ENHANCED SAMPLE PREPARATION AND DATA INTERPRETATION STRATEGIES USING MASSIVELY PARALLEL SEQUENCING FOR HUMAN IDENTIFICATION IN MISSING PERSONS' AND DVI CASEWORK

_____

Dissertation

Presented to

The Faculty of the Department of Forensic Science

Sam Houston State University

_____

In Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

_____

by

Kyleen Elizabeth Elwick

December, 2018

ENHANCED SAMPLE PREPARATION AND DATA INTERPRETATION STRATEGIES USING MASSIVELY PARALLEL SEQUENCING FOR HUMAN IDENTIFICATION IN MISSING PERSONS' AND DVI CASEWORK

by

Kyleen Elizabeth Elwick

_____

APPROVED:

Sheree Hughes-Stamm, PhD
Committee Director

David Gangitano, PhD
Committee Co-Director

Bobby LaRue, Jr., PhD
Committee Member

Charla Marshall, PhD
Committee Member

Rachel Houston, PhD
Committee Member

Phillip Lyons, PhD
Dean, College of Criminal Justice

DEDICATION

"You sort of start thinking anything's possible if you've got enough nerve." – Ginny Weasley

I would first like to thank my wonderful advisor Dr. Sheree Hughes-Stamm for guiding me and helping me to accomplish more than I ever imagined possible. She provided me with numerous opportunities that have shaped me as a scientist and furthered my career. I also commend her for putting up with my endless sass and constant demands.

I would also like to thank my PhD cohort and Team DNA for their continual support, willingness to listen to my rants, all of the laughs and jokes we have shared, and for pulling me out of the Duckweeds on many occasions.

I must express my overwhelming gratitude to my family for their tremendous support and continuous encouragement throughout my many years of study. First to my grandparents, who have always shown support for any task I have taken on. My grandma Elwick graciously took me in when I started my college career and in that time we enjoyed many shopping trips and episodes of NCIS together. To my grandpa Zaborowski for the many hours spent sharing intellectual conversations, for which I am extremely grateful. And to my grandma Zaborowski for our fun trips together and always believing that one day I would be a doctor. Next, I would like to thank my amazing younger siblings for always being there for me when I needed them. Taylore and Cameron, you have always supported me and been proud of my accomplishments. Our laughs, jokes, and adventures have relieved all of my stresses that have come from my educational endeavors. Finally, I must thank my parents for their continuous encouragement and unfailing support. I've cherished all of the conversations and stern lectures I have received from the both of you.

Dad, you have always been there to help me through difficult situations and guide me in the right direction when I needed it. I also appreciate your interest in my research and the many questions about the work that I do. Mom, you have always pushed me the hardest to be the best that I can be and always encouraged me to do better. You have had the utmost faith in me since the beginning and you are the reason I have made it this far in my career.

To everyone who has supported me through my educational journey, I am eternally grateful. And with that… *"Mischief Managed." – Sirius Black*

ABSTRACT

Elwick, Kyleen Elizabeth, *Enhanced sample preparation and data interpretation strategies using massively parallel sequencing for human identification in missing persons' and DVI casework.* Doctor of Philosophy (Forensic Science), December, 2018, Sam Houston State University, Huntsville, Texas.

When the remains of victims from mass disasters, military conflicts, or missing persons' cases are recovered, identification is the most important objective. The recovered unidentified remains may be intact, fragmented, comingled, highly decomposed, or skeletonized. The DNA within these tissues is often degraded, damaged, and/or contains inhibitory agents depending on the environment in which the remains were discovered. This project explores the use of traditional genotyping and newer DNA sequencing technologies for the identification of challenging human remains commonly recovered from mass disasters and missing persons' cases. The results of this study will provide the forensic community with additional information on the comparative performance of massively parallel sequencing (MPS) chemistries and platforms with compromised samples, particularly highly inhibited samples.

This study was comprised of four projects. First, two CE-based STR megaplex kits (GlobalFiler® PCR Amplification and Investigator® 24plex QS kits) were evaluated for their tolerance to PCR inhibitors (humic acid, melanin, hematin, collagen, calcium) and overall sensitivity of detection for high and low quantity (1 ng and 0.1 ng) DNA samples. The results suggested that the GlobalFiler® kit was more sensitive down to 7.8 pg of DNA while the Investigator® kit was more tolerant to all PCR inhibitors at both DNA concentrations. The GlobalFiler® kit produced more alleles, higher peak heights, and higher peak height ratios when determining sensitivity. Conversely, the Investigator® kit

produced more alleles and balanced profiles for every inhibitor and inhibitor concentration than the GlobalFiler® kit.

