asked at DNR checkpoints to voluntarily allow these samples to be collected. The deer's age, sex, county and deer management unit (DMU) were documented on the collection tube for each sample collected. Muscle tissue was collected and samples were placed in polypropylene tubes with caps and kept in a -20° C freezer until needed for DNA extraction. To date approximately 1200 deer tissue samples have been collected with only a few counties not totally represented. Bay, Clinton, Macomb and Sanilac counties have only one collected sample, while Lapeer has four and Keweenaw has none. These counties will not be fully represented in the database on the basis of five random deer samples per county unless additional samples are collected prior to the completion of this study.

Tissues for many of the counties had DNA already extracted by Ms. Jean Robertson using a standard protein digestion, phenol/chloroform extraction and ethanol precipitation procedure (Sambrook *et al.*, 1989). The quality and quantity of these DNA samples when initially measured seemed to be good on paper however, many of these DNA samples have since degraded. The quality of these DNA's was tested on a 0.8% agarose gel (SeaKem LE Agarose from FMC) (see Appendix A for gel preparation). Good quality

used for database acquisition, while the remainder of the samples, about 185, were re-extracted using a slightly modified procedure (see Appendix B for extraction procedure). The reason for this modification was to obtain the highest quality and quantity DNA possible from these tissues.

DNA samples used for the initial screening of the eight fluorescently labeled microsatellite markers (see Table 2 for details) along with the 425 samples necessary for the database were selected at random. DNA samples with low quantitation (less than 15 ng/μl) and/or low absorbance ratio, 260/280, (less than 1.5) were ultimately excluded for failure to provide acceptable data. Twenty samples were randomly selected from the approximate 1200 deer samples collected in 1994, measured on a Milton Roy Spectronic Genesys 5 spectrophotometer for an estimate of DNA quality and on a Hoefer TKO 100 fluorometer for DNA concentration and adjusted to be approximately 30 ng/μl. The 425 samples for the database, five from each county and two islands, were processed in the same manner. All samples were thus quantitatively standardized prior to performing polymerase chain reactions.

Table 2. Characteristics of the microsatellite markers used in this study.

LOCUS SYMBOL	LOCUS NAME	PRIMER SEQUENCES (5′-3′)	FLUORO-CHROME AT 5′ END	ALLELE SIZE RANGE	MAMMALIAN SOURCE AND LOCATION
OBCAM	Opiod binding cell adhesion protein	CCTGACTATAATGTACAGATCCCTC GCAGAATGACTAGGAAGGATGGCA	FAM (Blue)	195-211	Bovine Chromosome 29
IGF1	Insulin-like growth factor 1	CATATTTTTCTGCATAACTTGAACCT GAGGGTATTGCTAGCCAGCTG	TET (Green)	128-146	Bovine Chromosome 5
CRFA	Corticotrophin releasing factor	GCTGAGCAGCCGTCTAAGTTGC CTCGCTCACCTGCAGAAGCACC	TET (Green)	226-252	Ovine Chromosome 9
JP23	Cervine DNA Segment, JP0023	GAAAATCCAAGCGACAAAGG CCGCAGAACAACTAAGCCCAAG	FAM (Blue)	140-200	Cervine Linkage Group 4
JP15	Cervine DNA Segment, JP0015	CCTTCTTTCTCATTGCTAACTTATATTAAATATCC GGAAATACCTTATCTTTCATTCTTGACTGTGG	TET (Green)	110-134	Cervine Linkage Group 2
IRBP2	Interphotoreceptor retinoid binding protein	GTATGATCACCTTCTATGCTTCC CCCTAAATACTACCATCTAGAAG	HEX (Yellow)	137-145	Bovine Chromosome 28
CSN3	Casein, Kappa	ATGCACCCTTAACCTAATCCC GCACTTTATAAGCACCACAGC	HEX (Yellow)	188-192	Bovine Chromosome 6
OarFCB304	Ovine DNA Segment, OarFCB304	CCCTAGGAGCTTTCAATAAAGAATCGG CGCTGCTGTCAACTGGGTCAGGG	HEX (Yellow)	121-149	Ovine Chromosome 19

Markers in bold were multiplexed together to create the database.

PCR conditions were optimized for amount of $MgCl_2$ (Perkin-Elmer, 25mM), dNTP's (GeneAmp dNTP's from Perkin-Elmer, 10 mM), oligonucleotide primers (Integrated DNA Technologies, Inc., Operon Technologies, Inc., Gibco BRL and MSU, 2 mM each), Ampli-Taq DNA polymerase (Perkin-Elmer, 5 U/μl), GeneAmp 10X PCR buffer containing no $MgCl_2$ (Perkin-Elmer) and deionized water. The working concentration of each component used in the PCR was as follows: 2.5 mM dNTP solution, 5 U/μl Ampli-Taq DNA polymerase, 20 μM each primer and 25 mM $MgCl_2$ solution. All PCR reactions were performed using 60 ng of template DNA, in a volume of 25 μl, overlayed with one drop of mineral oil. The amounts of each PCR component, depending on the primer set, are listed in Table 3. Multiplexing OBCAM, CRFA and IGF1 together in one reaction changed the amount of each primer from 0.5 μl to 0.25 μl and deionized water from 14.25 μl to 13.75 μl, while maintaining the same concentrations of the other components. PCR thermocycler running conditions using the PTC-100 Thermocycler from MJ Research were also optimized for each primer set and are listed in Table 4.

Table 3. Amount of each component for a 25 μl PCR.

Marker	10X Buffer	2.5mM dNTPs	25mM $MgCl_2$	20μM Primer	5U/μl TAQ	30ng/μl DNA	dH_2O
OBCAM	2.5μl	2.0μl	3.0μl	0.5μl	0.25μl	2.0μl	14.25μl
CRFA	2.5μl	2.0μl	3.0μl	0.5μl	0.25μl	2.0μl	14.25μl
IGF1	2.5μl	2.0μl	3.0μl	0.5μl	0.25μl	2.0μl	14.25μl
IRBP2	2.5μl	2.0μl	3.0μl	0.5μl	0.25μl	2.0μl	14.25μl
CSN3	2.5μl	2.0μl	3.0μl	0.5μl	0.25μl	2.0μl	14.25μl
JP15	2.5μl	2.0μl	3.0μl	0.5μl	0.25μl	2.0μl	14.25μl
FCB304	2.5μl	2.0μl	3.0μl	0.5μl	0.25μl	2.0μl	14.25μl
JP23	2.5μl	2.0μl	1.5μl	0.5μl	0.25μl	2.0μl	15.75μl

Table 4. PCR thermocycler temperature programming.

MARKER	INITIAL DENATURATION	DENATURATION	ANNEALING	EXTENSION	AUTO-EXTEND	# OF CYCLES	FINAL EXTENSION
OBCAM	95°C for 3 minutes	95°C for 45 seconds	55°C for 45 seconds	72°C for 45 seconds	1 second per cycle	30	N/A
CRFA	95°C for 3 minutes	95°C for 45 seconds	55°C for 45 seconds	72°C for 45 seconds	1 second per cycle	30	N/A
IGF1	95°C for 3 minutes	95°C for 45 seconds	55°C for 45 seconds	72°C for 45 seconds	1 second per cycle	30	N/A
IRBP2	95°C for 3 minutes	95°C for 45 seconds	50°C for 45 seconds	72°C for 45 seconds	1 second per cycle	30	N/A
CSN3	95°C for 3 minutes	95°C for 45 seconds	55°C for 45 seconds	72°C for 45 seconds	1 second per cycle	30	N/A
JP15	95°C for 3 minutes	95°C for 45 seconds	50°C for 45 seconds	72°C for 45 seconds	N/A	30	N/A
FCB304	95°C for 3 minutes	95°C for 45 seconds	58°C for 45 seconds	72°C for 45 seconds	N/A	30	N/A
JP23	95°C for 3 minutes	95°C for 45 seconds	55°C for 45 seconds	72°C for 45 seconds	N/A	30	72°C for 5 minutes

Verification of optimal conditions was performed on 2% agarose gels, with ethidium bromide staining. All sample runs during optimization contained positive and negative controls, two random deer samples and DNA size standards in the first and last lanes of the gel. 5 µl of each standard combined with 1 µl of loading dye (Blue Dextran and Ficoll from Sigma) mixed with fragment comigration indicator dyes xylene cyanole FF (Sigma) and bromphenol blue (Sigma) were loaded in the wells of the gel. 15 µl of each sample combined with 3 µl of loading dye as described above were loaded in the wells of the gel. The test samples used were 95201 and RP85, the positive control was a sample of muscle from a single white-tailed deer and the negative control contained no DNA. The size standards used were 100 bp ladder (Gibco BRL), pUC19 digested with Hae III purified from *Haemophilus aegyptius* and/or pUC19 digested with Hpa II purified from *Haemophilus parainfluenzae*. These size standards have known fragment sizes in the range similar to the expected size fragments of the eight markers. The appearance of a single band in the approximate fragment size range was an indicator that the PCR conditions were specifically amplifying only the locus of interest. Once optimization was achieved for each marker, the 20 random samples were amplified using each marker's optimal PCR conditions. These samples were subsequently analyzed on a Perkin-Elmer ABI Prism 377 DNA Sequencer for more accurate size determinations of each allele and to determine the extent of genetic diversity among the species.

Fluorescently labeled PCR samples were electrophoresed on a standard 4% polyacrylamide denaturing gel of 0.2mm thickness (see Appendix C for gel preparation). 1.5 µl of each sample was loaded on the ABI containing the following components: 0.5

μl of PCR product, 0.5 μl ABI Prism GeneScan-350 TAMRA (Perkin-Elmer) internal size standard, 2.0 μl formamide (Amresco) and 1.0 μl ABI loading buffer (Blue Dextran, 50 mg/ml and EDTA, 25 mM). Samples were electrophoresed at 3000 Volts for 2 hours using 1X TBE (Tris/boric acid/EDTA) buffer, data was collected using standard ABI protocols of the ABI 377XL DNA Sequencer Data Collection 2.0 software and analyzed using the recommended local Southern method algorithm of the GeneScan Analysis 3.1 software.

Scoring of alleles was performed after the analysis software extracted the raw data from each lane and applied the local Southern algorithm when calculating the molecular length of unknown fragments using the selected internal size standard. The fragment sizes listed for each sample were represented as a number to the 100^{th} decimal. The GeneScan Analysis 3.1 software generated sizes for all peaks detected and then alleles were manually scored as whole numbers. Similar alleles, differing by one base pair, were "binned" together because alleles of a dinucleotide microsatellite marker will differ by multiples of 2. The scoring of alleles differing by only one dinucleotide repeat was performed based on the signal intensity for each allele. That is, a true heterozygote should display a more intense signal for the smaller sized fragment compared to the larger fragment, which is caused by the overlap of the smaller allele with the stutter band of the larger allele. The presence of only one fragment was scored as a homozygote with both alleles being the same size. All other samples were genotyped based on the presence of fragments within the allele size range showing the characteristic stutter bands.

Initial screening using the twenty random samples with each of the eight markers, was performed to identify which markers were polymorphic and the percent heterozygosity. Three of the eight markers tested were able to be multiplexed together and analyzed simultaneously on approximately 425 Michigan white-tailed deer to create a database with approximately 850 data points. A large database is necessary for probability estimates in order to individualize a sample and match it to evidentiary material with a high degree of probability. These selected three markers for database acquisition were tested for Mendelian inheritance using nine two-generation known pedigree white-tailed deer families from a captive deer population in Savannah River, Georgia.

FINDINGS

DNA QUALITY ASSESSMENT

The first set of results were obtained after an initial assessment of DNA quality by electrophoresing random samples of genomic DNA on agarose gels. Figures 3 and 4 exhibit highly variable DNA quality between samples previously extracted beginning in 1994. Samples with detected bands near the loading wells represent good quality DNA, while samples showing a smear or at the lower end of the gel represent highly degraded DNA. Several samples have since degraded and show an appreciable amount of low molecular weight DNA. High molecular weight DNA is not a requirement for PCR technology however, the presence of high quality DNA will make the optimization of PCR's and evaluation of results easier to interpret. Figure 5 shows genomic DNA extracted using both Qiagen lysis Buffer ATL and homemade lysis buffer, along with Proteinase K (Gibco BRL) and Qiagen Protease enzymes. The selection of the Qiagen lysis Buffer ATL with Proteinase K resulted in the highest quality DNA and was used for all subsequent DNA extractions. The quality or purity of DNA, absorbance ratio (260nm/280nm), of these samples ranged from 1.7-1.9, which represents high quality DNA because the amount of DNA in the sample (i.e. absorbance at wavelength 260nm) exceeds the amount of protein (i.e. absorbance at wavelength 280nm) by almost two-fold. A_{260} corresponds to the amount of absorbance of light by DNA in the sample at wavelength 260nm and A_{280} corresponds to the amount of absorbance of light by protein and other contaminants in the sample at wavelength 280nm. The Milton Roy Spectronic Genesys 5 spectrophotometer, used in this study, measures absorbances and calculates

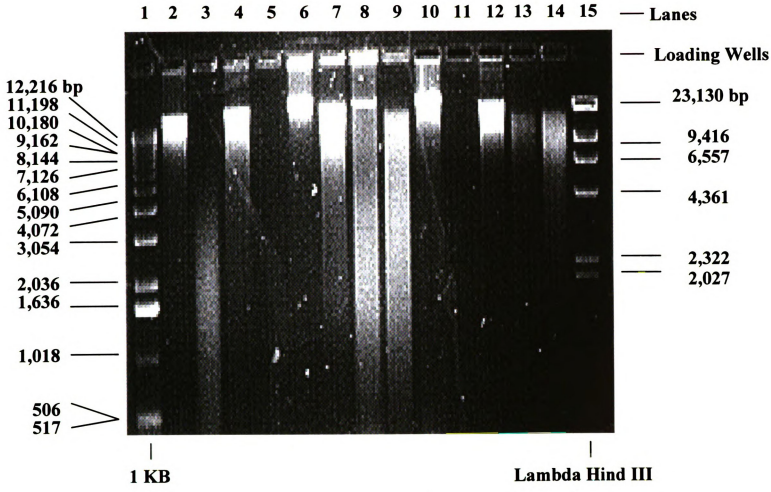


Figure 3. DNA quality of previously extracted genomic DNA samples.

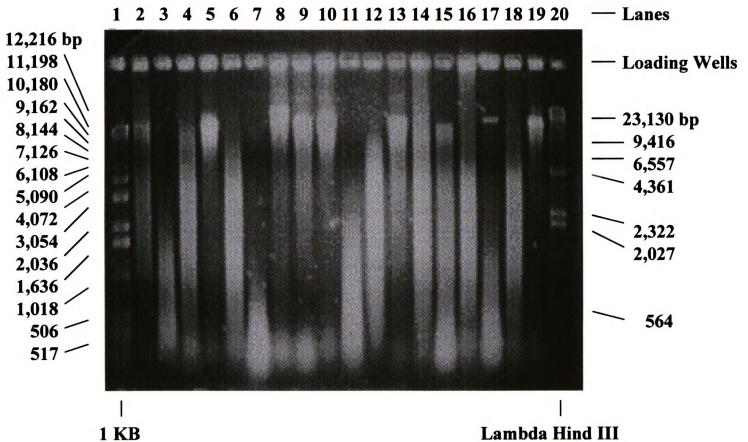


Figure 4. Additional genomic DNA samples previously extracted showing highly variable DNA quality.

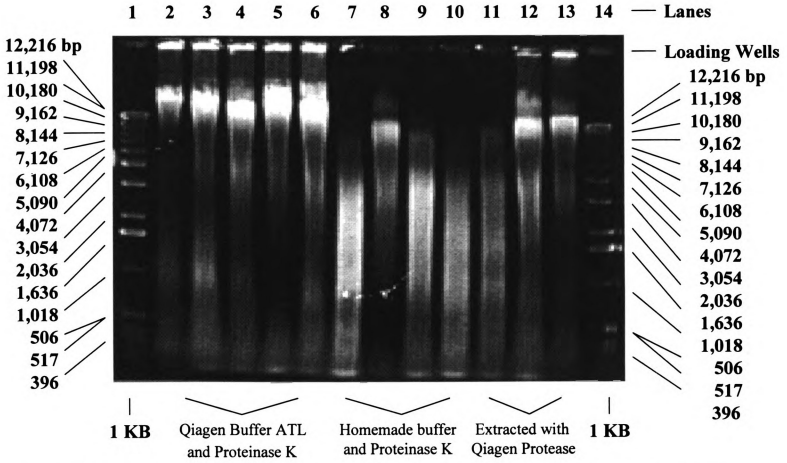


Figure 5. Comparison of genomic DNA samples extracted using two different lysis buffers and enzymes.

the 260nm/280nm absorbance ratio from the absorbance readings at wavelengths 260nm and 280nm. The ratio provides an estimate of the purity of the DNA.

PCR OPTIMIZATION OF EIGHT MARKERS

Two samples along with negative and positive controls were selected and used for initial PCR optimization of the eight markers. Figures 6 through 11 show the eight markers optimized and electrophoresed on agarose gels along with known size standards in external lanes. The presence of a single band in the approximate fragment size range based on species homology listed in Table 2 indicated that the PCR conditions were optimally amplifying the locus of interest. The absence of a band for the negative control indicates no DNA contamination and no PCR artifacts present in the size range of interest.

INITIAL SCREENING OF MARKERS WITH TWENTY SAMPLES

Two of the most important aspects of DNA markers are heterozygosity and diversity. A goal of this study was to identify a set of markers with high heterozygosity and numerous possible alleles in the test population. The next set of results were obtained to determine the extent of genetic diversity and to select those markers exhibiting high polymorphism among the species using twenty random samples containing high quality DNA. These samples were amplified with fluorescent primers using each marker's optimal PCR conditions and electrophoresed on an ABI 377XL DNA Sequencer. The initial screening data for the twenty samples for each marker along with the number of heterozygous samples, percent heterozygous, number of different alleles, the allele size range and polymorphic information content (PIC) are presented in Table 5. The PIC value was calculated using the formulation of Botstein *et al.* (1980) and is indicative of a

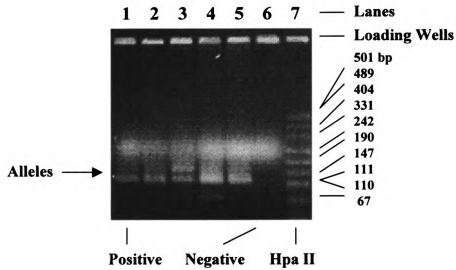


Figure 6. Agarose gel of marker: JP15.

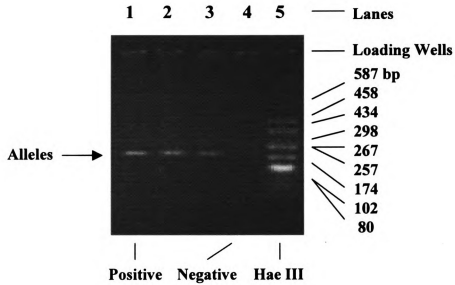


Figure 7. Agarose gel of marker: CSN3.

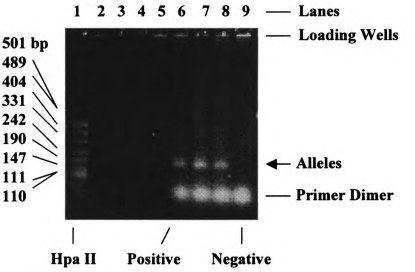


Figure 8. Agarose gel of marker: IRBP2.

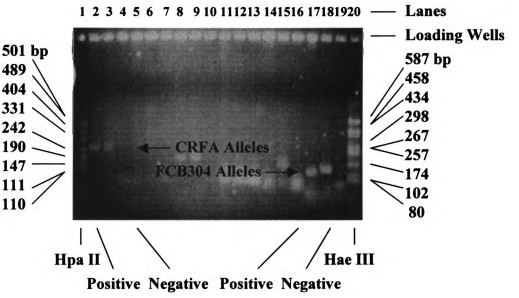


Figure 9. Agarose gel of markers: CRFA and FCB304.

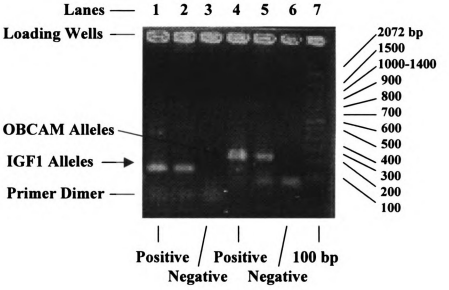


Figure 10. Agarose gel of markers: IGF1 and OBCAM.

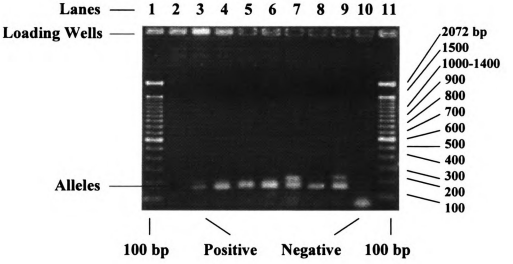


Figure 11. Agarose gel of marker: JP23.

Table 5. Initial screening results in base pairs from an ABI 377 of eight markers using random deer samples.

COUNTY	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	CSN3 HEX		FCB304 HEX		CRFA TET		JP15 TET		OBCAM FAM		IRBP2 HEX		IGF1 TET		JP23 FAM	
Delta	AL 94	28	1.6	192	192	120	120	254	254	111	113	205	211	173	179	127	127	NP	NP
Gogebic	AL 90	30	1.8	192	192	120	120	238	240	107	113	207	215	175	175	127	127	NP	NP
GD??	BR 191	25	1.8	192	192	120	134	244	244	111	111	195	203	177	177	127	141	154	154
Kalamazoo	PW 109	65	1.7	188	188	120	134	240	244	107	111	207	207	175	179	127	127	158	158
Hillsdale	95168	51	1.8	192	192	120	120	244	254	111	111	209	221	173	181	127	141	158	158
Emmet	RP 273	25	1.8	188	192	134	134	240	246	111	111	203	207	173	175	127	131	158	158
Beaver Island	95295	125	1.9	188	188	120	120	244	252	111	111	203	219	177	177	127	141	152	152
??	RL 19	21	1.5	188	188	134	134	238	244	111	111	209	215	173	177	133	141	154	168
Emmet	RP 190	20	1.8	188	188	120	134	238	250	111	111	195	201	169	173	127	127	NP	NP
Beaver Island	95290	125	1.9	188	188	120	134	252	258	111	111	203	205	177	181	127	127	NP	NP
Iosco	BR 108	63	1.8	188	192	120	134	238	254	107	115	203	213	177	187	129	141	170	170
N. Manitu Island	95045	93	1.8	192	192	120	120	242	242	107	111	205	221	175	181	129	135	152	152
Washtenaw	JK 99	56	1.7	188	188	120	120	248	252	111	111	195	209	177	177	127	127	158	158
Dickinson	AL 5	75	1.8	188	188	120	134	240	248	111	111	201	205	173	175	135	141	NP	NP
??	BG 43	40	1.5	188	188	120	120	242	246	111	111	205	211	173	173	129	131	158	162
Wayne	95083	67	1.8	188	188	134	134	250	250	111	111	201	209	173	175	127	135	154	154
Ottawa	95256	17	1.5	188	188	120	120	238	238	111	111	205	209	173	175	135	141	152	152
Monroe	95087	116	1.9	192	192	120	120	246	254	111	111	209	209	179	179	127	133	154	154
Alcona	BR 254	38	1.7	188	188	134	134	242	250	111	111	205	207	177	179	127	129	154	170
Leelanau	95005	45	1.8	188	192	120	134	230	252	107	111	203	215	175	181	135	143	158	158
Van Buren	95114	100	1.6															154	154
Wayne	95082	80	1.7	NP means no amplified product														152	152
Ottawa	95259	85	1.7	Shaded markers were selected to be multiplexed together for database acquisition														152	152
Number of Heterozygous Samples				3		7		15		6		18		14		14		3	
Percent Heterozygous Samples				15		35		75		30		90		70		70		17	
Polymorphic Information Content				0.36		0.36		0.88		0.32		0.87		0.77		0.68		0.70	
Number of Different Alleles				2		2		11		4		11		7		7		6	
Allele Size Range				188-192		120-134		230-258		107-115		195-221		169-187		127-143		152-170	

marker's ability to differentiate individuals. Heterozygosity was calculated by dividing the number of heterozygous individuals by the total number of individuals. Heterozygosity of the eight markers screened on twenty random deer samples is graphically represented in Figure 12. Marker loci CRFA, OBCAM, IRBP2 and IGF1 displayed a minimum 70% heterozygosity and appeared to be highly polymorphic. These markers were selected for use in analysis of database samples. Figure 13 shows a graphical representation of the degree of polymorphism for the three selected markers CRFA, OBCAM and IGF1. These three markers could be amplified under similar PCR conditions and thermocycler temperature programming (multiplexed). The ability to multiplex was considered critical since the final test needed to conserve cost as well as time. In addition, these markers also have non-overlapping allele size ranges so as to be able to be analyzed in one lane of a polyacrylamide gel.

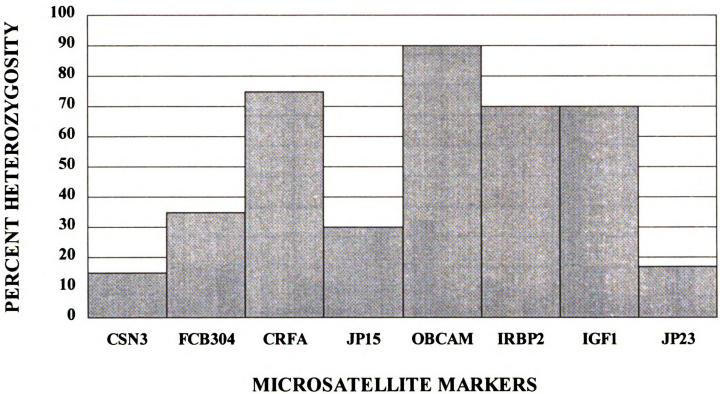


Figure 12. Heterozygosity of eight markers screened on twenty random deer samples.

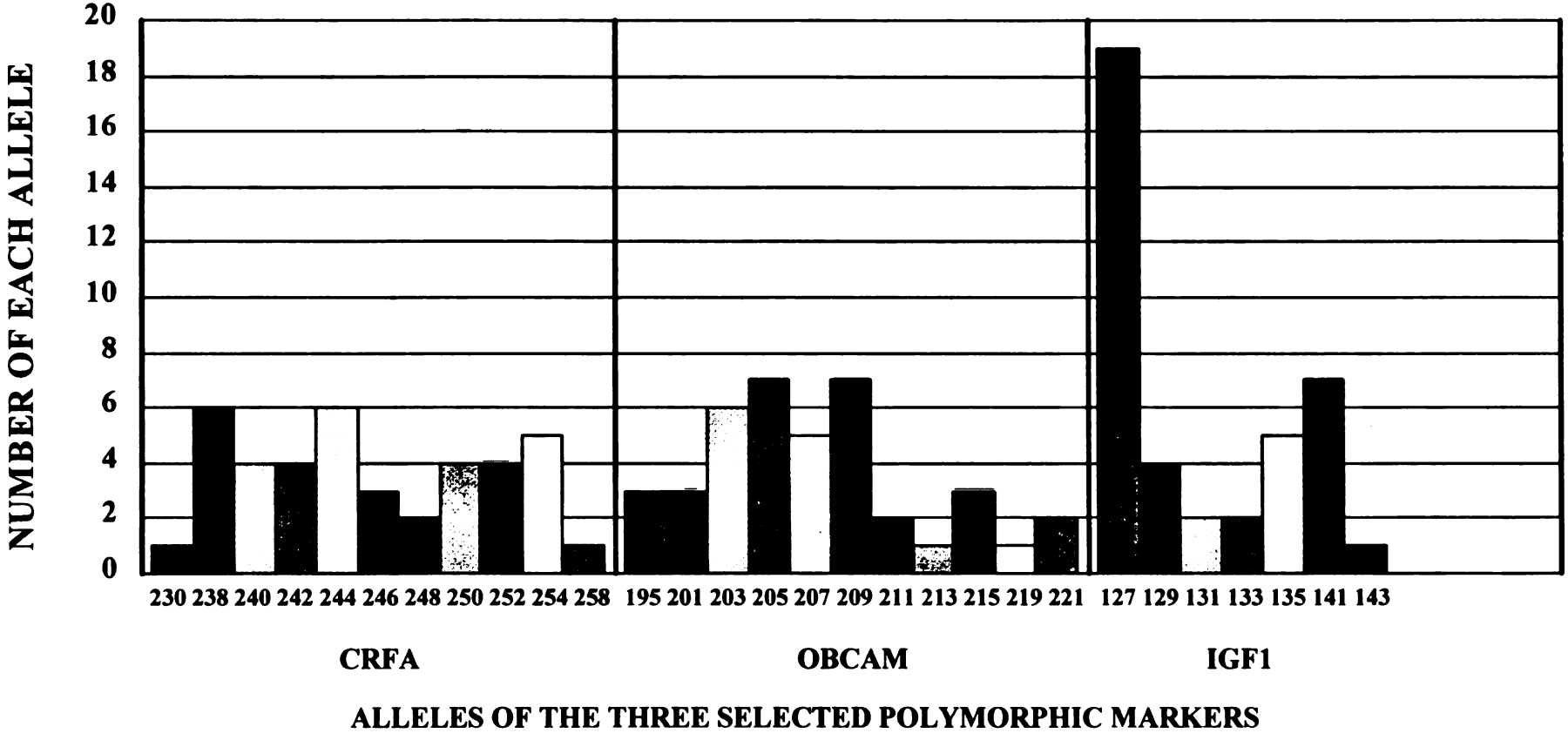


Figure 13. Graphical representation of the degree of polymorphism for the three selected markers screened on twenty random deer samples.

WHITE-TAILED DEER FAMILY STUDY

The three selected markers, in order to be considered genetic markers must follow Mendel's laws of inheritance, prior to database acquisition. This study analyzed nine two-generation white-tailed deer families originating from a captive deer population in Savannah River, Georgia, and supplied by Dr. Jerry L. Ruth with the Fish and Wildlife Service of the United States Department of the Interior. The results from these families prove Mendelian inheritance of the three markers used for database acquisition in this study. Figure 14 shows the ABI gel file containing nine pedigreed white-tailed deer families for the three selected markers and Table 6 shows the allelic data for the nine pedigreed white-tailed deer families contained in Figure 14. Notice that each offspring contains one allele from each of the two parents, thus, proving Mendel's principle of segregation of alleles for the three markers.

DATABASE ACQUISITION AND STATISTICAL ANALYSIS

Database results were obtained by amplifying the selected three markers (CRFA, OBCAM and IGF1) together and simultaneously running them in one lane of an ABI 377XL DNA Sequencer (multiplexing). An example of an ABI 377 gel file containing 50 deer DNA samples amplified by this method using the three multiplexed markers together in each lane is shown in Figure 15. Data from these gel files is displayed in an electropherogram (Figure 16) using GeneScan software to analyze and "call" allele sizes, based on comparison to internal size standards. Note that the table in Figure 16 accounts for the presence of two alleles per locus and that all three loci are discernable in the electropherogram. To construct a "representative" database against which allele frequencies could be determined for forensic samples, 450 different deer DNA samples

were analyzed by multiplex PCR using the three markers CRFA, OBCAM and IGF1 on the ABI 377 (Table 7). NP means that there was no amplified product for that marker for that deer sample and ? means that the allele for that deer sample and marker was unable to be scored confidently. Statewide, average heterozygosity of the three loci, OBCAM and CFRA exceeded 80% while that for IGF1 exceeded 60% (Figure 17). Figures 18, 19 and 20 show a graphical representation of the frequency of IGF1, CRFA and OBCAM alleles in the database, respectively.

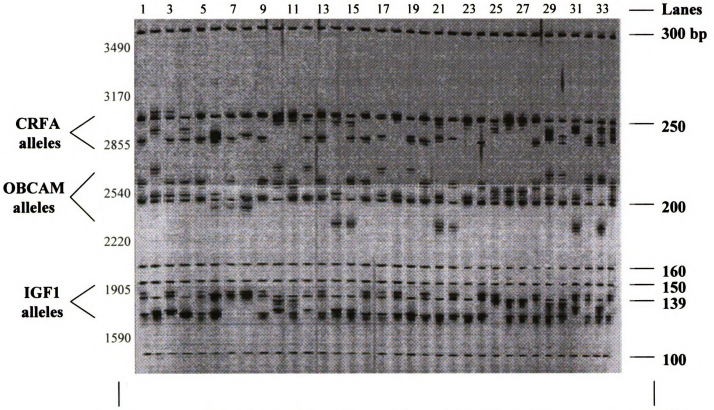


Figure 14. ABI gel file containing nine pedigreed white-tailed deer families.

Table 6. Allelic data for the nine pedigreed white-tailed deer families from Figure 14.

GEL LANE #	SAMPLES	MARKERS					
		IGF1 (bp)		OBCAM (bp)		CRFA (bp)	
1	Father 1	127	143	201	211	236	250
2	Mother 1	127	129	203	219	242	254
3	Offspring 11A	129	143	203	211	236	254
4	Offspring 11B	127	127	203	211	236	242
5	Father 1	127	143	201	211	236	250
6	Mother 2	127	141	195	203	236	238
7	Offspring 12A	141	143	195	211	236	250
8	Offspring 12B	141	143	195	201	238	250
9	Father 1	127	143	201	211	236	250
10	Mother 3	129	135	209	219	248	248
11	Offspring 13A	127	135	201	209	248	250
12	Offspring 13B	129	143	201	219	236	248
13	Father 1	127	143	201	211	236	250
14	Mother 4	129	129	185	203	246	250
15	Offspring 14A	127	129	185	211	236	246
16	Father 1	127	143	201	211	236	250
17	Mother 5	127	143	203	219	238	250
18	Offspring 15A	127	143	201	203	250	250
19	Offspring 15B	127	127	201	219	236	238
20	Father 1	127	143	201	211	236	250
21	Mother 6	127	127	185	203	238	248
22	Offspring 16A	127	143	185	211	238	250
23	Offspring 16B	127	127	201	203	248	250
24	Father 1	127	143	201	211	236	250
25	Mother 7	137	143	203	207	246	248
26	Offspring 17A	127	137	201	207	248	250
27	Offspring 17B	127	137	201	207	248	250
28	Father 1	127	143	201	211	236	250
29	Mother 8	127	135	203	217	240	242
30	Offspring 18A	127	135	201	217	236	240
31	Father 2	133	143	185	203	246	246
32	Mother 9	127	143	211	215	238	240
33	Offspring 29A	127	133	185	215	238	246
34	Offspring 29B	127	143	203	211	240	246