Second, two MPS chemistries and platforms (Ion AmpliSeq™ kit on the Ion PGM and the ForenSeq™ kit on the MiSeq FGx™) were evaluated side-by-side using the same inhibited DNA samples. The AmpliSeq™ and ForenSeq™ kits were found to be tolerant and susceptible to different common PCR inhibitors. The AmpliSeq™ chemistry demonstrated tolerance to collagen and calcium; however, it was highly susceptible to humic acid and hematin. Conversely, the ForenSeq™ kit showed extreme tolerance to hematin and calcium inhibitors but was greatly affected by melanin.

The third study focused on determining the effectiveness of common DNA extraction methods to remove inhibitors from forensically relevant samples and their downstream compatibility with two MPS chemistries. Three substrates (blood, hair, and bone) were spiked with high concentrations of four inhibitors (humic acid, melanin, hematin, and calcium) and extracted using five DNA extraction methods (DNA IQ™, QIAamp® DNA Investigator, PrepFiler®, and two total demineralization protocols (bone only)). The results showed that all extraction methods were able to efficiently remove all PCR inhibitors with no sign of inhibition and provide sufficiently pure DNA extracts for sequencing. Although the amount of DNA recovered using the different extraction methods differed, the sequencing data indicated that none of the extraction methods negatively influenced the downstream sequencing performance on either MPS system.

The fourth and final study reports the comparative performance of two MPS systems when sequencing challenging human skeletal remains. Thermally degraded, embalmed, cremated, burned, and decomposed bones and teeth (N = 24) were extracted

using a total demineralization protocol and processed with two MPS chemistries and platforms in addition to traditional CE-based STR typing. The results demonstrated that CE-based STR profiling was still a valuable approach by providing at least a partial DNA profile for every sample, whereas MPS did fail to produce a profile in some instances. However, these MPS chemistries are still not fully optimized to tolerate such difficult samples and further optimization is warranted. Conversely, MPS has the capability to analyze more markers and multiple marker systems (STRs, SNPs, etc.) simultaneously. Therefore, even though some CE samples produced more complete profiles, the additional markers within MPS multiplexes may result in higher powers of discrimination for identification, and thereby provide results to assist with solving missing persons', forensic, and DVI cases.

Keywords: Forensic DNA, Short tandem repeats, Massively parallel sequencing, Missing persons, Mass disaster victim identification, Human remains, PCR inhibitors

ACKNOWLEDGEMENTS

This project was funded by NIJ Award No. NIJ 2015-DN-BX-K066, awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect those of Department of Justice.

The author would also like to thank Thermo Fisher Scientific, especially Joseph Chang, Joshua Abernathy, Sheri Olson, and Angela Lackey for all of their help and support throughout this project.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Page

ABBREVIATIONS

AFDIL	Armed Forces DNA Identification Laboratory
AFIS	Automated Fingerprint Identification System
AFMES	Armed Forces Medical Examiner System
aiSNP	Ancestry informative SNP
bp	Base pair
BSA	Bovine serum albumin
CE	Capillary electrophoresis
CODIS	Combined DNA Index System
DNA	Deoxyribonucleic acid
ddNTP	Dideoxynucleotide triphosphate
dNTP	Deoxyribonucleotide triphosphate
DoD	Department of Defense
HID	Human identification
FBI	Federal Bureau of Investigation
FDIS	Federal DNA Index System
ICMP	International Commission on Missing Persons
iiSNP	Identity informative SNP
INTERPOL	International Criminal Police Organization
LDIS	Local DNA Index System
liSNP	Lineage informative SNP
MFI	Mass fatality incident
MP	Missing person

MPS	Massively Parallel Sequencing
mtDNA	Mitochondrial DNA
mtGenome	Mitochondrial Genome
NamUs	National Missing and Unidentified Persons System
NDIS	National DNA Index System
NGS	Next Generation Sequencing
NIJ	National Institute of Justice
NMPDD	National Missing Person DNA Database
PCR	Polymerase chain reaction
piSNP	Phenotypic informative SNP
qPCR	Real-time quantitative PCR
SDIS	State DNA Index System
SGS	Second Generation Sequencing
SNP	Single nucleotide polymorphism
STR	Short tandem repeat
UNTCHI	University of North Texas Center for Human Identification
UNTHSC	University of North Texas Health Science Center
UP	Unidentified Persons