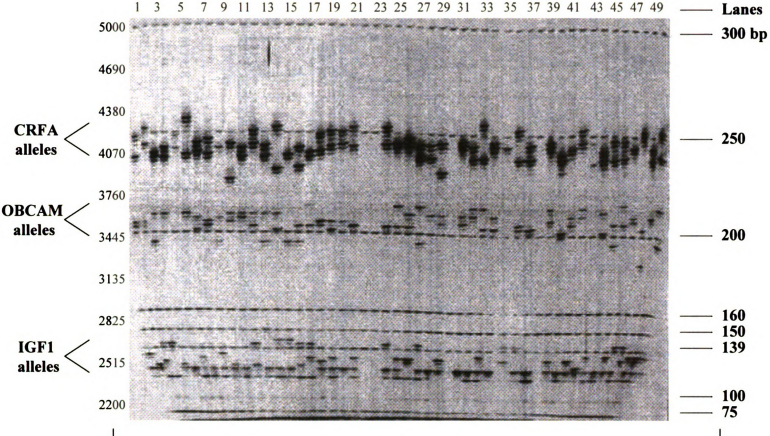
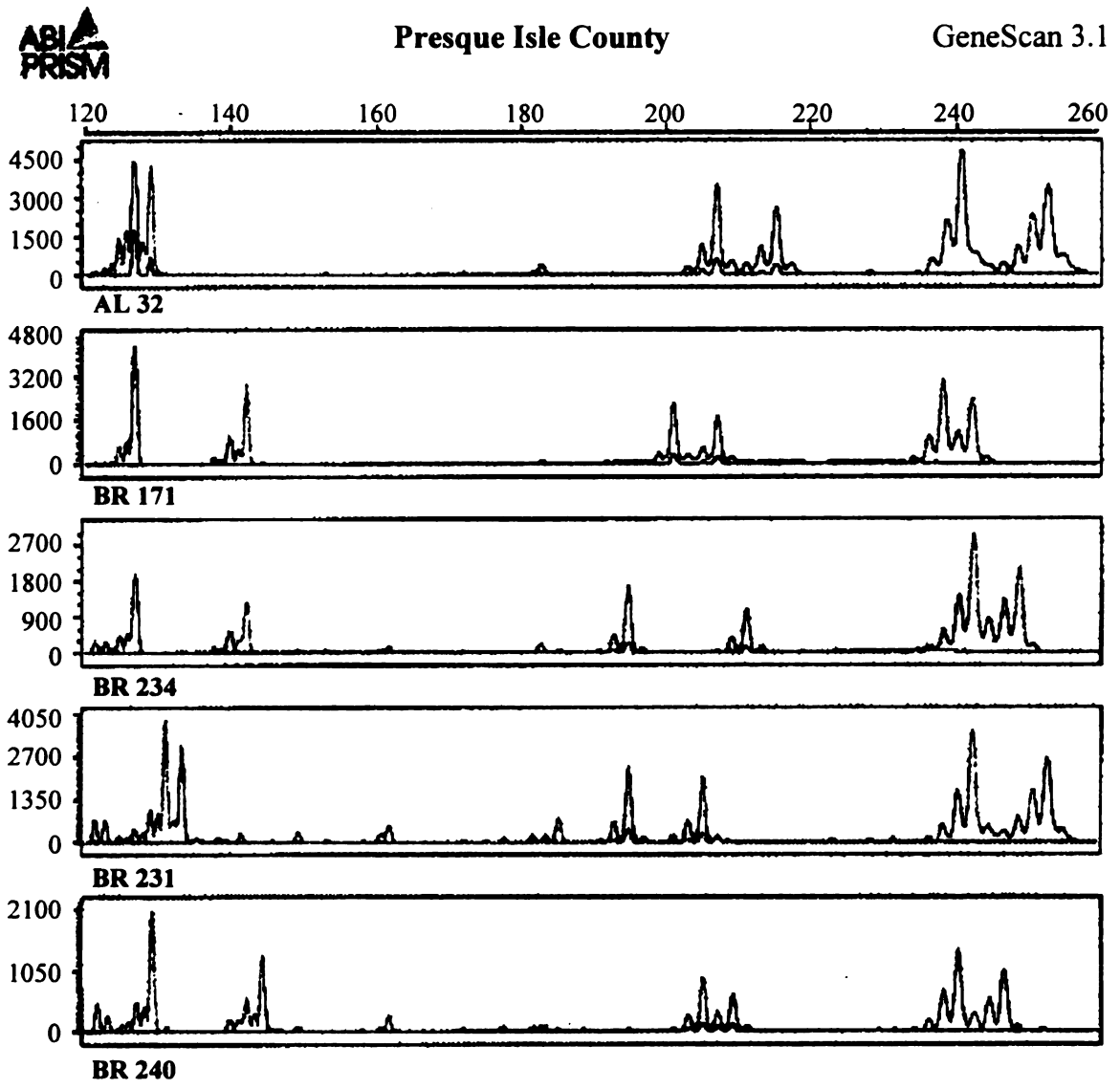


Figure 15. Example of an ABI gel file containing fifty samples with the three multiplexed markers.



SAMPLE	IGF1-1	IGF1-2	OBCAM-1	OBCAM-2	CRFA-1	CRFA-2
AL 32	127	129	207	215	241	253
BR 171	127	142	201	207	238	242
BR 234	127	142	195	211	243	249
BR 231	131	133	195	205	242	253
BR 240	129	144	205	209	240	247

Figure 16. Example of ABI electropherograms of five samples from Presque Isle.

Table 7. Michigan white-tailed deer database.

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
1	ALCONA	BR 173	100	1.8	203	211	129	137	244	244
		BR 230	48	1.7	203	207	133	137	242	252
		BR 238	62	1.6	195	215	127	135	242	248
		BR 37	57	1.9	203	215	127	141	240	254
		BR 90	68	1.9	201	205	129	141	244	244
		BR 254	38	1.7	205	207	127	129	242	250
2	ALGER	RP 278	50	1.9	205	213	127	127	238	252
		AL 114	85	2	205	213	127	127	238	252
		RP 113	344	1.9	203	203	127	127	252	252
		ES 58	62	1.7	205	213	127	127	238	248
		RP 154	112	1.9	209	213	127	127	238	248
		RP 277	104	2	203	205	127	135	**NP**	**NP**
3	ALLEGAN	PW 101	100	1.8	205	205	127	141	238	238
		PW 108	52	1.6	201	205	127	141	244	248
		PW 116	80	1.7	203	205	127	127	244	252
		PW 118	63	1.6	203	205	127	127	248	252
		PW 120	32	1.5	203	205	127	129	238	244
4	ALPENA	BR 246	104	1.9	209	213	127	133	246	246
		BR 208	25	2	207	207	127	127	244	248
		BR 218	90	1.9	207	207	141	141	238	246
		BR 116	112	1.7	209	213	127	135	238	240
		RP 81	260	1.6	201	207	121	127	238	238
		BR 147	159	1.9	205	209	127	127	248	252

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
5	ANTRIM	AL 55	70	2	207	213	129	141	242	246
		RP 40	130	1.8	205	219	127	135	244	246
		AL 119	70	1.9	203	205	127	?	238	250
		BR 200	90	1.8	205	209	135	?	232	240
		BR 43	91	1.9	195	205	127	127	240	240
		RP 23	54	1.9	207	213	127	135	244	244
		BR 248	70	1.9	209	209	127	127	238	244
6	ARENAC	BR 4	100	1.8	201	205	141	141	252	252
		BR 3	95	1.8	205	205	127	141	242	246
		BR 65	70	1.7	207	211	131	141	238	238
		BR 33	165	1.8	195	205	127	143	238	244
		BR 11	21	1.8	203	209	127	127	**NP**	**NP**
		BR 174	74	1.7	201	205	131	135	240	242
		BR 164	72	1.8	207	215	127	141	252	?
7	BARAGA	RP 210	25	1.8	203	205	127	127	238	240
		AL 110	85	1.8	205	205	127	141	242	242
		RP 199	15	1.8	195	205	127	129	242	250
		RP 136	25	1.9	211	221	135	135	244	248
		RP 157	78	1.8	213	219	127	127	240	246
8	BARRY	95029	435	1.9	195	205	127	127	244	244
		95025	163	1.8	205	215	141	143	244	248
		95032	85	1.9	185	205	127	141	238	240
		BR 100	145	1.9	205	207	129	143	236	236
		95023	140	1.8	203	203	127	141	238	248

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
9	BAY	BR 96	30	1.8	209	215	127	127	238	242
10	BENZIE	AL 146	72	1.9	195	203	121	127	238	244
		BR 127	66	1.9	205	209	137	141	244	248
		BR 229	15	1.9	203	205	127	127	**NP**	**NP**
		BR 118	70	1.9	195	211	127	135	250	250
		BR 50	101	1.9	203	213	131	141	238	244
		BR 223	74	1.8	203	209	127	135	250	250
11	BERRIEN	95304	80	1.6	203	205	127	127	244	256
		95306	70	1.6	203	209	121	127	252	256
		95307	80	1.6	195	205	127	141	244	244
		95308	50	1.7	203	205	127	127	238	246
		95309	68	1.7	195	207	127	127	238	244
12	BRANCH	JK 90	33	1.9	203	205	133	141	238	238
		JK 92	25	1.8	203	209	127	135	238	240
		95149	74	2	203	215	133	141	240	248
		95151	52	1.9	203	205	127	127	242	254
		95152	130	1.9	203	209	127	133	238	252
13	CALHOUN	JK 100	80	1.8	205	213	127	127	238	244
		95156	100	1.8	201	209	121	127	244	246
		95160	170	1.8	201	213	141	141	238	244
		95162	90	1.9	205	205	127	141	238	240
		95164	100	1.9	201	205	127	127	244	246

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
14	CASS	95312	172	1.8	205	207	127	141	238	244
		95313	80	1.9	203	205	127	141	244	248
		95316	135	1.9	185	209	127	127	238	246
		95319	35	1.8	185	201	127	127	238	242
		95321	70	1.8	203	205	127	135	242	248
15	CHARLEVOIX	RP 116	45	1.9	205	207	127	127	234	244
		BR 66	62	2	201	203	127	129	232	244
		RP 259	20	1.6	195	195	127	135	242	250
		AL 133	25	1.7	209	209	127	141	232	250
		RP 263	325	1.9	203	203	127	135	238	244
15B	BEAVER ISLAND	95299	90	1.8	203	209	125	125	244	244
		95284	58	2	207	219	127	127	246	250
		95286	134	1.8	203	203	127	133	252	256
		95289	50	1.8	203	203	127	127	244	256
		95295	125	1.9	203	219	127	141	244	252
		95290	125	1.9	203	205	127	127	252	258
16	CHEBOYGAN	BR 44	61	1.8	207	207	127	141	240	244
		BR 117	51	1.8	195	205	135	141	230	246
		BR 85	110	1.7	201	203	127	135	**NP**	**NP**
		BR 115	148	1.8	203	209	127	135	238	246
		BR 215	60	1.7	205	209	127	127	240	240
		RP 134	250	1.9	203	207	127	129	242	244

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
17	CHIPPEWA	RP 213	180	1.9	201	207	127	133	238	238
		AL 30	78	1.5	213	221	127	139	244	256
		RP 197	160	1.9	205	205	127	131	240	248
		RP 91	350	1.9	201	205	127	133	242	250
		BR 243	195	1.9	201	205	127	127	244	250
18	CLARE	RP 201	150	1.9	201	209	127	141	240	252
		RP 200	95	1.9	205	213	121	127	238	238
		AL 9	75	1.5	201	209	127	127	238	248
		AL 7	200	1.9	203	209	127	127	226	252
		RP 204	150	1.9	201	207	127	141	240	242
19	CLINTON	RP 53	375	1.9	205	205	135	143	238	244
20	CRAWFORD	BR 67	325	1.9	205	215	127	133	238	244
		AL 17	220	1.6	205	211	127	127	234	246
		AL 41	130	1.5	203	205	127	133	248	248
		BR 123	232	1.8	201	201	127	143	NP	NP
		AL 125	320	1.8	205	209	141	141	246	246
		BR 58	473	1.9	209	209	127	127	248	?
21	DELTA	BR 270	90	1.8	207	207	127	127	244	252
		BR 251	136	1.5	205	215	127	135	238	238
		BR 258	166	1.7	201	211	127	135	244	250
		RP 125	70	1.9	203	205	127	135	240	244
		RP 146	160	1.8	195	205	127	129	246	248
		AL 94	28	1.6	205	211	127	127	254	254

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
22	DICKINSON	AL 160	17	1.7	203	205	?	?	242	246
		AL 159	30	1.9	213	215	127	127	248	248
		AL 5	75	1.8	201	205	?	?	240	248
		RP 257	90	1.8	203	203	127	127	238	242
		95104	60	1.8	205	207	127	127	244	254
		RP 71	37	1.8	201	209	127	127	246	252
		RP 42	272	1.8	203	203	127	135	NP	NP
23	EATON	RL 1	55	1.9	203	213	127	133	238	252
		RL 8	48	1.8	205	209	121	127	244	246
		RL 18	115	2	195	209	127	127	238	242
		95064	360	1.9	195	205	127	135	246	248
		95063	250	1.9	201	215	133	141	242	246
		RL 3	30	1.7	195	203	127	141	238	238
24	EMMET	RP 190	35	1.8	195	201	127	127	238	248
		RP 151	25	1.8	201	203	127	145	232	238
		BR 98	35	1.8	203	213	127	135	240	240
		RP 150	16	1.8	203	213	127	137	242	248
		RP 273	38	1.8	203	207	127	131	240	246
	GD	BR 63	15	1.8	203	209	127	135	238	246
		BR 191	25	1.8	195	203	127	141	244	244

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
25	GENESEE	BR 256	130	1.9	203	209	127	141	238	252
		95138	70	1.9	203	209	127	127	240	246
		95125	320	1.9	203	209	127	127	236	250
		BR 257	200	2	201	205	135	135	238	242
		95147	255	1.9	195	201	141	141	248	252
26	GLADWIN	BR 189	28	1.8	203	203	127	143	238	252
		RP 100	8	1.8	203	213	127	141	238	244
		BR 255	20	1.7	201	205	127	135	244	244
		BR 22	480	1.9	203	209	133	151	226	238
		BR 16	480	1.8	201	205	127	135	226	232
27	GOGEBIC	RP 169	60	1.6	195	207	135	141	240	242
		AL 90	30	1.8	207	215	127	127	238	240
		RP 43	20	1.8	195	205	127	127	238	240
		RP 186	30	1.8	209	213	127	127	246	248
		RP 143	53	1.7	207	207	127	141	238	252
28	GRAND TRAVERSE	95013	70	1.8	195	203	127	127	246	250
		95014	140	1.8	195	209	127	129	240	250
		95015	110	1.9	195	211	127	141	238	252
		95018	300	1.8	195	205	127	127	230	246
		95020	100	2	195	203	127	135	244	254

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
29	GRATIOT	AL 149	25	1.8	201	205	127	141	240	252
		95124	70	1.6	203	205	127	135	238	242
		95130	116	1.6	203	203	127	143	248	248
		95131	132	1.7	209	215	127	127	244	244
		AL 2	135	2.3	201	203	141	141	238	240
30	HILLSDALE	JK 83	50	1.7	203	203	127	141	240	252
		95171	135	1.7	205	205	133	141	248	252
		95176	54	1.7	203	209	127	141	238	244
		95178	80	1.7	203	213	127	135	252	252
		95179	105	1.8	201	203	127	141	244	246
		95168	51	1.8	209	221	127	141	244	254
31	HOUGHTON	AL 45	80	1.9	203	211	127	141	236	250
		RP 30	295	1.8	213	?	127	127	252	254
		RP 29	175	2	205	209	127	141	246	262
		RP 215	120	1.7	209	211	127	127	240	242
		RP 198	102	1.9	205	209	127	133	244	244
		RP 57	414	1.9	203	205	135	141	246	248
32	HURON	SP 2	70	1.7	209	211	127	127	250	252
		SP 4	70	1.7	209	221	127	127	246	252
		95180	85	1.8	195	203	127	141	248	248
		SP 6	50	1.6	203	205	127	133	242	244
		SP 5	70	1.6	203	205	127	133	240	244

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
33	INGHAM	95070	105	1.8	205	205	129	135	244	246
		95132	70	1.9	205	209	129	141	232	238
		95136	125	1.8	205	209	131	141	244	244
		95146	85	1.9	205	207	121	141	244	252
		95071	50	1.6	205	205	127	?	238	242
34	IONIA	95137	265	1.7	203	203	127	127	242	244
		95181	98	1.9	203	203	127	127	238	248
		95183	95	2	195	203	127	141	238	246
		95185	60	1.8	205	205	121	135	242	248
		95188	68	1.9	205	205	127	131	242	248
35	IOSCO	BR 181	75	1.8	201	213	131	133	238	246
		BR 210	85	1.8	205	205	127	141	240	252
		BR 73	48	1.8	209	213	127	135	242	248
		BR 57	120	1.9	201	213	141	141	230	252
		BR 95	32	1.5	209	209	127	131	238	240
		BR 108	63	1.8	203	213	129	141	238	254
36	IRON	AL 22	115	1.8	203	209	135	139	244	252
		RP 97	110	1.9	205	209	127	141	246	248
		BR 159	218	1.8	195	205	127	127	234	238
		AL 104	240	1.9	203	207	133	135	238	240
		AL 74	250	1.9	195	203	127	135	238	240
		AL 106	353	1.9	211	211	135	139	226	246

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
37	ISABELLA	BR 104	228	1.8	203	207	127	127	238	238
		AL 132	267	1.9	203	205	127	141	238	238
		AL 131	276	1.9	207	215	133	135	246	246
		BR 62	272	1.9	205	209	135	141	238	246
		AL 11	278	1.8	209	215	127	141	238	246
38	JACKSON	JK 81	118	1.8	195	203	127	141	242	244
		95192	100	1.8	203	205	127	127	244	248
		95194	90	1.9	201	203	127	141	244	246
		95200	85	1.9	205	205	127	127	238	244
		JK 76	124	1.7	203	213	127	127	244	252
39	KALAMAZOO	FCTC 3	155	1.7	203	203	127	141	244	248
		FTC 1	145	1.7	205	209	127	135	242	244
		FTC 6	200	2	209	209	141	141	244	244
		PW 105	92	1.8	205	215	127	127	238	252
		PW 104	65	1.7	195	203	127	141	244	244
		PW 109	65	1.7	207	207	127	127	240	244
40	KALKASKA	AL 164	150	2	205	209	127	127	238	248
		BR 226	163	1.9	195	209	127	135	240	242
		BR 131	160	1.9	203	213	135	135	242	242
		RP 241	127	1.8	209	213	127	135	238	240
		AL 43	136	1.9	205	205	131	141	242	244

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
41	KENT	RP 243	95	1.9	203	213	127	131	244	244
		95206	160	1.8	203	209	127	141	242	250
		95209	100	1.9	203	205	127	127	244	246
		95211	60	1.9	201	203	127	141	238	244
		95213	50	1.9	195	203	127	127	238	254
42	KEWEENAW									
43	LAKE	AL 156	58	1.7	209	215	127	127	244	248
		RP 182	170	2	203	209	131	141	238	238
		BR 111	225	1.8	195	205	127	141	238	244
		AL 153	160	1.9	203	213	127	127	238	246
		AL 137	170	1.8	201	209	127	127	238	244
44	LAPEER	BR 160	440	1.9	203	205	127	127	250	?
		95098	38	1.6	203	209	127	127	240	252
		95128	290	1.6	203	205	133	143	242	244
		95097	61	1.6	195	205	127	141	244	246
45	LEELANAU	BR 221	280	1.7	203	219	127	131	232	240
		RP 28	110	1.9	195	201	127	129	232	254
		95005	45	1.8	203	215	135	143	230	252
		RP 167	125	1.8	203	205	127	133	?	?
		RP 209	128	1.6	201	209	127	135	238	254
		95009	50	1.7	205	209	129	141	250	250

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
45B	N. MANITU ISLAND	95043	50	1.8	213	221	127	133	240	244
		95038	66	1.9	207	213	135	135	244	244
		95036	140	1.8	195	209	127	129	244	252
		95042	90	1.8	195	221	133	135	244	244
		95049	156	1.8	205	207	127	135	240	242
		95045	93	1.8	205	221	129	135	242	242
46	LENAWEE	95072	105	1.8	205	209	127	141	238	252
		95215	55	1.7	201	203	135	141	**NP**	**NP**
		95217	80	1.7	203	209	131	135	244	248
		95220	40	1.6	203	205	127	143	238	248
		95222	78	1.8	201	209	121	127	244	244
		JK 97	40	1.7	205	213	135	141	238	240
		95216	40	1.6	203	203	141	143	**NP**	**NP**
		95223	38	1.6	205	209	141	141	**NP**	**NP**
47	LIVINGSTON	BR 206	55	1.9	195	213	127	127	244	246
		RL 11	55	1.8	213	221	127	133	246	248
		95101	585	1.7	205	209	133	141	238	242
		95106	145	1.9	203	205	127	141	230	240
		95145	80	1.8	205	209	141	141	238	238
48	LUCE	RP 121	328	1.9	205	213	127	131	238	252
		RP 86	158	1.8	203	215	127	127	246	248
		AL 108	230	1.9	203	211	127	141	242	254
		BR 167	275	1.9	207	207	127	127	246	250
		RP 188	460	1.9	205	205	127	127	238	244

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
49	MACKINAC	RP 152	215	2	203	207	127	127	250	252
		BR 101	225	1.9	203	209	141	141	242	244
		BR 201	336	1.9	207	213	135	141	246	248
		RP 249	212	1.9	205	209	127	127	238	252
		AL 85	195	1.9	203	205	127	141	244	252
50	MACOMB	RP 233	406	1.8	205	211	127	141	238	246
51	MANISTEE	BR 139	166	1.9	195	213	127	127	238	246
		RP 225	300	1.8	203	205	127	131	238	246
		BR 48	190	1.8	203	215	127	127	242	252
		AL 40	275	1.8	207	207	127	133	238	248
		RP 65	230	1.8	209	213	127	133	238	240
52	MARQUETTE	BR 262	146	1.9	205	215	129	141	238	246
		AL 54	247	1.9	195	221	131	135	242	248
		AL 88	215	1.9	211	221	127	135	248	252
		RP 170	235	1.9	205	205	127	127	242	246
		RP 10	265	1.9	207	211	129	129	246	252
53	MASON	RP 211	173	1.9	203	219	127	127	238	238
		RP 214	222	1.8	203	209	127	127	?	?
		AL 62	222	1.9	209	219	127	127	238	244
		AL 66	302	1.8	203	219	127	127	238	250
		RP 192	382	1.8	209	219	127	127	240	246
		BR 244	555	1.8	203	215	127	127	238	252

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
54	MECOSTA	RP 11	370	1.8	203	209	135	141	238	246
		BR 219	172	1.9	203	205	127	135	238	248
		RP 14	139	1.9	203	205	127	129	244	252
		BR 197	212	1.9	195	209	131	141	238	240
		RP 141	200	1.8	209	209	141	141	238	242
55	MENOMINEE	BR 142	230	1.7	**NP**	**NP**	127	129	242	248
		BR 92	206	1.9	209	209	127	127	238	256
		BR 141	174	1.9	201	207	127	133	240	244
		RP 46	188	1.9	203	205	127	127	238	246
		BR 124	45	1.9	195	207	127	135	242	242
		RP 48	490	1.9	205	207	127	127	238	250
56	MIDLAND	BR 103	205	1.8	205	209	127	129	226	244
		AL 161	250	1.9	207	209	127	127	238	242
		AL 148	275	1.9	203	209	129	141	244	250
		AL 100	240	1.9	195	209	127	151	238	242
		BR 77	700	1.9	203	207	127	143	230	242
57	MISSAUKEE	AL 107	240	1.8	205	213	127	141	242	246
		BR 183	208	1.9	201	209	127	143	232	252
		BR 86	135	1.8	195	195	127	143	238	238
		RP 77	170	1.9	195	203	129	141	232	242
		BR 151	270	1.9	203	213	133	141	238	240

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
58	MONROE	95075	205	1.8	203	205	127	135	242	252
		95076	130	1.9	209	213	127	127	236	250
		95079	127	1.9	205	217	135	135	250	252
		95084	80	1.9	209	215	127	141	242	246
		95085	120	1.9	203	205	127	135	246	248
		95087	116	1.9	209	209	127	133	246	254
59	MONTCALM	95224	90	1.9	201	209	135	141	238	250
		95226	168	1.7	?	207	127	135	238	238
		95229	50	1.7	203	205	133	141	238	238
		95231	66	1.8	?	207	141	141	242	248
		95232	40	1.9	203	209	127	135	230	238
		95228	92	1.7	203	205	127	127	244	252
		95233	138	2	203	215	127	141	238	244
		95225	88	1.8	205	209	127	135	238	238
60	MONTMORENCY	BR 68	328	1.7	209	209	133	135	240	250
		RP 267	178	1.9	203	203	127	133	242	252
		BR 88	188	1.8	195	?	127	127	228	238
		RP 140	220	1.9	203	203	127	141	244	252
		BR 91	490	1.9	201	205	127	133	232	232
		BR 79	600	1.8	203	207	127	141	242	244
61	MUSKEGON	95241	208	1.9	203	213	127	133	244	244
		95239	195	1.9	203	209	133	133	244	246
		95243	190	1.8	195	213	127	141	238	244
		95236	165	1.9	209	209	127	127	238	244
		95242	290	1.8	205	209	127	133	230	244

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
62	NEWAYGO	RP 73	125	2	205	219	127	133	238	244
		RP 179	310	1.8	207	?	127	127	244	244
		RP 132	245	1.9	205	209	127	127	244	246
		RP 128	385	1.9	209	213	127	127	238	244
		RP 195	348	1.9	203	213	127	135	238	254
		RP 33	762	1.7	203	205	127	133	244	244
63	OAKLAND	95091	50	1.7	195	203	127	137	238	242
		95093	46	1.8	209	211	129	141	244	248
		95094	76	1.9	205	209	137	137	240	256
		95092	84	2	203	209	127	135	248	254
		95090	68	1.8	205	213	127	127	240	244
		95089	60	2	205	205	127	127	238	248
64	OCEANA	95253	346	1.9	203	209	127	127	244	246
		95245	180	1.8	195	213	133	141	244	244
		95250	360	1.9	?	205	127	141	238	252
		RP 6	195	1.8	203	203	127	127	238	244
		RP 232	210	1.8	195	209	127	141	238	238
		95252	572	1.9	203	209	127	127	244	246
65	OGEMAW	BR 89	252	1.7	205	213	127	127	232	244
		BR 14	235	1.8	207	213	127	133	244	246
		BR 19	510	1.8	199	199	127	127	238	240
		BR 84	385	1.9	205	209	127	133	242	242
		BR 259	208	1.8	209	213	127	129	246	252

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
66	ONTONAGON	BR 83	160	1.9	205	219	127	135	238	244
		RP 32	242	1.8	201	201	127	135	236	246
		RP 119	488	1.9	201	211	127	127	238	242
		RP 268	350	1.9	195	205	127	135	238	246
		RP 138	830	1.8	205	215	127	141	238	246
67	OSCEOLA	RP 218	264	1.6	195	195	127	141	238	238
		RP 105	168	1.7	207	209	133	141	244	244
		RP 108	455	1.7	185	203	127	135	250	252
		RP 103	218	1.7	209	209	133	135	238	244
		RP 58	285	1.8	195	213	127	135	240	252
		RP 224	328	1.9	203	203	127	143	240	240
68	OSCODA	BR 106	100	1.9	201	203	127	127	242	242
		BR 110	100	1.8	213	213	127	143	240	246
		BR 21	50	1.8	195	209	127	127	244	244
		BR 245	120	2	201	201	141	141	250	252
		BR 36	30	1.8	195	205	127	141	252	252
69	OTSEGO	BR 250	345	1.8	203	209	127	127	244	250
		BR 192	350	1.8	201	205	127	127	232	242
		BR 180	250	1.8	207	209	127	143	238	238
		BR 242	350	1.9	201	215	129	135	242	242
		BR 145	510	1.9	195	207	129	135	244	246

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
70	OTTAWA	95254	50	1.8	203	209	127	127	238	242
		95257	35	1.8	201	209	127	145	238	250
		95259	85	1.7	203	205	127	133	244	244
		95263	22	1.7	203	209	127	127	240	244
		95255	30	1.9	203	203	127	135	254	254
		95256	17	1.5	205	209	135	141	238	238
71	PRESQUE ISLE	AL 32	470	1.7	207	215	127	129	240	252
		BR 171	140	1.9	201	207	127	141	238	242
		BR 234	135	1.9	195	211	127	141	242	248
		BR 231	210	1.8	195	205	131	133	242	252
		BR 240	55	1.8	205	209	129	143	240	246
72	ROSCOMMON	RP 55	194	1.8	201	205	127	127	240	242
		RP 79	166	1.8	195	201	127	135	240	252
		BR 239	400	1.9	201	203	137	141	248	252
		BR 199	404	1.8	205	213	127	141	248	250
		RP 19	410	1.8	195	209	127	135	242	248
73	SAGINAW	95074	68	1.7	205	209	121	127	242	242
		AL 13	225	1.8	201	201	127	133	238	242
		BR 236	420	1.9	205	209	133	133	244	244
		AL 12	288	1.9	203	203	127	131	238	246
		AL 109	275	1.9	203	203	127	127	244	254

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
74	ST. CLAIR	LI 04	65	1.8	203	205	127	133	244	252
		LI 06	80	1.8	205	205	127	141	240	244
		LI 07	76	1.8	195	215	127	135	248	252
		LI 08	105	1.7	203	209	129	129	244	252
		LI 01	90	1.6	205	205	127	129	242	250
75	ST. JOSEPH	95328	140	1.9	203	205	127	141	242	246
		95331	165	1.7	203	205	135	141	238	244
		95334	70	1.8	205	207	127	141	246	254
		95330	118	1.6	195	203	127	143	238	252
		95329	80	1.7	203	205	129	141	242	252
76	SANILAC	95264	322	1.9	203	205	127	141	?	?
77	SCHOOLCRAFT	RP 265	380	1.9	209	215	127	131	236	246
		RP 115	440	1.8	205	213	127	149	242	244
		RP 264	410	1.9	205	209	127	127	240	248
		AL 3	165	1.5	205	213	127	127	240	252
		AL 97	300	1.9	201	205	127	141	238	248

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
78	SHIAWASSEE	RL 14	85	1.8	195	209	127	127	232	244
		95069	95	1.8	205	217	127	127	?	?
		95123	110	1.8	203	205	127	133	238	238
		95134	50	2	201	203	127	127	248	248
		95119	46	1.9	205	209	127	133	238	238
		RL 13	15	1.9	205	215	127	141	248	252
		RL 7	64	1.9	205	215	127	135	242	248
		RL 22	68	1.9	205	207	127	141	244	246
		95068	40	1.8	195	203	127	141	244	248
79	TUSCOLA	BR 113	284	1.8	205	209	127	135	240	252
		BR 70	338	1.9	195	221	127	135	242	242
		BR 2	316	1.9	203	209	141	141	244	252
		95268	370	1.8	203	203	127	127	248	248
		BR 38	315	1.9	203	205	127	135	240	244
80	VAN BUREN	PW 103	65	1.7	205	209	121	127	240	244
		95111	220	1.6	203	205	127	127	238	240
		95114	100	1.6	205	205	127	141	248	250
		95336	50	1.6	203	205	135	151	248	256
		95338	95	1.8	205	205	141	141	242	246

Table 7. (cont'd).

COUNTY NUMBER	COUNTY NAME	SAMPLE	CONC. (ng/uL)	O.D. (260/280)	OBCAM-1	OBCAM-2	IGF1-1	IGF1-2	CRFA-1	CRFA-2
81	WASHTENAW	JK 86	45	1.9	203	203	129	135	238	252
		95272	95	1.8	195	203	127	141	238	?
		95275	55	1.8	205	207	129	141	238	244
		95278	35	1.8	205	209	127	141	?	?
		95280	35	1.7	209	209	141	141	240	250
		BR 158	91	1.8	207	209	127	141	244	248
		JK 99	56	1.7	195	209	127	127	248	252
		95274	30	1.8	203	203	127	141	?	?
		95277	90	1.8	195	203	121	141	242	248
82	WAYNE	95057	40	1.7	209	209	121	129	238	240
		95058	50	1.8	203	209	127	135	238	246
		95082	80	1.7	209	217	127	133	250	?
		95086	100	1.8	203	221	127	127	248	252
		95088	60	1.8	213	213	129	133	244	250
		95083	67	1.8	201	209	127	135	250	250
83	WEXFORD	RP 24	120	1.9	195	205	127	127	240	250
		RP 114	452	1.9	195	205	127	135	232	238
		RP 235	702	1.9	205	215	127	127	238	244
		BR 260	430	1.9	205	205	135	143	240	242
		RP 178	250	1.8	205	211	127	131	230	244
		RP 106	525	1.9	203	203	127	141	230	244

? means allele unable to be scored confidently. NP means no amplified product. Allele sizes are in base pairs.

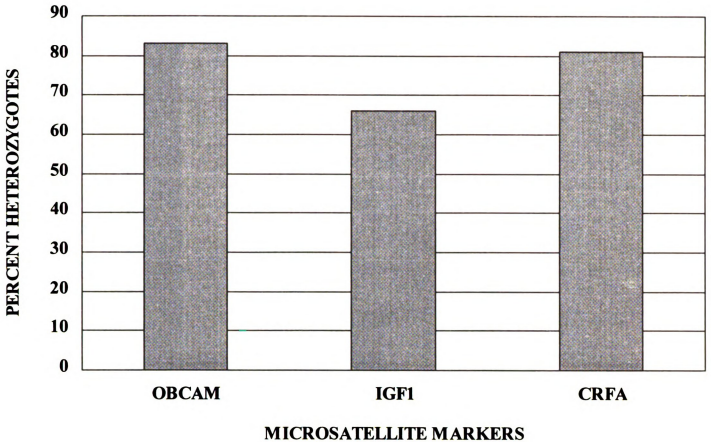


Figure 17. Heterozygosity of the three microsatellite markers for Michigan white-tailed deer.

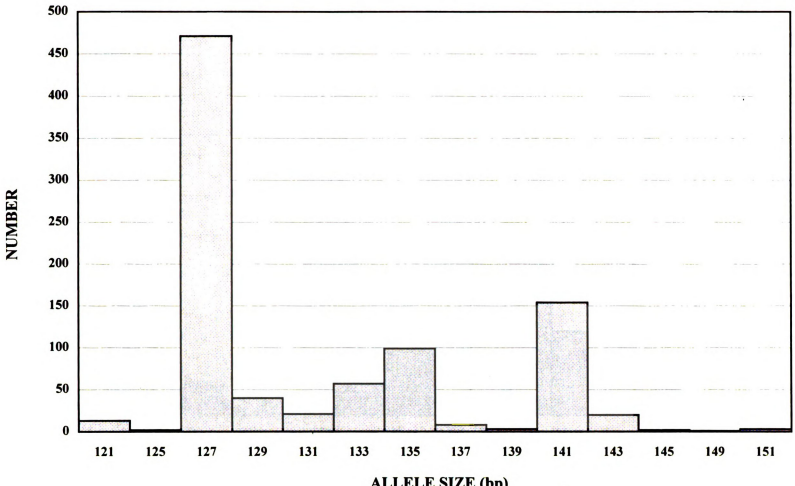


Figure 18. Graphical representation of IGF1 alleles in the database.

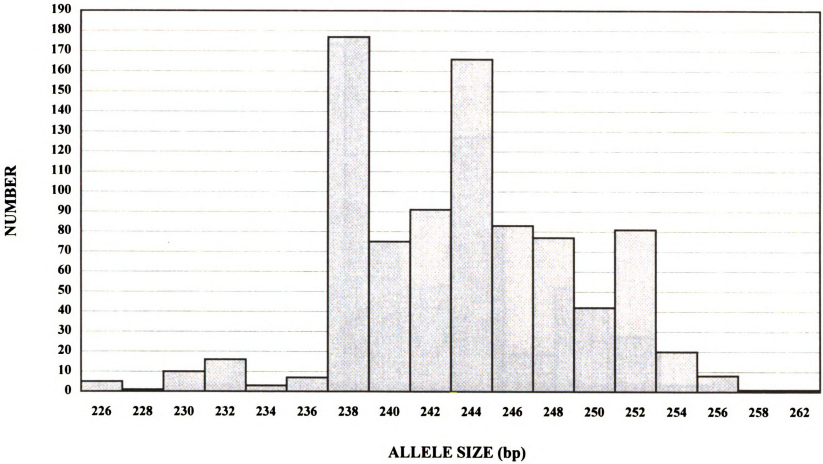


Figure 19. Graphical representation of CRFA alleles in the database.

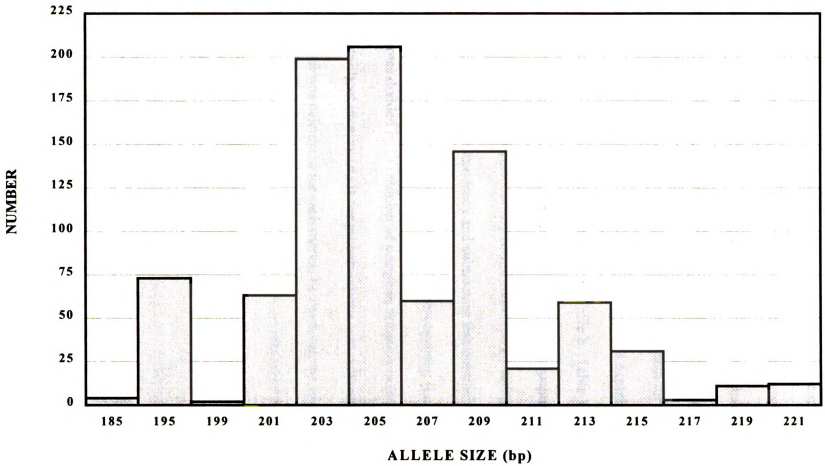


Figure 20. Graphical representation of OBCAM alleles in the database.