GLOSSARY

Allele	Version of a gene or locus
Allelic dropout	Failure to detect an allele within a sample or failure to amplify during PCR
Capillary electrophoresis (CE)	An electrophoretic technique for size-based separation of DNA molecules through a glass capillary
Deoxyribonucleic acid	Double stranded genetic material of many organisms
Heterozygosity	Different alleles at one or more loci on homologous chromosomes
Inhibitor	A biological or chemical matrix interference that prevents polymerase chain reaction (PCR)
Noise	Background signal detected by the instrument performing analysis
Peak Height	The intensity of a peak (quantity of PCR product) measured by the relative fluorescent units (RFUs)
Peak height imbalance	A notable difference in peak heights based on RFUs
Peak height ratio	The measurement of one allele peak height over a second allele peak height at a single locus
Polymerase chain reaction (PCR)	A process that yields millions of copies of a DNA target region
Power of discrimination (PD)	The potential power of a marker to differentiate between any two people at random
Short tandem repeats (STRs)	Microsatellite repeat units between 2-6 bp

Single nucleotide polymorphism (SNP)	A polymorphic variation at a single nucleotide

CHAPTER I

Introduction

Missing Persons and Mass Disasters

Thousands of migrants and refugees have gone missing or died while attempting to cross borders and seas around the world [1]. In addition, several military conflicts worldwide have resulted in mass fatalities and mass graves. As a result, the remains of many victims need to be identified and repatriated to their home country and/or family members. Missing persons' (MP) investigations consist of the search and recovery of bodies, identification of remains, and recovery of evidence that may determine the cause of death [2].

There are several programs worldwide dedicated to processing the remains of missing and unidentified persons. The United States has a limited number of programs to record and help determine the identity of unknown remains. Laboratories performing missing persons' casework in the United States must be able to process autosomal STRs, Y-STRs, and mitochondrial DNA (mtDNA). Currently, very few laboratories in the US have those capabilities including the Federal Bureau of Investigation (FBI), the University of North Texas Center for Human Identification (UNTCHI) [3], and the Armed Forces DNA Identification Laboratory (AFDIL).

The FBI created the Combined DNA Index System (CODIS), and it contains the National DNA Index System (NDIS), which receive DNA profiles from federal (FDIS), state (SDIS), and local (LDIS) forensic laboratories. CODIS was created in 1990 and there are over 190 public law enforcement laboratories participating in NDIS throughout the US and over 90 laboratories internationally. NDIS allows all levels of laboratories to compare

DNA profiles and associate results to known offenders. Also within NDIS is the National Missing Person DNA Database (NMPDD). The NMPDD program contains three indexes for DNA profiles: biological relative of missing persons, unidentified human remains, and missing persons [4, 5].

Texas contains the largest missing persons' identification program in the United States and one of the largest internationally. Texas was the first state with a missing persons' database and the first to participate in CODIS at the federal level. The University of North Texas Health Science Center (UNTHSC) created the Texas Missing Persons DNA Database and the Center for Human Identification (CHI) in 2001. They collaborate with law enforcement, medical examiner systems, and families of missing persons to collect reference samples for testing [3]. The UNTCHI provides services including forensic genetic and anthropological examinations for criminal casework and missing persons identification. UNTCHI also administers the National Missing and Unidentified Persons System (NamUS) that was established by the National Institute of Justice (NIJ) in 2005 [6]. It was the first and one of the only laboratories capable of analyzing nuclear DNA and mtDNA at the time [3]. Currently, UNTCHI has collected and processed the majority of missing person samples contained within CODIS in the United States [7].

NamUs was created in 2005 as a result of the National Missing Persons Task Force attempting to solve missing and unidentified person cases. In 2007, NamUs launched the Unidentified Persons (UP) database, and the Missing Persons (MP) database followed a year later. In 2011, UNTCHI assumed responsibility for the operations and management of NamUs with continuous regulation and financial support by the NIJ. The following year, an Analytic Division and the NamUs Automated Fingerprint Identification System

(AFIS)/Fingerprint Unit were created to further facilitate the identification of missing and unidentified persons. According to NamUs, over 600,000 people go missing in the US each year and tens of thousands of those missing persons remain missing for more than a year. They estimate that 4,400 unidentified bodies are recovered every year and around 1,000 of those remain unidentified for more than a year [6, 9].