Table 8 lists the frequency of alleles, heterozygosity data, PIC values, Fisher Exact test probabilities, and the probability of identity for the three markers present in the database of 450 deer samples. The data presented in this table is based on only complete genotypic information for each sample and marker, with incomplete genotypes being eliminated from further testing. If a sample had an incomplete genotype, where one or both alleles was unable to be scored for a locus, then the data for that marker was not included in the frequency data or in the calculations. This explains the difference in the number of complete samples for each marker in Table 8. Table 7 contains all the complete and incomplete genotypic information for the test population. The expected heterozygosity under Hardy-Weinberg equilibrium (HWE) was calculated by summing the squares of each allele frequency and subtracting that number from one. This calculation was used to compare the observed heterozygosity (i.e. observed number of heterozygotes divided by the number of complete samples) in the sampled population. The probability of identity (P_i) is the probability that two individuals will have the same genotype and was calculated for each marker by summing the squares of each genotype frequency (National Research Council, 1996). This probability was calculated in order to determine the power of discrimination (P_d) for this microsatellite system, which by multiplying the P_i for each marker and subtracting from one resulted in a P_d of 0.99984 or 99.98 %. The P_d signifies the ability of the system to discriminate between individuals with this degree of probability. The Fisher Exact test, or probability test, tests for exactness to Hardy-Weinberg equilibrium by rejecting the hypothesis (i.e. the population of white-tailed deer sampled in this study are in HWE) if the observed genotypic frequencies deviate from what is expected under HWE (Weir, 1990) and was calculated

by the software program GENEPOP V3.1d (Raymond and Rousset, 1995). The Fisher Exact test was selected to test the population for HWE because of the possibility of rare alleles occurring in a rather large sampled population, such that, the frequency of the alleles and genotypes would be too small for the chi-square statistic. Even though the chi-square test statistic was also calculated by this program using the same inputed genotypic data, both of these tests resulted in compatible outcomes at the 0.004 level of significance for each test after correcting to account for multiple testing. This correction was done using the Bonferroni procedure which, simply stated, is the desired significance level (0.05) divided by the number of tests conducted. Along with the overall HWE testing for the whole population of sampled deer in Michigan, Table 9 also contains the statistical data testing Hardy-Weinberg equilibrium for four subdivisions of Michigan in order to assess regional differences in HWE.

Table 8. Summary data of all 450 sampled Michigan white-tailed deer.

Number of Alleles	CRFA		OBCAM		IGF1	
	Allele Size (bp)	Number	Allele Size (bp)	Number	Allele Size (bp)	Number
1	226	5	185	4	121	13
2	228	1	195	73	125	2
3	230	10	199	2	127	471
4	232	16	201	63	129	40
5	234	3	203	199	131	21
6	236	7	205	206	133	57
7	238	177	207	60	135	99
8	240	75	209	146	137	8
9	242	91	211	21	139	3
10	244	166	213	59	141	154
11	246	83	215	31	143	20
12	248	77	217	3	145	2
13	250	42	219	11	149	1
14	252	81	221	12	151	3
15	254	20				
16	256	8				
17	258	1				
18	262	1				
# of Complete Samples		432		445		447
Total # of Alleles		864		890		894
# of Different Alleles		18		14		14
Observed # of Heterozygous Samples		350		367		294
Expected # of Heterozygous Samples		376		378		299
Observed Heterozygosity		0.81		0.83		0.66
Expected Heterozygosity		0.87		0.85		0.67
Fisher Exact Test Probability		0.0555		0.3722		0.2424
Polymorphic Information Content (PIC)		0.86		0.83		0.64
Probability of Identity (P_i)		0.0287		0.0408		0.1357
Allele Size Range		226-262		185-221		121-151
Power of Discrimination (P_d) for this system is 0.9998 or 99.98%						

Table 9. Statistical data testing Hardy-Weinberg equilibrium of Michigan white-tailed deer population subdivided into four regions of Michigan.

POPULATION	MARKER	FISHER EXACT TEST PROBABILITY	OVERALL CHI-SQUARE (X^2)	CHI-SQUARE PROBABILITY
All Michigan	CRFA IGF1 OBCAM	0.0555 0.2424 0.3722	10.6	0.1018
Upper Peninsula	CRFA IGF1 OBCAM	0.5811 0.6217 0.1983	5.3	0.5095
Southern Lower Michigan	CRFA IGF1 OBCAM	0.1723 0.2545 0.4555	7.8	0.2511
North East Lower Michigan	CRFA IGF1 OBCAM	0.0131 0.2919 0.1016	15.7	0.0154
North West Lower Michigan	CRFA IGF1 OBCAM	0.0751 0.4162 0.5362	8.2	0.2254

*** Significant at P = 0.004 after Bonferroni correction.**

To identify possible substructuring or inbreeding within the total sampled population, it is necessary to subdivide Michigan into regions for an assessment of the structuring of the deer population. Tables 10, 11, 12, 13 and 14 show the distribution of alleles among the subdivided regions of Michigan and the contributing counties are listed below each table. Table 10 contains data from all 31 counties south of M-57 and including the 3 counties in the thumb region north of M-57 (i.e. southern lower Michigan). Table 11 contains data from all 39 counties north of M-57 excluding the 3 counties in the thumb region (i.e. northern lower Michigan). Table 12 contains data from all 15 counties in the upper peninsula (i.e. upper peninsula). Lastly, northern Michigan was further subdivided into the east and west and Tables 13 and 14, respectively, contain data from those counties listed beneath the tables. Michigan and thus, the database was subdivided this way in order to roughly identify possible regions of substructuring or inbreeding. The selection was not geographically based on natural barriers (e.g. streams, steep hills, highways, cities, etc.), but based on the DNR's accepted subdivision of Michigan. Therefore, the interpretation of the data can only be an estimate of what is possibly occurring in those regions. A more accurate evaluation of the data would take into account other factors, outside the realm of this study, when subdividing the deer population in Michigan.

Establishing a database for future forensic applications in paternity testing or identification of unknown samples for the purpose of matching them to evidentiary material requires the use of allele frequencies of the markers in this system in order to calculate an accurate probability of a match. Along with the previous tables listing the number of occurrences of each allele for each marker for each region of Michigan, Table

15 shows the frequencies of each allele for each marker for each region. This presentation of the data allows for a visual comparison of frequencies for each marker across regions, with a noticeable variation between regions for many alleles. Statistical testing of this inter-region variance in allele frequencies was performed using the F-statistic (F_{st}) and the chi-square and are also listed in this table. Statistical analysis using the F-statistic was performed by GENEPOP V3.1d software with this statistic measuring the variance in the genetic structure (allelic frequencies) in the subdivided population. The chi-square approximation for F_{st} was calculated using the formula $2N \times F_{st} \times (K-1)$, with degrees of freedom = (S-1) x (K-1) where N = # of samples, K = # of alleles and S = # of populations (Workman & Niswander, 1970). A significant chi-square would be indicative of heterogeneity in allele frequencies. Table 16 shows the comparison between regions for variance in allele frequencies between each other using the F-statistic and the chi-square. The inter-region variance in allele frequencies is low, but after taking into account the large sample size, the chi-square values for all but three comparisons become significant. The level of significance for the F-statistic and chi-square has been corrected in order to account for multiple testing, which again was done using the Bonferroni procedure.

Table 10. Southern Michigan lower peninsula white-tailed deer data.

Number of Alleles	CRFA		OBCAM		IGF1	
	Allele Size (bp)	Number	Allele Size (bp)	Number	Allele Size (bp)	Number
1	230	1	185	3	121	9
2	232	2	195	23	127	159
3	236	4	201	16	129	13
4	238	57	203	87	131	4
5	240	21	205	95	133	18
6	242	28	207	10	135	30
7	244	68	209	54	137	3
8	246	26	211	3	141	73
9	248	38	213	13	143	7
10	250	12	215	8	145	1
11	252	31	217	3	151	1
12	254	8	221	5		
13	256	4				
# of Complete Samples		150		160		159
Total # of Alleles		300		320		318
Observed # of Heterozygous Samples		127		130		107
Expected # of Heterozygous Samples		129		128		108
Observed Heterozygosity		0.85		0.81		0.67
Expected Heterozygosity		0.86		0.80		0.68
Fisher Exact Test Probability		0.1723		0.4555		0.2545

Counties: Allegan, Barry, Berrien, Branch, Calhoun, Cass, Clinton, Eaton, Genessee, Hillsdale, Huron, Ingham, Ionia, Jackson, Kalamazoo, Kent, Lapeer, Lenawee, Livingston, Macomb, Monroe, Oakland, Ottawa, Sanilac, Shiawassee, St. Clair, St. Joseph, Tuscola, Van Buren, Washtenaw, Wayne.

Table 11. Northern Michigan lower peninsula white-tailed deer data.

Number of Alleles	CRFA		OBCAM		IGF1	
	Allele Size (bp)	Number	Allele Size (bp)	Number	Allele Size (bp)	Number
1	226	4	185	1	121	4
2	228	1	195	40	125	2
3	230	9	199	2	127	214
4	232	14	201	36	129	21
5	234	2	203	89	131	14
6	238	91	205	69	133	33
7	240	40	207	33	135	48
8	242	46	209	77	137	5
9	244	80	211	7	141	65
10	246	36	213	34	143	13
11	248	21	215	16	145	1
12	250	22	219	9	151	2
13	252	34	221	3		
14	254	7				
15	256	2				
16	258	1				
# of Complete Samples		205		208		211
Total # of Alleles		410		416		422
Observed # of Heterozygous Samples		155		172		145
Expected # of Heterozygous Samples		178		179		148
Observed Heterozygosity		0.76		0.83		0.69
Expected Heterozygosity		0.87		0.86		0.70
Fisher Exact Test Probability		0.0034		0.3434		0.1990

Counties: Alcona, Alpena, Antrim, Arenac, Bay, Beaver Island, Benzie, Charlevoix, Cheboygan, Clare, Crawford, Emmet, Gladwin, Grand Traverse, Gratiot, Iosco, Isabella, Kalkaska, Lake, Leelanau, Manistee, Mason, Mecosta, Midland, Missaukee, Montcalm, Montmorency, Muskegon, Newaygo, North Manitu Island, Oceana, Ogemaw, Osceola, Oscoda, Otsego, Presque Isle, Roscommon, Saginaw, Wexford.

Table 12. Michigan upper peninsula white-tailed deer data.

Number of Alleles	CRFA		OBCAM		IGF1	
	Allele Size (bp)	Number	Allele Size (bp)	Number	Allele Size (bp)	Number
1	226	1	195	9	127	96
2	234	1	201	11	129	6
3	236	3	203	21	131	4
4	238	28	205	42	133	5
5	240	14	207	17	135	20
6	242	17	209	14	139	3
7	244	16	211	11	141	15
8	246	20	213	12	149	1
9	248	18	215	7		
10	250	8	219	2		
11	252	16	221	4		
12	254	5				
13	256	2				
14	262	1				
# of Complete Samples		75		75		75
Total # of Alleles		150		150		150
Observed # of Heterozygous Samples		67		62		40
Expected # of Heterozygous Samples		67		65		42
Observed Heterozygosity		0.89		0.83		0.53
Expected Heterozygosity		0.89		0.86		0.56
Fisher Exact Test Probability		0.5811		0.1983		0.6217

Counties: Alger, Baraga, Chippewa, Delta, Dickinson, Gogebic, Houghton, Iron, Keweenaw, Luce, Mackinac, Marquette, Menominee, Ontonagon, Schoolcraft.

Table 13. North East Michigan lower peninsula white-tailed deer data.

Number of Alleles	CRFA		OBCAM		IGF1	
	Allele Size (bp)	Number	Allele Size (bp)	Number	Allele Size (bp)	Number
1	226	4	195	11	121	3
2	228	1	199	2	127	5
3	230	2	201	28	129	12
4	232	5	203	34	131	6
5	234	1	205	35	133	16
6	238	36	207	23	135	17
7	240	19	209	35	137	3
8	242	30	211	4	141	36
9	244	27	213	14	143	8
10	246	19	215	10	151	2
11	248	13				
12	250	6				
13	252	20				
14	254	3				
# of Complete Samples		93		98		99
Total # of Alleles		186		196		198
Observed # of Heterozygous Samples		69		79		71
Expected # of Heterozygous Samples		82		84		71
Observed Heterozygosity		0.74		0.81		0.72
Expected Heterozygosity		0.88		0.86		0.72
Fisher Exact Test Probability		0.0131		0.1016		0.2919

Counties: Alcona, Alpena, Arenac, Bay, Cheboygan, Clare, Crawford, Gladwin, Gratiot, Iosco, Isabella, Midland, Montmorency, Ogemaw, Oscoda, Otsego, Presque Isle, Roscommon, Saginaw.

Table 14. North West Michigan lower peninsula white-tailed deer data.

Number of Alleles	CRFA		OBCAM		IGF1	
	Allele Size (bp)	Number	Allele Size (bp)	Number	Allele Size (bp)	Number
1	230	6	185	1	121	1
2	232	9	195	29	125	2
3	234	1	201	8	127	119
4	238	55	203	55	129	9
5	240	21	205	34	131	8
6	242	16	207	10	133	17
7	244	53	209	42	135	31
8	246	20	211	3	137	2
9	248	8	213	20	141	29
10	250	16	215	6	143	5
11	252	14	219	9	145	1
12	254	4	221	3		
13	256	2				
14	258	1				
# of Complete Samples		113		110		112
Total # of Alleles		226		220		224
Observed # of Heterozygous Samples		86		93		74
Expected # of Heterozygous Samples		96		94		75
Observed Heterozygosity		0.76		0.85		0.66
Expected Heterozygosity		0.85		0.85		0.67
Fisher Exact Test Probability		0.0751		0.5362		0.4162

Counties: Antrim, Beaver Island, Benzie, Charlevoix, Emmet, Grand Traverse, Kalkaska, Lake, Leelanau, Manistee, Mason, Mecosta, Missaukee, Montcalm, Muskegon, Newaygo, North Manitu Island, Oceana, Osceola, Wexford.

Table 15. Frequency of alleles for each marker in the four regions of Michigan.

CRFA					OBCAM					IGF1				
Allele Size (bp)	**North East Lower**	**North West Lower**	**South Lower**	**U.P.**	**Allele Size (bp)**	**North East Lower**	**North West Lower**	**South Lower**	**U.P.**	**Allele Size (bp)**	**North East Lower**	**North West Lower**	**South Lower**	**U.P.**
226	0.021			0.007	**185**		0.005	0.009		**121**	0.015	0.004	0.028	
228	0.005				**195**	0.056	0.132	0.072	0.060	**125**		0.009		
230	0.016	0.027	0.003		**199**	0.010				**127**	0.480	0.531	0.500	0.640
232	0.027	0.041	0.007		**201**	0.143	0.036	0.050	0.073	**129**	0.061	0.040	0.041	0.040
234	0.005	0.005		0.007	**203**	0.173	0.250	0.272	0.140	**131**	0.030	0.036	0.013	0.027
236			0.013	0.020	**205**	0.179	0.155	0.297	0.280	**133**	0.081	0.076	0.057	0.033
238	0.191	0.248	0.190	0.187	**207**	0.117	0.045	0.031	0.113	**135**	0.086	0.138	0.094	0.133
240	0.101	0.095	0.070	0.093	**209**	0.179	0.191	0.169	0.093	**137**	0.015	0.009	0.009	
242	0.160	0.072	0.093	0.113	**211**	0.020	0.014	0.009	0.073	**139**				0.020
244	0.144	0.239	0.227	0.107	**213**	0.071	0.091	0.041	0.080	**141**	0.182	0.129	0.230	0.100
246	0.106	0.072	0.087	0.133	**215**	0.051	0.027	0.025	0.047	**143**	0.040	0.022	0.022	
248	0.069	0.036	0.127	0.120	**217**			0.009		**145**		0.004	0.003	
250	0.032	0.072	0.040	0.053	**219**		0.041		0.013	**149**				0.007
252	0.106	0.063	0.103	0.107	**221**		0.014	0.016	0.027	**151**	0.010		0.003	
254	0.016	0.018	0.027	0.033										
256		0.009	0.013	0.013										
258		0.005												
262				0.007										
Fst for CRFA = 0.0069					**Fst for OBCAM = 0.0168**					**Fst for IGF1 = 0.0091**				
Chi-Square = 101.3 (*df* = 51) *					**Chi-Square = 194.4 (*df* = 39) ***					**Chi-Square = 105.8 (*df* = 39) ***				

Blank spaces = absence of that allele in that region of Michigan.

*** Significant at P = 0.004 after Bonferroni correction.**

Table 16. F_{st} and Chi-Square values from comparing the frequency of alleles in the four regions of Michigan.

Regions of Michigan		North East Lower		North West Lower		Southern Lower	
	Markers	F_{st}	Chi-Square (X^2)	F_{st}	Chi-Square (X^2)	F_{st}	Chi-Square (X^2)
North West Lower	CRFA	0.0097	59.655 (*df* 15) *				
	IGF1	0.0019	8.820 (*df* 11)				
	OBCAM	0.0134	66.893 (*df* 12) *				
Southern Lower	CRFA	0.0054	39.528 (*df* 15) *	0.0061	44.579 (*df* 14) *		
	IGF1	-0.0008	0 (*df* 10)	0.0067	39.945 (*df* 11) *		
	OBCAM	0.0210	130.032 (*df* 12) *	0.0140	90.720 (*df* 12) *		
Upper Peninsula	CRFA	-0.0012	0 (*df* 16)	0.0164	97.613 (*df* 16) *	0.0052	35.100 (*df* 15) *
	IGF1	0.0240	91.872 (*df* 11) *	0.0067	30.070 (*df* 12) *	0.0257	144.331 (*df* 12) *
	OBCAM	0.0096	36.538 (*df* 11) *	0.0249	101.343 (*df* 11) *	0.0176	99.264 (*df* 12) *

*** Significant at P = 0.0028 after Bonferroni correction.**

DISCUSSION

DNA quality and quantity is of major concern when utilized in molecular biology procedures, especially for forensic scientists handling minute amounts of degraded DNA samples. PCR does not require the highest quality or quantity DNA, however, amplification of the markers in this study was affected by these two variables. Samples exhibiting low quality and/or low quantity DNA often lead to unpredictable and unreliable results. Samples of this nature, containing either minimal amount of intact target DNA and flanking primer sites or very few copies of the target sequence, resulted in weak amplification products which were unable to be scored confidently. Problems were encountered in accurately scoring the alleles present for many samples with optical densities less than 1.5 and/or concentrations less than 15 ng/μl. Therefore, DNA was extracted from many deer samples in order to alleviate any scoring discrepancies. Freshly extracted DNA samples resulted in high molecular weight DNA facilitating unambiguous scoring of alleles.

The most successful method for extracting these tissue samples, resulting in the highest quality and quantity of DNA, was Proteinase K protein digestion using Qiagen lysis Buffer ATL followed by phenol/chloroform extraction and ethanol precipitation. Digestion with Proteinase K instead of Protease and the use of a commercial lysis buffer instead of a homemade lysis buffer yielded higher quality DNA. Also, the use of the Qiagen QIAamp tissue kit using spin columns resulted in low quality and quantity DNA, (data not shown). The options tried in this research project are not all inclusive of the

many different possible combinations of tissue extraction materials and methods that can be used. However, this combination clearly worked the best for the needs of this research project.

Along with the benefits of using PCR-based systems for individualizing forensic evidence, such as extreme sensitivity and specificity, simple and quick to perform, speed of amplification and the applicability to small quantities of degraded samples, comes the concern of PCR misincorporation rate, error and contamination. The issue of contamination was not observed because samples were handled under good laboratory practices and when they were genotyped, the results clearly comprised of no more than one or two alleles for each sample. The appearance of shadow bands, two nucleotides shorter than the correct allele, is a resultant error on the part of the polymerase due to slippage during chain elongation. These bands were present for all the markers used in this study and their appearance did not affect the scoring of the correct alleles. To the contrary, having these shadow bands made the identification of alleles easier whenever the allele intensity was weak. The other source of error by the polymerase is the addition of a nontemplated nucleotide, usually an A, to the 3′ end of the PCR product. This situation is marker specific with some markers being affected and some not at all, as was the case in this study. All the markers were not affected by this polymerase error, except the marker JP23, which resulted in both products, the true allele and the true allele + A. Correction for this error was to facilitate the nontemplated nucleotide addition by adding an additional five minute final extension step to the end of the thermocycler program in order to favor the production of the allele + A product. The misincorporation rate of *Taq* polymerase used in this study is approximately 10^{-4} nucleotides/cycle (Saiki *et al.*, 1988).

This error rate is minimal and would only be detectable by electrophoretically visible bands if the misincorporation was during the beginning cycles of PCR, with very limited amount of template and affecting many copies (Saiki *et al.*, 1988). Misincorporated nucleotides will not be detected because the visible product on the gel will be the majority of the amplified product. The incorporation of an incorrect nucleotide will not change the length of the fragment.

Eight microsatellite markers were initially used to assess genetic diversity among twenty random deer DNA samples resulting in four highly polymorphic loci. CRFA, OBCAM, IGF1 and IRBP2 all exhibit high heterozygosity and high polymorphic information content (PIC) and therefore show genetic variability. The calculated PIC value, according to the formulation of Botstein *et al.* (1980), is an indicator of a genetic markers' ability to distinguish between individuals and was used in selecting markers for database acquisition. A marker with a high PIC value is indicative of numerous allelic variants in the sampled population, and this makes it less likely that two random individuals will have the same alleles. The presence of several different alleles occurring in heterozygous individuals allows for the greatest power of discrimination, and the larger the excluded population, the greater the weight of the evidence. While CSN3, FCB304 and JP15 do not show a high enough degree of variability to have high discriminatory potential, JP23 has a high PIC value, but very low heterozygosity and thus could be an informative marker if combined with other polymorphic markers. The low heterozygosity with a high PIC value could be due to the presence of nonamplifying or 'null' alleles, which would result in mistyping a heterozygous individual as a homozygous individual. This was probably due to the loss of complementarity in

nucleotide sequence or mutation at the primer binding sites (Callen *et al.*, 1993; Pemberton *et al.*, 1995), or too few intact copies of target DNA at that locus resulting in weak amplification of alleles, unable to be scored accurately. This marker, JP23, was not selected for database acquisition because many samples had amplification problems resulting in no products. Amplification problems often occur when using primers derived from species other than the one which it was cloned.

Marker selection criteria for database acquisition focused on the development of a practical, economical, time-efficient DNA typing system for individualizing forensic evidence. This entails being able to multiplex several markers in one PCR and run them simultaneously in one electrophoretic lane of an ABI 377 DNA Sequencer in order to save on both money and time. Three of the four highly polymorphic markers were selected for database acquisition based on these criteria: OBCAM, CRFA and IGF1. The marker IRBP2 was eliminated because the annealing temperature was five degrees lower than the other three markers and therefore, could not be multiplexed together.

The three selected markers have high heterozygosities, high PIC values and a high number of allelic variants making them the obvious choices for the development of a marker panel with a potentially high power of discrimination. These markers also satisfy the quest for one multiplex PCR to be analyzed simultaneously in a single lane of a polyacrylamide gel. Optimal PCR conditions for all three markers were performed using the same amounts of each PCR component and the same annealing temperature of the thermocycler running conditions. The allele size ranges of these markers do not overlap, thus allowing them to be run simultaneously in one lane of a polyacrylamide gel. Each

marker is also separated by a different fluorescent dye, which eliminates any miscalling of alleles between each marker.

The use of heterologous primer sequences for these markers derived from bovine and ovine sources are homologous in cervines. Many studies have shown that within the order Artiodactyla, primer pairs from other species can amplify homologous products in related species of deer (DeWoody *et al.*, 1995; Engel *et al.*, 1996; Roed, 1998; Slate *et al.*, 1998; Kuhn *et al.*, 1996; Pepin *et al.*, 1995; O'Connell and Denome, 1999; Talbot *et al.*, 1996). This conservation of heterologous PCR primer pairs between related species has facilitated the development of this microsatellite marker system for individual identification by saving time and money that would have been required for cervid primer development.

Primer pairs for the loci OBCAM and IGF1 were derived from bovine DNA and CRFA was derived from ovine DNA. These microsatellite loci must follow Mendel's laws of inheritance in order to be considered genetic markers. Moore *et al.* in 1992 proved Mendel's principle of segregation for both OBCAM and CRFA by showing that in one bovine family, one allele from each parent was present in the five offspring sampled. Adams & Maddox in 1994 proved Mendelian inheritance of the IGF1 microsatellite in two two-generation ovine families. This study analyzed nine two-generation white-tailed deer families of a captive deer population from Savannah River, Georgia, and supplied by Dr. Jerry L. Ruth with the Fish and Wildlife Service of the United States Department of the Interior. The results from these families proved Mendelian inheritance of all three microsatellite markers used in this study. Mendel's principle of independent assortment states that alleles for one marker segregate

independently of alleles for the other loci being analyzed simultaneously. The location of IGF1 is on bovine chromosome 5, OBCAM is on 29 and CRFA is on ovine chromosome 9. As all three markers are on separate chromosomes for the related species bovine and ovine, which verifies independent assortment of alleles, and with the proven homology between species of Artiodactyls, the same assumption was made for cervids. There is no evidence of association of alleles among the three loci selected for database acquisition. The alleles for each marker show no signs of being linked to each other. These loci are inherited independently, which allows for the product rule to be applied when calculating multiple loci profiles from the observed allele frequencies under the assumption of Hardy-Weinberg equilibrium.

To address the issue of deviations from random mating in the sampled population of Michigan white-tailed deer, tests for Hardy-Weinberg equilibrium (HWE) were conducted. HWE was statistically tested for the sampled population in this study by using the Fisher Exact test, which tests for exactness to HWE by comparing the observed genotypic frequencies with those expected under HWE. That is, Hardy-Weinberg law states that genotypic frequencies will remain constant from generation to generation as long as the following assumptions remain true: the absence of selection, migration and mutation along with the continuation of random mating (King and Stansfield, 1990). The Fisher Exact test was selected because of the possibility of rare alleles occurring in very low frequency in the sampled population, and these small numbers will have very little effect on other tests performed at the 5% level of significance even in large sampled populations. The Fisher Exact test can identify significant differences between low frequency alleles and genotypes.

The P-values for the three markers for the sampled population along with the four separate regions of the population all show no significant deviation from HWE at the Bonferroni corrected significance level of P = 0.004 (Table 9). This probability or significance level measures the probability of a type I error, that is, the probability of falsely rejecting a true hypothesis. However, the marker CRFA appears to approach significance for the overall population and especially for the North East Lower Michigan region, which indicates the possibility of population substructure or inbreeding in that region. This could be particularly true for this region because of the feeding arrangement set up by hunters for deer in the area where tuberculosis was recently detected at higher levels. The microsatellite marker CRFA is associated with the gene corticotropin-releasing factor, which is involved in the body's response to stress and has an effect on the immune system. The incidence of tuberculosis in this region may therefore have resulted in selection pressure at that locus. This makes sense because the deer in the North East Lower Michigan were exposed to tuberculosis around the time the tissue samples were collected in 1994 and CRFA was the only marker of the three approaching significant deviation from HWE. The deer contained to this area, because of the excellent feeding arrangement, would thus constantly inbreed amongst each other and result in homozygote excess at the CRFA locus. This excess of homozygotes or deficiency of heterozygotes can readily be seen in Table 8 for CRFA, 81% observed heterozygosity compared to 87% expected heterozygosity. Table 11 shows that much of this excess of homozygotes is attributed to the Northern part of the lower peninsula with Table 13 showing that the North East region is causing much of the deviation from the expected heterozygosity. There are obvious deviations in observed heterozygosity compared to

HWE expectations, but only for CRFA. Another possible reason for the deviations could be the migration of deer northward because of fires in the southern part of the state years ago, causing the deer to be in close proximity to each other and to breed among siblings. The scope of this study cannot fully address the reasons for the homozygote excess for the marker CRFA in the population without additional testing however, this study did address the question of whether the population is consistent with HWE. The results (Table 9) were that HWE holds true at each locus in each region of Michigan and the state as a population. This means that the sampled population exists in equilibrium to the expectations of the Hardy-Weinberg law and genotype frequencies can be reliably estimated from allele frequency data. Further continuation of microsatellite marker testing on Michigan white-tailed deer can be based on random mating and statistical analyses assuming HWE can be performed. If not in HWE, then future testing would have to use different statistical analyses to account for disequilibrium.

An overall chi-square (X^2) was also calculated along with the Fisher Exact test by the GENEPOP V3.1d software program resulting in the same insignificant results when comparing all three markers for each region of Michigan for deviations from HWE at the same P = 0.004 (Table 9). These differences have a high probability of occurring by chance and are statistically insignificant, thus no deviation from Hardy-Weinberg equilibrium. A goodness of fit chi-square test means that the observed results do not differ significantly from what would be expected. This test aims to test the hypothesis that there is close agreement between what was observed in the sampled population and the expectations of HWE. These two statistical tests for Hardy-Weinberg equilibrium are not good indicators of population substructure therefore, the results cannot imply the

absence of substructure. To detect substructure, the population of Michigan was subdivided into selected regions and statistical analyses were performed on them. However, the best and most accurate way to detect substructuring is to sample the individual subgroups and observe the genotype frequencies between them.

To address this issue of a possible substructured population of Michigan white-tailed deer sampled in this study, tests for inter-population variance in allele frequency were conducted. That is, if there are differences in allele frequencies between the four regions in Michigan, then population substructuring exists. Even though tests for random mating under HWE were insignificant for each marker within each region of Michigan, testing for differences between these regions using the F-statistic (F_{st}) to estimate spatial heterogeneity in allele frequencies was conducted. Table 15 lists the allele frequencies for each marker in each region along with the calculated F_{st} value for each marker analyzed across the four regions. The calculated chi-square value measures the degree of significance of the F_{st} value, taking into account the large sample size of approximately 450 samples. The F_{st} values for the three markers across the four regions were low but statistically significant based on the chi-square values. This means that there is not a lot of variance between regions for allele frequency, but considering the large amount of samples, the F_{st} values for each marker become statistically significant after the chi-square test. Chi-square values greater than the level of significance (P = 0.004) after Bonferroni correction rejects the hypothesis of population homogeneity ($F_{st} = 0$) in allele frequencies and thus, indicates spatial heterogeneity within the sampled population of Michigan white-tailed deer.

Table 16 shows pair-wise comparisons between all regions for allelic frequency differences. The chi-square values are statistically significant at P = 0.0028 after Bonferroni correction. All combinations are significant except when comparing the three regions to the North East Lower region. The marker IGF1 for North West Lower and Southern Lower and CRFA for the Upper Peninsula are insignificant when compared to the allele frequencies in the North East Lower. Overall, allele frequencies differ between the four regions in Michigan, also known as the Wahlund effect because the population consists of a number of subpopulations having different allelic frequencies. Accurate calculations of probability estimates for linking forensic evidence to a suspect should apply the appropriate allele frequencies for the region of Michigan where the crime was committed. In the event of a match, the calculated probability using allele frequencies at the population level compared to using subpopulation frequencies, will not have a significant effect on the value of the match; the match will still be a match whether the probability is one in a million or one in ten million. A match between evidence at the crime scene and evidence on the suspect or in the suspect's possession clearly places the suspect at the crime scene.

The purpose of this research project was to develop a practical, economical, time-efficient DNA typing system for wildlife forensic scientists to utilize for individualizing forensic evidence. This DNA typing system has the ability to match two physically separated deer samples, thus placing the suspect at the scene of the crime and providing a powerful law enforcement tool for wildlife officials. This research project has developed a microsatellite marker panel which can be multiplexed in one PCR and run simultaneously in one electrophoretic lane of a DNA sequencer. This panel of markers

along with the application of molecular biology techniques will save forensic wildlife scientists time and money by assisting them in achieving their goal of enforcing wildlife laws. This DNA typing system has a power of discrimination of 0.9998 or 99.98 %, and in the event of matching two physically separated samples with forty-six possible allelic variants, accuracy in calculating probability estimates can be achieved, even when the population is substructured. Forensic wildlife scientists can also use this system for parentage determinations, which was validated by the testing of the nine pedigreed families in this study, and distinguishing native white-tailed deer from either imported white-tailed deer or deer from outside the Michigan area.

FUTURE DIRECTIONS

Recommendations for future improvements to this study for continued research in this area of forensic identification of Michigan white-tailed deer will be discussed. The group of samples used in the present study was limited to five samples per county, representing a random sampling of all Michigan deer, but not a representative sampling from each county. Some counties were underrepresented in the total sampled population of Michigan deer. Therefore, the frequency of alleles will not be an accurate representation of the whole Michigan deer population at large and when calculating the probability of a match, the estimate will not be truly accurate. If more samples were genotyped from each county, the allele frequencies would be exceedingly more accurate and new or rare alleles would arise, thus resulting in higher genetic variability. The goal of any identification system is to obtain all possible genetic variants in the population prior to analyzing evidentiary material.

This study resulted in four highly polymorphic markers of which three were ultimately used for database acquisition. The fourth marker, IRBP2, would be an excellent marker to add to the panel for many reasons. The percent heterozygosity, polymorphic information content and number of different alleles is relatively high. The allele size range would allow the marker to be analyzed simultaneously with the other three markers in one lane of a DNA sequencer. The fluorescent tag is HEX and with the neighboring markers in the panel being labeled with FAM and TET, IRBP2 could be used without the risk of any interfering problems. The PCR conditions are compatible with the other three markers, however, the annealing temperature used in the thermocycler

temperature programming is 50 degrees instead of 55 degrees as for the developed panel. This means that in order to add this marker to the panel, the length of the oligonucleotide primers would need to be extended an additional few nucleotide bases so that the T_m for each primer would be five degrees higher. The addition of an AC to the 3′ end of the primer 5′-GTATGATCACCTTCTATGCTTCC-3′ would raise the T_m by approximately six degrees because for every A or T added the T_m raises approximately two degrees and for every C or G added the T_m raises approximately four degrees. The addition of a CA to the 5′ end of the primer 5′-CCCTAAATACTACCATCTAGAAG-3′ would also raise the T_m by approximately six degrees. In order to have optimal PCR conditions, both primers need to have similar T_m's and a G or C at the 3′ end to anchor the primer so efficient elongation by the polymerase can occur. Other considerations when adding these two bases are sequence complementarity between primers and the formation of secondary structures, so checking these new sequences through an oligonucleotide primer design program prior to synthesizing the primers is highly recommended. These additional bases were selected from adjacent sequence information for the primers for this marker, located in GenBank (Accession # M20748). This sequence information was unavailable during this research project, therefore, not pursued for this microsatellite panel. The addition of this marker along with other polymorphic markers to the developed panel will increase the power of discrimination of the system, thus making it an even more powerful tool for individualizing forensic evidence.

A more accurate evaluation of the data presented here for testing the substructure of the population of white-tailed deer in Michigan would take into account other factors

when subdividing the deer population in Michigan. For this study, the assessment of the structure of the population was based on dividing the state into four regions based on county lines, which has no relevance biologically. A future study could accurately assess the division of Michigan based on actual geographical barriers, which would be more relevant to answering questions concerning the structuring of the Michigan white-tailed deer population.

Wildlife forensic scientists will benefit from the results of this study in several ways. They will have an easy, low cost, reliable scientific test to perform, with results in less than one day, on evidence linking a suspect to a crime when a suspect has been identified. That is, to determine whether the field evidence matches the evidence found in the possession of the suspect with a high degree of probability. This DNA typing system will assist law enforcement officials in convicting poachers by providing critical evidence linking evidentiary samples and parentage testing for illegal importation cases. The mere fact of having the ability to individualize a deer for forensic purposes will hopefully deter the criminal from committing the crime. As for many crimes committed by humans, members in the criminal justice system use deterrence, usually harsh sentencing and high penalties, to decrease the number of committed crimes. However, the effectiveness of this deterrence remains to be seen, but for wildlife law enforcement officers it would be a step in the right direction to minimizing poaching crimes. A future study could evaluate the deterrent effect and the number of convictions based on having this DNA typing system available to wildlife forensic scientists.

The data presented in this study has laid the groundwork for several future applications, other than individual identification, of Michigan white-tailed deer.

Database acquisition is a necessary requirement for calculating probability estimates for individual identification, and exclusion/inclusion of parentage as well as the determination of the extent of population substructure, genetic diversity, population size and the detrimental effects of inbreeding. Evolutionary and comparative mapping studies between related species can also be accomplished utilizing the accumulated data from this study. The assessment of captive breeding programs for the loss of genetic diversity is another possibility. Identification of relatedness among pedigreed and captive populations, so that the selection of matings will preserve genetic variability in order to maintain the species, can furthermore be accomplished.

APPENDIX A

APPENDIX A

AGAROSE GEL PREPARATION PROCEDURE

1. Determine the gel percentage and the number of wells needed.

2. A small gel (15cm x 15cm) requires at least 125 ml of agarose solution.

3. Weigh the correct amount of agarose (SeaKem LE Agarose, Cat # 50007 from FMC) based on the gel percentage and place in a 250 ml Erlenmeyer flask.

4. A 0.8% gel (for genomic DNA) would need 1 g agarose and a 2% gel (for PCR products) would need 2.5 g agarose in 125 ml of 1X TBE (Tris/boric acid/EDTA).