The Armed Forces DNA Identification Laboratory (AFDIL) is a division of the Armed Forces Medical Examiner System (AFMES) and the only DNA testing facility for human remains used by the Department of Defense (DoD). The AFDIL is tasked with the identification of human remains from current and past conflicts including the Korean War, World War II, and the Cold War, as well as any major mass disaster they are called upon to assist [8].

In addition to the United States, the European Union (EU) also houses two major programs for missing and unidentified persons including the International Commission on Missing Persons (ICMP) and the International Criminal Police Organization (INTERPOL). ICMP was created in 1996 after the Dayton Peace Agreement, which ended the conflict in the former Yugoslavia. The Yugoslavian conflict resulted in around 40,000 missing persons between the years of 1991 and 1995 [10], creating the need of an organization to address these types of situations. A few years later, in 2001, ICMP's missing persons DNA identification system was established along with a specialized missing persons database, the Identification Data Management System (iDMS) for the management of over 150,000 international missing persons cases. The ICMP work with international governments and organizations around the world to resolve missing persons' cases resulting from mass disasters, conflict, and crime. ICMP has aided governments with the excavation of over

3,000 mass gravesites and processes the largest amount of human remains worldwide; to date, they have identified over 19,000 missing persons from events including the Asian Tsunami in 2004, Hurricane Katrina in 2005, and many military and societal conflicts in Iraq, Colombia, Chile, and Libya [11]. They were also the first MP agency to implement a massively parallel sequencing (MPS) workflow for casework. In October 2017, the ICMP collaborated with QIAGEN to implement the complete GeneReader NGS System and workflow into the MP casework laboratory [12].

Mass disasters or mass fatality incidents (MFI) are defined as an unexpected event causing death and/or injury more people than local agencies can manage [13-15], and may occur locally, nationally, or internationally [15]. Mass disaster events are commonly referred to as either "open" incidents because the number of victims is unknown, or "closed" incidents consisting of a known number of victims, such as plane crashes [14, 15]. In addition, they can be classified as environmental or natural disasters (hurricanes, earthquakes, and tsunamis), medical (disease and famine), vehicle (plane, car, train, and watercraft), industrial (fires and explosions), and terrorist attacks (biological, chemical, explosive, and nuclear attacks) [13, 14, 16-18]. They can be subcategorized into major, mass, or catastrophic events, depending on the number of fatalities [15]. Missing persons' cases can also be referred to as mass disasters taking place over a longer period of time [16]. The type of incident, time elapsed since death, and the local environmental conditions are all factors that will determine the state of preservation of the human remains, and therefore may influence which methods may be most successful when identifying victims.

Mass disasters and missing persons' cases have the same common goal of victim identification. The most important reason for victim identification is to bring closure to

family and friends, in addition to being necessary for civil or criminal investigations. Other reasons may include inheritance, collection of insurance policies, and getting remarried [13-15, 19]. Missing persons', war remains, and mass disaster cases can be challenging to solve and involve considerable effort by investigators and laboratory personnel. However, forensic genetics is continuously evolving and forensic scientists have the tools to effectively repatriate families with their loved ones.

Identification Methods

Methods used to identify remains include the analysis of skeletal features (forensic anthropology), dental records (forensic odontology), fingerprints, characteristic marks (tattoos or scars), medical devices, unique personal effects, and DNA [14, 16, 20, 21]. However, only DNA analysis, fingerprints, and odontology are considered primary methods of identification on which a death certificate can be issued [22].

Autopsies performed by a forensic pathologist or skeletal analysis performed by a forensic anthropologist may be helpful in identifying a biological profile (gender, sex, height, race, approximate age, and any unique marks or medical implants) of the remains. Anthropologists are able to distinguish between human and animal remains and reassemble commingled remains, which is vital in these types of situations [19]. They are also able to estimate the number of deceased individuals as well as offer an opinion on potential trauma (ante-, peri-, or post-mortem) [23] and the best skeletal elements to use for DNA analysis [16].

Forensic odontologists are commonly requested in the event of mass disaster or missing persons' cases for identification and/or to estimate the age of children [19] based

on their dentition. Post-mortem dental charts are compared to ante-mortem (AM) records for identification purposes [23]; however, AM records are not always available.

Fingerprints are another primary tool for identification if the human remains are not too fragmented or decomposed. Fingerprints consist of friction ridge skin that do not change throughout one's lifetime (except scarring), and are unique to each individual [14], even twins. They can be taken from the fingers, palms, and feet of victims and compared with prints taken from the victim's home, possessions, or national databases such as AFIS. However, fingerprints are not useful when there are no prints to compare to, amputation of hands and feet has occurred, or if remains are highly fragmented, burnt, or decomposed [14, 19].

Other factors that could make identification difficult after a mass disaster include the number of fatalities, the extent of fragmentation, commingling, condition of the bodies, availability of medical and dental records, accessibility of remains, and the availability of AM reference samples [13, 15, 24]. In some cases, the identification of individuals is not possible due to the lack of sufficient information available for analysis or comparison [15]. Often, the only means available for the identification of remains is to use a DNA-based approach. For identification, a DNA profile from the human remains is either compared directly to profiles generated from the victim's own personal affects or via kinship analysis with relatives. DNA reference material for kinship analysis usually includes blood or buccal swabs from multiple members of the immediate family. For direct matching, DNA profiles can be produced from personal effects such as razors, toothbrush, hair brush, dirty laundry, etc. and be compared to a DNA profile generated from a particular set of human remains [13, 14, 16, 17]. However, in most cases, personal effects cannot be obtained, and

family reference samples must be used to determine the identity of the victim through family pedigrees and kinship analysis [16]. This may be particularly difficult if several relatives perish together or if only distant relatives are available to provide AM samples for analysis [25].

The most reliable identification method for mass disaster and missing persons investigations is DNA profiling [14, 16]. Identification by DNA typing is invaluable because even the smallest fragments of tissue can be identified and repatriated to the correct set of human remains [14, 19]. DNA typing can be used to determine familial relationships between sets of unknown remains. However, the use of DNA for human identification also has some limitations such as being a more expensive and time-consuming process than alternative approaches. DNA analysis must be performed by highly trained staff in a dedicated laboratory and identification is not always successful depending on the quantity and quality of the DNA extracted from challenging samples [19].

There are many potential approaches to sampling human remains after an incident for identification. According to Interpol [22], collection of blood or saliva on Flinders Technology Australia (FTA®) or a cotton swab is recommended if an individual is not decomposed and is intact [24].

For remains that are not intact, muscle tissue is the suggested sample type for DNA analysis barring decomposition of the deceased [22]. However, bones and teeth are among the most reliable DNA sources [17, 24, 26] from highly decomposed remains because they are fairly resilient to DNA degradation [16]. Additionally, they are often the only material available for identification when the remains are skeletonized [26, 27]. This is particularly true in missing persons' cases.

Bone density is one major factor in determining the preservation of DNA in bone [28]. Cancellous bone may contain a high amount of DNA, but may not be as well protected from the elements and degradation as dense cortical bone [17]. Cortical bone from weight bearing bones (long bones such as the femur and tibia) is the traditional choice for obtaining DNA for identification from bones [17, 24, 28]. However, a few studies have shown that small cancellous bones like fingers and toes can provide comparable or even higher concentrations of DNA than cortical bones [29-31]. Teeth are also valuable samples for DNA analysis because they are relatively protected from the environment and provide a rich DNA source. DNA may be recovered from the cementum and pulp, with molars typically being the first choice when sourcing DNA from teeth. However, teeth may not be as valuable if they are decayed or absent [24].

Hair and nails are often recovered from highly decomposed or skeletonized remains for DNA identification. Allouche et al. [32] was able to obtain full DNA profiles from fingernails of decomposed cadavers up to 6 months, but fingernails may include exogenous DNA, resulting in mixed profiles unfavorable for identification. When hair samples are recovered, DNA can be extracted from either the root or the shaft. Habib et al. [33] was able to recover full profiles from four out of five samples using 6 hair roots at a time. In addition, Pfeiffer et al. [34] demonstrated successful recovery of mitochondrial genetic profiles from head, pubic, and axillary hair shafts.

Challenging Remains

Mass disasters and missing persons' cases often present with remains that are fragmented and highly decomposed or skeletonized, commingled from multiple victims, contaminated with environmental elements, and/or severely heat damaged [15, 28]. Human

remains in these cases may range from pristine to extremely compromised [16]. They may also be exposed to adverse climates (high temperatures and humidity), which can increase the rate of decomposition [28, 35]. The DNA contained within these tissues may be highly degraded, damaged, and/or inhibited as a result of these adverse conditions.

DNA Damage and Degradation

Highly degraded, damaged, or environmentally effected samples can be problematic and may reduce the success of downstream DNA typing. Environmental insults such as UV exposure, temperature, fire, humidity, and microbial infestation may result in severe damage and degradation of DNA in biological samples [28, 36].