5. Prepare a 10X TBE solution with 108 g Tris, 55 g boric acid, 8.3 g EDTA and quantity sufficient to 1 L with deionized water. Dilute to 1X TBE with deionized water for use in gel preparation and for the electrophoresis buffer.

6. Mix the agarose and 1X TBE in the flask by either heating on top of a heated stir plate or on low in a microwave until the agarose has completely dissolved. Watch the solution in the microwave to prevent the solution from boiling over and becoming a different concentration. Once dissolved, add deionized water to correct for any loss due to evaporation during heating.

7. Let the solution cool until the flask is able to be handled by the hands; too cool will show signs of hardening in the flask and too warm will warp the plastic tray.

8. Ethidium bromide (Cat # E-8751 from Sigma) is not added to the solution for possible interference with fragment migration. The gel after electrophoresis is submerged in an ethidium bromide solution in order to detect the fragments.

9. Place appropriate comb(s) into the level gel tray and carefully pour the agarose solution into the tray so as to not introduce bubbles.

10. Allow the gel to stand and harden for at least 30 minutes at room temperature. Quicker hardening may be achieved by placing the tray in the refrigerator.

11. Remove the comb(s) with the gel submerged in the electrophoresis buffer (1X TBE) to prevent tearing the wells.

12. Add enough buffer to the electrophoresis chamber to cover the gel and load the samples into the wells when ready.

APPENDIX B

APPENDIX B

TISSUE EXTRACTION PROCEDURE

A. TISSUE DIGESTION

1. Cut 50 mg of frozen tissue into very tiny pieces with a sterile scalpel and place in a 1.5 ml polypropylene tube.

2. Add 250 µl of Qiagen Buffer ATL (Cat # 19076) to each tube.

3. Add 25 µl of 20 mg/ml Proteinase K (Cat # 25530-015 from Gibco BRL) to each tube.

4. Vortex each tube vigorously for 30 seconds until a homogeneous mixture is obtained.

5. Incubate all tubes in a 55° waterbath overnight or until the tissues are lysed completely.

B. PHENOL/CHLOROFORM EXTRACTION

1. Add 300 µl of phenol:chloroform:isoamyl alcohol, 25:24:1, saturated with 10mM Tris, pH 8.0, 1mM EDTA (Cat # 0883-100mL from Amresco) to each tube.

2. Invert the tubes several times to thoroughly mix the contents.

3. Spin the tubes for 5 minutes at full speed (12,000 rpm) in a microcentrifuge.

4. Remove the top aqueous layer (approx. 300 µl) and add to a fresh tube.

5. Repeat steps 1 through 4.

6. Add 300 µl of chloroform (Cat # 9180-03 from J.T. Baker) to each tube.

7. Invert the tubes several times to thoroughly mix the contents.

8. Spin the tubes for 5 minutes at full speed (12,000 rpm) in a microcentrifuge.

9. Remove the top aqueous layer (approx. 300 µl) and add to a fresh tube.

10. Perform steps 1-9 in a vented hood due to the carcinogenicity of the chemicals.

C. ETHANOL PRECIPITATION

1. Add 600 µl of 100% ethanol and 30 µl of 3M sodium acetate to each tube.

2. Invert the tubes several times to fully precipitate the DNA.

3. Spin the tubes for 5 minutes at full speed (12,000 rpm) in a microcentrifuge.

4. Carefully decant the contents of each tube without losing the pelleted DNA at the bottom.

5. Add 1 ml of 70% ethanol to each tube.

6. Invert the tubes several times to thoroughly wash the DNA.

7. Spin the tubes for 5 minutes at full speed (12,000 rpm) in a microcentrifuge.

8. Carefully decant the contents of each tube without losing the pelleted DNA at the bottom.

9. Resuspend the DNA in the tubes, after the DNA has air dried, in 200 µl Tris/EDTA (TE), pH 8.0. This amount should give a yield of approximately 200 ng/µl of high quality DNA with an absorbance ratio (260nm/280nm) in the range of 1.7-1.9.

10. Measure the quantity of DNA in the samples using a fluorometer which yields a measurement of only DNA and not RNA. Measure the quality of DNA in the samples using a spectrophotometer which yields a ratio of both DNA and RNA to protein in the samples.

11. Dilute the DNA samples to approximately 30 ng/µl with deionized water. These samples are now ready for a polymerase chain reaction (PCR).

APPENDIX C

APPENDIX C

ACRYLAMIDE GEL PREPARATION PROCEDURE

1. Prepare a 10X Tris/boric acid/EDTA solution (TBE) with 108 g Tris, 55 g boric acid, 8.3 g EDTA and quantity sufficient to 1 L with deionized water.

2. Add 18 g urea (Cat # IB72064 Molecular Biology Grade from Kodak), 5.3 ml 40% Acrylamide/Bis 19:1 (5% C) (Cat # 161-0120 from Biorad), 25 ml deionized water, 5.5 ml 10X TBE and approximately 0.5 g of Amberlite MB-150 (Cat # A-5710 from Sigma) to an Erlenmeyer flask.

3. Mix the contents using a magnetic stir bar on a stir plate until dissolved completely.

4. Filter and degas the solution using a 0.45 micron bottle top filter (Cat # 290-3345 from Nalgene).

5. Prepare a 10% ammonium persulfate (Cat # 0486-100G-APP from Amresco) solution (APS) with deionized water.

6. Add 250 µl of the 10% APS to the acrylamide solution.

7. Gently swirl the bottle to mix the contents without introducing bubbles.

8. Add 35 µl TEMED (Cat # 161-0800 from Biorad) to the acrylamide solution.

9. Gently swirl the bottle to mix the contents without introducing bubbles.

10. Pour the solution quickly into the clamped ABI 377 DNA sequencing plates without introducing bubbles and before the solution polymerizes.

11. Place the appropriate sized comb into the plate opening and clamp together.

12. Let the gel stand for at least 1 ½ hours before using to assure complete polymerization.

13. Acrylamide is a neurotoxin so extreme care should be taken and appropriate safety equipment should be worn during the above steps.

APPENDIX D

APPENDIX D

MULTIPLEX PCR PROTOCOL

A. MULTIPLEX PCR COMPONENTS FOR A 25 µl REACTION

1. Pipet 2 µl of each 30 ng/µl DNA sample into separate microcentrifuge tubes.
2. Pipet 13.75 µl deionized water to each tube.
3. Pipet 3.0 µl $MgCl_2$ (25mM) to each tube.
4. Pipet 2.5 µl 10X buffer (100 mM KCl, 20 mM Tris-HCl, pH 8.0, 0.1 mM EDTA) to each tube.
5. Pipet 2.0 µl dNTP's (2.5 mM) to each tube.
6. Pipet 0.25 µl each primer (20 µM) to each tube.
7. Pipet 0.25 µl AmpliTaq DNA polymerase (5 Units/µl) to each tube.
8. Overlay each sample mixture with 1 drop of mineral oil.

B. MULTIPLEX PCR THERMOCYCLER TEMPERATURE PROGRAMMING

1. Initial denaturation – 95°C for 3 minutes.
2. Denaturation – 95°C for 45 seconds.
3. Annealing – 55°C for 45 seconds.
4. Extension – 72°C for 45 seconds.
5. Autoextend – 1 second per cycle.
6. Repeat steps 1 through 5 for 29 cycles.
7. End with a 4°C hold.

REFERENCES

REFERENCES

Adams, L. J. & Maddox, J. F. (1994). A dinucleotide repeat polymorphism in the ovine insulin-like growth factor-I gene 5′ flanking region. Animal Genetics, 25(1), 61.

Andre, P., Kim, A., Khrapko, K. & Thilly, W. G. (1997). Fidelity and mutational spectrum of *Pfu* DNA polymerase on a human mitochondrial DNA sequence. Genome Research, 7, 843-852.

Baird, J. (1992). Forensic DNA in the Trial Court 1990-1992: A Brief History. In P. R. Billings (Ed.), DNA on Trial: Genetic Identification and Criminal Justice. (pp.61-77) New York: Cold Spring Harbor Laboratory Press.

Beckmann, J. S. & Weber, J. L. (1992). Survey of human and rat microsatellites. Genomics, 12, 627-631.

Blake, E., Mihalovich, J., Higuchi, R., Walsh, P. S. & Erlich, H. (1992). Polymerase chain reaction (PCR) amplification and human leukocyte antigen (HLA)-DQa oligonucleotide typing on biological evidence samples: Casework experience. Journal of Forensic Sciences, 37(3), 700-726.

Botstein, D., White, R. L., Skolnick, M. & Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American Journal of Human Genetics, 32, 314-331.

Brownstein, M. J., Carpten, J. D. & Smith, J. R. (1996). Modulation of non-templated nucleotide addition by *Taq* DNA polymerase: primer modifications that facilitate genotyping. BioTechniques, 20, 1004-1010.

Budowle, B., Moretti, T. R., Keys, K. M., Koons, B. W. & Smerick, J. B. (1997). Validation studies of the CTT STR multiplex system. Journal of Forensic Sciences, 42(2), 701-707.

Callen, D. F., Thompson, A. D., Shen, Y., Phillips, H. A., Richards, R. I., Mulley, J. C. & Sutherland, G. R. (1993). Incidence and origin of "null" alleles in the $(AC)_n$ microsatellite markers. American Journal of Human Genetics, 52, 922-927.

Cha, R. S. & Thilly, W. G. (1993). Specificity, efficiency and fidelity of PCR. PCR Methods and Applications, 3, S18-S29.

Chamberlain, J. S., Gibbs, R. A., Ranier, J. E., Nguyen, P. N. & Thomas, C. (1988). Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. Nucleic Acids Research, 16(23), 11141-11156.

Chang, J. C. & Kan, Y. W. (1982). A sensitive new prenatal test for sickle-cell anemia. New England Journal of Medicine, 307(1), 30-32.

Clark, J. M. (1988). Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA. Nucleic Acids Research, 16(20), 9677-9686.

DeWoody, J. A., Honeycutt, R. L. & Skow, L. C. (1995). Microsatellite markers in white-tailed deer. The Journal of Heredity, 86(4), 317-319.

Engel, S. R., Linn, R. A., Taylor, J. F. & Davis, S. K. (1996). Conservation of microsatellite loci across species of artiodactyls: Implications for population studies. Journal of Mammalogy, 77(2), 504-518.

Erlich, H. A., Sheldon, E. L. & Horn, G. (1986). HLA typing using DNA probes. BioTechnology, 4, 975-981.

Frank, T. S. Topics in molecular oncology: Microsatellite instability. [Online] http://zapruder.pds.med.umich.edu/users/Frank/MIN.html, September 10, 1997.

Fregeau, C. J., Bowen, K. L. & Fourney, R. M. (1999). Validation of highly polymorphic fluorescent multiplex short tandem repeat systems using two generations of DNA sequencers. Journal of Forensic Sciences, 44(1), 133-166.

Gelfand, D. H. (1989). *Thermus aquaticus* DNA polymerase. In H. A. Erlich, R. Gibbs & H. H. Kazazian, Jr. (Eds.), Current Communications in Molecular Biology: Polymerase Chain Reaction. (pp. 11-17) New York: Cold Spring Harbor Laboratory Press.

Gelfand, D. H. & White, T. J. (1990). Thermostable DNA polymerases. In M. A. Innis, D. H. Gelfand, J. J. Sninsky & T. J. White (Eds.), PCR Protocols: A Guide to Methods and Applications. (pp. 129-141) New York: Academic Press.

Ghosh, S., Karanjawala, Z. E., Hauser, E. R., Ally, D., Knapp, J. I., Rayman, J. B., Musick, A., Tannenbaum, J., Te, C., Shapiro, S., Eldridge, W., Musick, T., Martin, C., Smith, J. R., Carpten, J. D., Brownstein, M. J., Powell, J. I., Whiten, R., Chines, P., Nylund, S. J., Magnuson, V. L., Boehnke, M., Collins, F. S. & the FUSION (Finland-U.S. Investigation of NIDDM Genetics) Study Group (1997). Methods for the precise sizing, automated binning of alleles, and reduction of error rates in large-scale genotyping using fluorescently labeled dinucleotide markers. Genome Research, 7, 165-178.

Hamada, H., Petrino, M. G. & Kakunaga, T. (1982). A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. Proceedings of the National Academy of Sciences, USA, 79, 6465-6469.

Hamada, H., Seidman, M., Howard, B. H. & Gorman, C. M. (1984). Enhanced gene expression by the poly(dT-dG)·poly(dC-dA) sequence. Molecular and Cellular Biology, 4(12), 2622-2630.

Hauge, X. Y. & Litt, M. (1993). A study of the origin of 'shadow bands' seen when typing dinucleotide repeat polymorphisms by the PCR. Human Molecular Genetics, 2(4), 411-415.

Helmuth, R., Fildes, N., Blake, E., Luce, C., Chimera, J., Madej, R., Gorodezky, C., Stoneking, M., Schmill, N., Klitz, W., Higuchi, R. & Erlich, H. A. (1990). HLA-DQa allele and genotype frequencies in various human populations, determined by using enzymatic amplification and oligonucleotide probes. American Journal of Human Genetics, 47, 515-523.

Henegariu, O., Heerema, N. A., Dlouhy, S. R., Vance, G. H. & Vogt, P. H. (1997). Multiplex PCR: Critical parameters and step-by-step protocol. BioTechniques, 23, 504-511.

Jeffreys, A. J., Wilson, V. & Thein, S. L. (1985a). Hypervariable 'minisatellite' regions in human DNA. Nature, 314, 67-73.

Jeffreys, A. J., Wilson, V. & Thein, S. L. (1985b). Individual-specific 'fingerprints' of human DNA. Nature, 316, 76-79.

Jeffreys, A. J., Brookfield, J. F. Y. & Semeonoff, R. (1985c). Positive identification of an immigration test-case using human DNA fingerprints. Nature, 317, 818-819.

Jeffreys, A. J., Wilson, V., Neumann, R. & Keyte, J. (1988). Amplification of human minisatellites by the polymerase chain reaction: Towards DNA fingerprinting of single cells. Nucleic Acids Research, 16(23), 10953-10971.

Kan, Y. W. & Dozy, A. M. (1978). Polymorphism of DNA sequence adjacent to human B-globin structural gene: Relationship to sickle mutation. Proceedings of the National Academy of Sciences, USA, 75(11), 5631-5635.

Kantanen, J., Vilkki, J., Elo, K. & Tanila, A. M. (1995). Random amplified polymorphic DNA in cattle and sheep: Application for detecting genetic variation. Animal Genetics, 26, 315-320.

Kemp, S. J., Hishida, O., Wambugu, J., Rink, A., Longeri, M. L., Ma, R. Z., Da, Y., Lewin, H. A., Barendse, W. & Teale, A. J. (1995). A panel of polymorphic bovine, ovine and caprine microsatellite markers. Animal Genetics, 26, 299-306.

Kimpton, C. P., Gill, P., Walton, A., Urquhart, A., Millican, E. S. & Adams, M. (1993). Automated DNA profiling employing multiplex amplification of short tandem repeat loci. PCR Methods and Applications, 3, 13-22.

King, R. C. and Stansfield, W. D. (1990). A Dictionary of Genetics (fourth edition). New York: Oxford University Press.

Kobilinsky, L. (1993). Deoxyribonucleic acid structure and function – A review. In R. Saferstein (Ed.), Forensic Science Handbook Vol. III. (pp. 287-357) New Jersey: Regents/Prentice-Hall.

Kostia, S., Palo, J. & Varvio, S-L. (1996). DNA sequences of RAPD fragments in artiodactyls. Genome, 39, 456-458.

Krawczak, M. & Schmidtke, J. (1994). DNA Fingerprinting. Oxford, UK: Bios Scientific Publishers Limited.

Kuhn, R., Anastassiadis, C. & Pirchner, F. (1996). Transfer of bovine microsatellites to the cervine (*Cervus elaphus*). Animal Genetics, 27, 199-201.

Lindquester, G. J. Eukaryotic DNA content. [Online] http://glindquester.biology.rhodes.edu/molbiol/eukcontent.html, 1997.

Lins, A. M., Sprecher, C. J., Puers, C. & Schumm, J. W. (1996). Multiplex sets for the amplification of polymorphic short tandem repeat loci - silver stain and fluorescence detection. BioTechniques, 20, 882-889.

Litt, M. & Luty, J. A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. American Journal of Human Genetics, 44, 397-401.

Lorente, M., Lorente, J. A., Alvarez, C., Budowle, B., Wilson, M. R. & Villanueva, E. (1997). Sequential multiplex amplification: Utility in forensic casework with minimal amounts of DNA and partially degraded samples. Journal of Forensic Sciences, 42(5), 923-925.

Melvin, Jr., J. R., Kateley, Jr., J. R., Oaks, M. K., Simson, Jr., L. R. & Maldonado, W. E. (1988). Paternity testing. In R. Saferstein (Ed.), Forensic Science Handbook Vol. II. (pp. 273-346) New Jersey: Prentice-Hall.

Moore, S. S., Sargeant, L. L., King, T. J., Mattick, J. S., Georges, M. & Hetzel, D. J. S. (1991). The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. Genomics, 10, 654-660.

Moore, S. S., Barendse, W., Berger, K. T., Armitage, S. M. & Hetzel, D. J. S. (1992). Bovine and ovine DNA microsatellites from the EMBL and GENBANK databases. Animal Genetics, 23, 463-467.

Moxon, E. R. & Wills, C. (1999). DNA microsatellites: Agents of evolution? Scientific American, January, 94-99.

Mullis, K. B. & Faloona, F. A. (1987). Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. Methods in Enzymology, 155, 335-350.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. & Erlich, H. (1986). Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. Cold Spring Harbor Symposium of Quantitative Biology, 51, 263-273.

Murray, V., Monchawin, C. & England, P. R. (1993). The determination of the sequences present in the shadow bands of a dinucleotide repeat PCR. Nucleic Acids Research, 21(10), 2395-2398.

Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E. & White, R. (1987). Variable number of tandem repeat (VNTR) markers for human gene mapping. Science, 235, 1616-1622.

National Research Council. (1992). DNA Technology in Forensic Science. Washington, D. C.: National Academy Press.

National Research Council. (1996). The Evaluation of Forensic DNA Evidence. Washington, D. C.: National Academy Press.

Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K. & Sekiya, T. (1989a). Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. Proceedings of the National Academy of Sciences, USA, 86, 2766-2770.

Orita, M., Suzuki, Y., Sekiya, T. & Hayashi, K. (1989b). Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. Genomics, 5, 874-879.