DNA degradation and/or damage can occur from multiple processes that include either enzymatic degradation or nonenzymatic degradation. Enzymatic degradation causes nucleases in the body to fragment the DNA during cell death. When cell membranes rupture, fluids from the cell are released, which increases the growth of microorganisms [37]. Nonenzymatic degradation can be produced by hydrolytic reactions, DNA crosslinkages, oxidative reactions, and radiation. The weakest bond in a DNA strand is the glycosidic bond between a sugar and a nucleotide base and is the primary site for a hydrolytic attack, causing the loss of a nucleotide base [38]. DNA crosslinkages may occur when an abasic site on a DNA strand is available and can transpire with proteins or between the sugar and amino group [39, 40]. However, DNA crosslinking can be slowed or avoided by storing DNA in cooler environments [41] such as refrigerators and freezers. Another type of reaction causing DNA damage is oxidative reactions, which are triggered by the actions of aerobic microorganisms [38]. Pyrimidines, especially thymine, are more prone to oxidative damage than purines [42]; however, most oxidative damage occurs in the form

of base removal, crosslinkages, and modifications to the sugar, cytosine, and thymine molecules [38]. UV radiation has the ability to produce many types of DNA damage including oxidative damage, breaks in the DNA strand to one or both strands, crosslinkages, primer dimers, and modification or destruction of sugar and nucleotide molecules [43]. Any combination of these insults can cause DNA damage and/or degradation in samples that make DNA typing more difficult in mass disaster and missing persons' cases.

Inhibitors

Inhibitors are chemical or biological matrix interferences that affect DNA extraction and/or PCR amplification processes during DNA analysis [44]. Environmental conditions such as burial in soil can introduce added complications for DNA typing in the form of PCR inhibitors such as humic and fumic acid [45, 46]. In addition, biological tissues themselves including bone, hair, teeth, and blood contain various PCR inhibitors that may be co-extracted with the DNA in these samples [46-49]. Commonly co-extracted inhibitors include humic acid, hematin, collagen, calcium, melanin, indigo, bile salt, and urea, each having different mechanisms by which they inhibit DNA amplification [44, 47, 48, 50, 51] (Table 1.1). These co-extracted inhibitors (humic acid, melanin, hematin) may also discolor the DNA extract a yellow to a red or brown color [47, 52].

There are three potential mechanisms by which to inhibit PCR: the inhibitor binds to the polymerase, the inhibitor binds to the DNA, or the inhibitor interacts with the polymerase during primer extension [53]. Calcium is the main inorganic component in bone making up two-thirds of its structure. Calcium may be co-extracted with DNA and likely inhibits the *Taq* DNA polymerase during PCR. Calcium and magnesium are both

divalent cations with similar structures so they may compete with each other during PCR, inhibiting *Taq* and reducing the total amount of product and PCR efficiency [44, 53].

Humic acid is one of the major components in soil, which is comprised of decomposed plant and animal tissue [44] and is often found in buried human skeletal remains [53]. Humic acid is a large molecule (227.2 g/mol) and most likely inhibits PCR by sequence specific binding to DNA, thereby affecting the availability of template DNA during amplification [44, 53]. Collagen is a protein comprising approximately 28% of organic bone tissue and other connective tissues, and may also be co-extracted from skeletal remains during the DNA extraction process. The triple helix structure of collagen may intercalate with and wrap around the DNA molecule inhibiting amplification by binding to DNA and reducing the efficiency of the *Taq* DNA polymerase [44, 53].

Hematin is a metal chelating agent found in red blood cells and most likely inhibits PCR by binding to the DNA itself [44, 53]; although other sources propose that it is an inhibitor of the *Taq* DNA polymerase [53]. Melanin is a pigment found in human hair and skin. The proposed mechanisms by which this inhibitor causes interference with PCR include intercalating between the DNA base pairs and reversibly binding to the DNA polymerase [44], but the most likely mechanism is sequence specific binding to the DNA, limiting the available template [44, 53].

Bile salts are found in feces and they inhibit amplification by reducing the availability of template DNA [44]. Urea is an organic waste product found in urine that inhibits PCR by binding to the DNA and reducing the activity of *Taq* DNA polymerase [44]. Ethylenediaminetetraacetic acid (EDTA) is a metal ion chelating agent often found in extraction or digestion buffers that can interfere with PCR by binding metal ions and

reducing the PCR efficiency. Phenol is an organic compound potentially carried over during DNA extraction from phenol/chloroform. The inhibition mechanism is likely due to the phenol binding to the DNA and *Taq* polymerase inhibition [44].

Table 1.1 Sources and mechanisms of common PCR inhibitors

Inhibitor	Source(s) of Inhibitor	Mechanisms of Inhibitors
Humic Acid	Soil	Binds DNA, affects availability of template DNA
Hematin	Blood	Binds DNA, inhibits *Taq* polymerase
Collagen	Bone, connective tissue	Binds DNA, inhibits *Taq* polymerase
Calcium	Bone	Inhibits *Taq* polymerase
Melanin	Hair, skin	Binds DNA, limiting the amount of template DNA
Bile Salt	Feces	Reduces available template DNA
Urea	Urine	Binds DNA, inhibits *Taq* polymerase
EDTA	Extraction buffer	Binds ions, reduces reaction efficiency
Phenol	PCIA	Binds DNA, inhibits *Taq* polymerase

PCR inhibition is the most common cause of PCR failure when adequate amounts of DNA are present [47, 51], and therefore it is important to remove PCR inhibitors from samples prior to PCR amplification for successful DNA typing [47, 48, 50]. Inhibitors can cause a total failure of the PCR resulting in no amplified products, or simply reduce the efficiency of the PCR and produce negative downstream effects such as allele dropout, lower peak heights, peak height imbalance, stutter, locus-specific dropout, and poor sensitivity [44, 53]. In addition, PCR inhibition can result in inaccurate DNA quantification

[51] when using PCR based methods such as real-time quantitative PCR (qPCR) resulting in an underestimation of the amount of amplifiable DNA in samples.

There are two basic approaches to eliminating the negative effects of PCR inhibitors on downstream DNA profiling: 1) reduce the effects of inhibitors in DNA extracts by altering the reagents in the PCR reaction, and 2) remove PCR inhibitors during the DNA extraction process [47,50,52,54]. Methods that reduce the effects of inhibitors include diluting the DNA sample, adding bovine serum albumin (BSA), or adding more *Taq* polymerase [51]. The most common method used to reduce the effects of inhibitors during PCR is adding BSA to the reaction master mix. BSA blocks PCR inhibitors and indirectly promotes polymerase activity by binding to the surface of the inhibitor allowing PCR components to be free in the reaction mix [55]. This approach has been widely used with degraded and inhibited biological samples to overcome low temperature co-fired ceramic (LTCC) mediated inhibition of PCR [55]. BSA is also a common component in most commercial and custom STR kits as a strategy to improve their kits and make them more tolerant to PCR inhibitors [56-59]. Another technique to reduce the effects of inhibitors during PCR involves diluting the extract with DNA-free water or low TE buffer prior to amplification, which dilutes the inhibitor to a level that allows successful amplification. Although this is a simple and effective technique, it may not be suitable for samples with very low amounts of DNA available such as from bone, teeth, and decomposing remains [47,48,54]. Increasing the amount of *Taq* polymerase to use as a decoy in the PCR reaction may be effective in samples with pristine DNA. However, if the sample is degraded or low-template, the added polymerase may intensify the possible contaminates in the sample [47]. Conversely, another approach to inhibitor removal is

effective DNA extraction or purification [48,50]. Silica-based extraction using chaotropic salts to bind to silica membranes has been shown to be more effective at removing inhibitors from degraded bone samples than the “classic” phenol/chloroform technique [60,61].

DNA Preparation and Extraction

For the successful identification of human remains, the extraction of adequate amounts (>100 pg) of clean, good quality DNA is necessary [16, 63]. Many DNA extraction methods have been developed to purify as much DNA as possible from biological material while minimizing the co-extraction of PCR inhibitors [26, 50]. The most common techniques include silica-based extraction methods [26, 47, 49], various phenol-chloroform (organic) protocols, and Chelex® [26, 48, 50, 64]. Purification of DNA by silica binding (via beads or membrane) wash and elute methods have become a favored method for the extraction of DNA from a wide variety of forensic samples such as buccal swabs, blood, muscle tissue, and cigarette butts [63]. When silica-based spin columns are used, nucleic acids are attracted to the silica filter membrane due to high concentrations of chaotropic salts, while all other contaminants are washed away. Additionally, systems that use magnetic-based silica beads such as DNA IQ™, PrepFiler®, and EZ1 DNA Investigator kit also depend on DNA binding to the beads in a favorable ionic environment where contaminants are washed away [18]. Silica-based methods have been designed to maximize the removal of PCR inhibitors, are less time consuming and are amenable to automation. However, they are more expensive [26, 65].