Pardue, M. L., Lowenhaupt, K., Rich, A. & Nordheim, A. (1987). $(dC\text{-}dA)_n \cdot (dG\text{-}dT)_n$ sequences have evolutionarily conserved chromosomal locations in *Drosophila* with implications for roles in chromosome structure and function. EMBO Journal, 6(6), 1781-1789.

Pemberton, J. M., Slate, J., Bancroft, D. R. & Barrett, J. A. (1995). Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. Molecular Ecology, 4, 249-252.

Pena, S. D. J. & Chakraborty, R. (1994). Paternity testing in the DNA era. Trends in Genetics, 10(6), 204-209.

Raymond, M., Rousset, F. & Slatkin, M. (1995). Genepop (V1.2): A population genetics software for exact tests and ecumenicism. Journal of Heredity, 86, 248-249.

Royle, N. J., Clarkson, R. E., Wong, Z. & Jeffreys, A. J. (1988). Clustering of hypervariable minisatellites in the proterminal regions of human autosomes. Genomics, 3, 352-360.

Saiki, R. K. (1989). Optimization of the polymerase chain reaction. In H. A. Erlich, R. Gibbs & H. H. Kazazian, Jr. (Eds.), Current Communications in Molecular Biology: Polymerase Chain Reaction. (pp. 25-30) New York: Cold Spring Harbor Laboratory Press.

Saiki, R. K., Bugawan, T. L., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1986). Analysis of enzymatically amplified B-globin and HLA-DQa DNA with allele-specific oligonucleotide probes. Nature, 324, 163-166.

Saiki, R. K., Walsh, P. S., Levenson, C. H. & Erlich, H. A. (1989). Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes. Proceedings of the National Academy of Sciences, USA, 86, 6230-6234.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science, 239, 487-491.

Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. & Arnheim, N. (1985). Enzymatic amplification of B-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science, 230, 1350-1354.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual. New York: Cold Spring Harbor Laboratory Press.

Schlotterer, C. & Tautz, D. (1992). Slippage synthesis of simple sequence DNA. Nucleic Acids Research, 20(2), 211-215.

Schwengel, D. A., Jedlicka, A. E., Nanthakumar, E. J., Weber, J. L. & Levitt, R. C. (1994). Comparison of fluorescence-based semi-automated genotyping of multiple microsatellite loci with autoradiographic techniques. Genomics, 22, 46-54.

Sensabaugh, G. F. & Blake, E. T. (1993). DNA analysis in biological evidence: Applications of the polymerase chain reaction. In R. Saferstein (Ed.), Forensic Science Handbook Vol. III. (pp. 416-452) New Jersey: Regents/Prentice-Hall.

Skolnick, M. H. & Wallace, R. B. (1988). Simultaneous analysis of multiple polymorphic loci using amplified sequence polymorphisms (ASPs). Genomics, 2, 273-279.

Slate, J., Coltman, D. W., Goodman, S. J., MacLean, I., Pemberton, J. M. & Williams, J. L. (1998). Bovine microsatellite loci are highly conserved in red deer (*Cervus elaphus*), sika deer (*Cervus nippon*) and Soay sheep (*Ovis aries*). Animal Genetics, 29, 307-315.

Smith, J. R., Carpten, J. D., Brownstein, M. J., Ghosh, S., Magnuson, V. L., Gilbert, D. A., Trent, J. M. & Collins, F. S. (1995). Approach to genotyping errors caused by nontemplated nucleotide addition by *Taq* DNA polymerase. Genome Research, 5, 312-317.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. Journal of Molecular Biology, 98, 503-527.

Stallings, R. L., Ford, A. F., Nelson, D., Torney, D. C., Hildebrand, C. E. & Moyzis, R. K. (1991). Evolution and distribution of $(GT)_n$ repetitive sequences in mammalian genomes. Genomics, 10, 807-815.

State v. Andrews, case no. 87-1565, Ninth Judicial Circuit Ct., Orange County, Fla., Division 15, Nov. 6, 1987.

Tautz, D. & Renz, M. (1984). Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Research, 12(10), 4127-4138.

Tautz, D. (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Research, 17(16), 6463-6471.

Technical Working Group on DNA Analysis Methods (TWGDAM) (1989). Guidelines for a quality assurance program for DNA restriction fragment length polymorphism analysis. Crime Laboratory Digest, 16, 40-59.

Technical Working Group on DNA Analysis Methods (TWGDAM) (1990). Guidelines for a quality assurance program for DNA restriction fragment length polymorphism analysis. Crime Laboratory Digest, 17, 50-60.

Tuggle, C. K. (1994). DNA Technology: Its use in animal breeding. Presented at a symposium on future genetics for the animal industry, St. Louis, MO, May 4.

Weber, J. L. & May, P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. American Journal of Human Genetics, 44, 388-396.

Weber, J. L. (1990). Informativeness of human $(dC{\cdot}dA)_n{\cdot}(dG{\cdot}dT)_n$ polymorphisms. Genomics, 7, 524-530.

Weedn, V. W. & Roby, R. K. (1993). Forensic DNA testing. Archives of Pathology and Laboratory Medicine, 117, 486-491.

Weir, B. S. (1990). Genetic Data Analysis. Sunderland, Massachusetts: Sinauer Associates, Inc. Publishers.

Weir, B. S. (1996). The second National Research Council report on forensic DNA evidence. American Journal of Human Genetics, 59, 497-500.

Wells, R. A., Green, P. & Reeders, S. T. (1989). Simultaneous genetic mapping of multiple human minisatellite sequences using DNA fingerprinting. Genomics, 5, 761-772.

Welsh, J. & McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Research, 18(24), 7213-7218.

Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A. & Tingey, S. V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research, 18(22), 6531-6535.

Wyman, A. R. & White, R. (1980). A highly polymorphic locus in human DNA. Proceedings of the National Academy of Sciences, USA, 77(11), 6754-6758.

Ziegle, J. S., Su, Y., Corcoran, K. P., Nie, L., Mayrand, E., Hoff, L. B., McBride, L. J., Kronick, M. N. and Diehl, S. R. (1992). Application of automated DNA sizing technology for genotyping microsatellite loci. Genomics, 14, 1026-1031.

Zumbo, J. (1999). Taking aim at poachers, part III. Outdoor Life, October, 204(3), 55-59, 102.

BIBLIOGRAPHY

BIBLIOGRAPHY

Balazs, I. (1990). Properties of hypervariable single locus polymorphisms and their application to identity testing. In J. Robertson, A. M. Ross & L. A. Burgoyne (Eds.), DNA in Forensic Science: Theory, Techniques and Applications. (pp. 112-124) New York: Ellis Horwood Limited.

Bruford, M. W. & Wayne, R. K. (1993). Microsatellites and their application to population genetic studies. Current Opinion in Genetics and Development, 3, 939-943.

Buchanan, F. C. & Crawford, A. M. (1993). Ovine microsatellites at the OarFCB11, OarFCB128, OarFCB193, OarFCB266 and OarFCB304 loci. Animal Genetics, 24, 145.

Budowle, B., Deadman, H., Murch, R. S. & Baechtel, F. S. (1988). An introduction to the methods of DNA analysis under investigation in the FBI laboratory. Crime Laboratory Digest, 15(1), 8-21.

Budowle, B., Sajantila, A., Hochmeister, M. N. & Comey, C. T. (1994). In K. B. Mullis, F. Ferre & R. A. Gibbs (Eds.), The Polymerase Chain Reaction. (pp. 244-256) Boston: Birkhauser.

Carrano, A. V., Lamerdin, J., Ashworth, L. K., Watkins, B., Branscomb, E., Slezak, T., Raff, M., de Jong, P. J., Keith, D., McBride, L., Meister, S. & Kronick, M. (1989). A high-resolution, fluorescence-based, semiautomated method for DNA fingerprinting. Genomics, 4, 129-136.

Chamberlain, J. S. & Chamberlain, J. R. (1994). Optimization of multiplex PCR's. In K. B. Mullis, F. Ferre & R. A. Gibbs (Eds.), The Polymerase Chain Reaction. (pp. 38-46) Boston: Birkhauser.

Cook, J. M. (1999). New Zealand's AgResearch maps 500 deer traits using rare hybrid breed. Agricultural Genomics, 2(13), 3-4.

Easteal, S., McLeod, N. & Reed, K. (Ed.) (1991). DNA Profiling: Principles, Pitfalls and Potential. Philadelphia: Harwood Academic Publishers.

Erlich, H. A., Gelfand, D. & Sninsky, J. J. (1991). Recent advances in the polymerase chain reaction. Science, 252, 1643-1651.

Erlich, H. A. (Ed.) (1989). PCR Technology: Principles and Applications for DNA Amplification. New York: Stockton Press.

Fowler, C. (1990). Polymorphism in the major human repetitive DNA sequences: Theory, practice, and comparison with other methods of measuring human genome variation. In J. Robertson, A. M. Ross & L. A. Burgoyne (Eds.), DNA in Forensic Science: Theory, Techniques and Applications. (pp. 87-111) New York: Ellis Horwood Limited.

Fowler, J. C. S., Burgoyne, L. A., Scott, A. C. & Harding, H. W. J. (1988). Repetitive deoxyribonucleic acid (DNA) and human genome variation-a concise review relevant to forensic biology. Journal of Forensic Sciences, 33(5), 1111-1126.

Garza, J. C., Slatkin, M. & Freimer, N. B. (1995). Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. Molecular Biology Evolution, 12(4), 594-603.

Gill, P., Jeffreys, A. J. & Werrett, D. J. (1985). Forensic application of DNA 'fingerprints'. Nature, 318, 577-579.

Haig, S. M. (1998). Molecular contributions to conservation. Ecology, 79(2), 413-425.

Herrera, R. J. & Tracey, Jr., M. L. (1992). DNA fingerprinting: Basic techniques, problems, and solutions. Journal of Criminal Justice, 20, 237-248.

Higuchi, R. & Blake, E. T. (1989). Applications of the polymerase chain reaction in forensic science. In J. Ballantyne, G. Sensabaugh & J. Witkowski (Eds.), DNA Technology and Forensic Science: Banbury Report 32. (pp. 265-281) New York: Cold Spring Harbor Laboratory Press.

Huang, P. L., Huang, P. L. & Lee-Huang, S. (1990). DNA polymorphism and forensic identification. In H. C. Lee & R. E. Gaensslen (Eds.), DNA and Other Polymorphisms in Forensic Science. (pp. 1-25) Chicago: Year Book Medical Publishers, Inc.

Hughes, A. E. (1993). Optimization of microsatellite analysis for genetic mapping. Genomics, 15, 433-434.

Inman, K. & Rudin, N. (1997). An Introduction to Forensic DNA Analysis. Florida: CRC Press, Inc.

Innis, M. A., Gelfand, D. H., Sninsky, J. J. & White, T. J. (1990). PCR Protocols: A Guide to Methods and Applications. New York: Academic Press.

Jarman, A. P., Nicholls, R. D., Weatherall, D. J., Clegg, J. & Higgs, D. R. (1986). Molecular characterization of a hypervariable region downstream of the human a-globin gene cluster. EMBO Journal, 5(8), 1857-1863.

Jeffreys, A. J., Wilson, V., Wong, Z., Royle, N., Patel, I., Kelly, R. & Clarkson, R. (1987). Highly variable minisatellites and DNA fingerprints. Biochemical Society Symposium, Great Britain, 53, 165-180.

Jeffreys, A. J. & Pena, S. D. J. (1993). Brief introduction to human DNA fingerprinting. In S. D. J. Pena, R. Chakraborty, J. T. Epplen & A. J. Jeffreys (Eds.), DNA Fingerprinting: State of the Science. (pp. 1-20) Boston: Birkauser Verlag.

Jeffreys, A. J. (1993). DNA typing: Approaches and applications. Journal of the Forensic Science Society, 33(4), 204-211.

Jeffreys, A. J., Wilson, V., Thein, S. L., Weatherall, D. J. & Ponder, B. A. J. (1986). DNA "fingerprints" and segregation analysis of multiple markers in human pedigrees. American Journal of Human Genetics, 39, 11-24.

Jin, Li, Underhill, P. A., Buoncristiani, M. & Robertson, J. M. (1997). Defining microsatellite alleles by genotyping global indigenous human populations and non-human primates. Journal of Forensic Sciences, 42(3), 496-499.

Jurka, J. & Pethiyagoda, C. (1995). Simple repetitive sequences from primates: Compilation and analysis. Journal of Molecular Evolution, 40, 120-126.

Kasai, K., Nakamura, Y. & White, R. (1990). Amplification of a variable number of tandem repeats (VNTR) locus (pMCT118) by the polymerase chain reaction (PCR) and its'application to forensic science. Journal of Forensic Sciences, 35(5), 1196-1200.

Kemp, S. J., Brezinsky, L. & Teale, A. J. (1993). A panel of bovine, ovine and caprine polymorphic microsatellites. Animal Genetics, 24, 363-365.

Kirkpatrick, B. W. (1992). Identification of a conserved microsatellite site in the porcine and bovine insulin-like growth factor-I gene 5′ flank. Animal Genetics, 23, 543-548.

Kobilinsky, L. (1993). Deoxyribonucleic acid structure and function - a review. In R. Saferstein (Ed.), Forensic Science Handbook Vol. III. (pp. 287-357) New Jersey: Regents/Prentice-Hall.

Lander, E. S. (1989). DNA fingerprinting on trial. Nature, 339, 501-505.

Lander, E. S. (1991). Invited editorial: Research on DNA typing catching up with courtroom application. American Journal of Human Genetics, 48, 819-823.

Lander, E. S. (1991). Lander Reply. American Journal of Human Genetics, 49, 899-903.

Lee, H. C. (1982). Identification and grouping of bloodstains. In R. Saferstein (Ed.), Forensic Science Handbook. (pp. 267-337) New Jersey: Prentice-Hall.

Lewontin, R. C. & Hartl, D. L. (1991). Population genetics in forensic DNA typing. Science, 254, 1745-1750.

Mansfield, D. C., Brown, A. F., Green, D. K., Carothers, A. D., Morris, S. W., Evans, H. J. & Wright, A. F. (1994). Automation of genetic linkage analysis using fluorescent microsatellite markers. Genomics, 24, 225-233.

Mullis, K. B. (1989). The polymerase chain reaction: Why it works. In H. A. Erlich, R. Gibbs & H. H. Kazazian, Jr. (Eds.), Current Communications in Molecular Biology: Polymerase Chain Reaction. (pp. 237-243) New York: Cold Spring Harbor Laboratory Press.

Mullis, K. B. (1990). The unusual origin of the polymerase chain reaction. Scientific American, April, 56-65.

Nelson, D. L., Ledbetter, S. A., Corbo, L., Victoria, M. F., Ramirez-Solis, R., Webster, T. D., Ledbetter, D. H. & Caskey, C. T. (1989). *Alu* polymerase chain reaction: A method for rapid isolation of human-specific sequences from complex DNA sources. Proceedings of the National Academy of Sciences, USA, 86, 6686-6690.

O'Connell, A. & Denome, R. M. (1999). Microsatellite locus amplification using deer antler DNA. BioTechniques, 26, 1086-1088.

Odelberg, S. J. & White, R. (1990). Repetitive DNA: Molecular structure, polymorphisms, and forensic applications. In H. C. Lee & R. E. Gaensslen (Eds.), DNA and Other Polymorphisms in Forensic Science. (pp. 26-44) Chicago: Year Book Medical Publishers, Inc.

Parker, P. G., Snow, A. A., Schug, M. D., Booton, G. C. & Fuerst, P. A. (1998). What molecules can tell us about populations: Choosing and using a molecular marker. Ecology, 79(2), 361-382.

Queller, D. C., Strassmann, J. E. & Hughes, C. R. (1993). Microsatellites and kinship. Trends in Ecology and Evolution, 8, 285-288.

Reed, P. W., Davies, J. L., Copeman, J. B., Bennett, S. T., Palmer, S. M., Pritchard, L. E., Gough, S. C. L., Kawaguchi, Y., Cordell, H. J., Balfour, K. M., Jenkins, S. C., Powell, E. E., Vignal, A. & Todd, J. A. (1994). Chromosome-specific microsatellite sets for fluorescence-based, semi-automated genome mapping. Nature Genetics, 7, 390-395.

Reynolds, R., Sensabaugh, G. & Blake, E. (1991). Analysis of genetic markers in forensic DNA samples using the polymerase chain reaction. Analytical Chemistry, 63(1), 2-15.

Scribner, K. T. & Pearce, J. M. (In Press). Microsatellites: Evolutionary and methodological background and empirical applications at individual, population, and phylogenetic levels. In A. Baker (Ed.), Molecular Methods in Ecology. (Chapter 7) London: Blackwell Scientific.

Sensabaugh, G. F. (1982). Biochemical markers of individuality. In R. Saferstein (Ed.), Forensic Science Handbook. (pp. 338-415) New Jersey: Prentice-Hall.

Sensabaugh, G. F. & von Beroldingen, C. (1989). The application of polymerase chain reaction in forensic science. In H. A. Erlich, R. Gibbs & H. H. Kazazian, Jr. (Eds.), Current Communications in Molecular Biology: Polymerase Chain Reaction. (pp. 147-150) New York: Cold Spring Harbor Laboratory Press.

Sensabaugh, G. F. (1991). Genetic typing of biological evidence using the polymerase chain reaction. In G. Berghaus, B. Brinkmann, C. Rittner & M. Staak (Eds.), DNA-Technology and Its Forensic Application. (pp. 33-40) New York: Springer-Verlag.

Sprecher, C. J., Puers, C., Lins, A. M. & Schumm, J. W. (1996). General approach to analysis of polymorphic short tandem repeat loci. BioTechniques, 20, 266-276.

Trent, R. J. (1997). Molecular Medicine: An Introductory Text, Second Edition. New York: Churchill Livingstone.

Waye, J. S. & Fourney, R. M. (1993). Forensic DNA typing of highly polymorphic VNTR loci. In R. Saferstein (Ed.), Forensic Science Handbook Vol. III. (pp. 358-415) New Jersey: Regents/Prentice-Hall.

Weir, B. S. (1992). Population genetics in the forensic DNA debate. Proceedings of the National Academy of Sciences, USA, 89, 11654-11659.

Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millasseau, P., Vaysseix, G. & Lathrop, M. (1992). A second-generation linkage map of the human genome. Nature, 359, 794-801.

Worley, J., Lee, S., Ma, M. S., Eisenberg, A., Chen, H. Y. & Mansfield, E. (1997). Fluorescence imaging in human identity testing. BioTechniques, 23, 148-153.

Zietkiewicz, E., Rafalski, A. & Labuda, D. (1994). Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. Genomics, 20, 176-183.