Phenol-chloroform extraction is a well-established extraction method involving the use of organic, hazardous solvents and many time-consuming steps including various

precipitation and filtration techniques. Phenol-chloroform itself can also act as a PCR inhibitor if not removed prior to genotyping [26, 64]. Chelex® is another DNA extraction method that has historically been used within forensic laboratories. The polarity of the Chelex® resin binds other polar molecules and leaves the non-polar DNA in solution. This method has proven to be rapid, but is unable to effectively remove PCR inhibitors [66]. This is most likely due to Chelex® extractions being performed in a single tube with no wash steps, and the Chelex® resin itself is also a PCR inhibitor if carried over to the PCR reaction [46].

Conflicting reports on the effectiveness of various DNA extraction methods to remove common PCR inhibitors from various samples prior to DNA typing have been published. When extracting forensic-type samples, organic extraction, Chelex®, and commercially available kits such as QIAamp® DNA Investigator, DNA IQ™, and PrepFiler® are commonly used in forensic crime laboratories. Several studies have suggested that commercial methods have been more effective in removing inhibitors and providing higher DNA yields than Chelex® ion-exchange resin or phenol/chloroform organic extraction [67-72]. However, one study demonstrated better results (higher yield and more full profiles) using Chelex® with added Proteinase K as opposed to DNA IQ™ for the extraction of cigarette butts [73]. A second study demonstrated higher DNA yield using organic extraction over silica-based methods (PrepFiler® and DNA IQ™); however, all samples produced full profiles and no inhibition was observed [74].

Even though there are many viable DNA extraction methods for forensic samples, skeletal remains are more complex, and additional steps prior to extraction are necessary to effectively recover DNA from hard tissues.

Bone and Tooth Samples

DNA extraction methods for hard tissues such as bone or teeth require more aggressive sample preparation and lysis methods prior to purification of DNA. Most commonly these include powdering of bone/tooth tissue into a fine powder and a total demineralization digestion step. Total demineralization of hard tissues helps to break down the difficult components or organic materials (hydroxyapatite) of bone and teeth using high concentrations of EDTA and abundant proteins with Proteinase K. DNA can be preserved in bone by crystal aggregates. Weiner and Price [75] discovered that DNA protected by crystal aggregates cannot be broken down when the collagen matrix is isolated by oxidation with sodium hypochlorite (bleach), especially when the bone has been powdered. Additionally, DeNiro and Weiner [76] demonstrated that fragments of collagen and possibly some proteins were also preserved within the aggregate crystals and were not damaged by bleach. Using a total demineralization protocol for hard tissue samples is important because EDTA demineralizes the hard tissue (crystal aggregates) causing complete dissolution of the bone or tooth [77].

In 2007, Loreille et al. [77] developed a total demineralization organic extraction protocol specifically for bone and tooth tissue. Fourteen bones ranging from 5-100 years old were powdered with two different methods (Freezer Mill and Waring MC2 blender cup). Samples from each powdering method were extracted using two separate protocols, the AFDIL casework protocol described in Edson et al. [78] and a total demineralization protocol. The AFDIL casework protocol described by Edson suggests incubating 1-2 g of bone powder in 3 mL of extraction buffer (10 mM Tris, pH 8; 100 mM NaCl; 50 mM EDTA, pH 8; 0.5% SDS) and 100 µL of Proteinase K at 56ºC overnight with gentle

agitation. Loreille et al. determined that 15 mL of 0.5 M EDTA was necessary to fully dissolve 1 gram of bone powder. In this study, 9-18 mL of extraction buffer (EDTA 0.5 M, 1% lauroyl-sarcosinate) and 200 µL of Proteinase K were incubated in shaker at 56ºC overnight. The lysates were then extracted with phenol/chloroform/isoamyl alcohol (25:24:1) and concentrated using a 30 kDa Amicon Ultra-15, Centricon+20, or Centriplus filter leaving 2 mL of extract. The remaining extract was transferred to a Centricon 30 column and washed three times for a final volume of 100 µL. The yields between the total demineralization protocol and standard protocol were observed as well as the differences in grinding hard tissues with the Freezer Mill versus the blender cup. For every sample, the total demineralization protocol yielded between 2.5 and 100+ times more DNA than the casework protocol with an approximate average of 4.6 times more DNA. They also determined that there were no benefits of using one grinding method over the other.

In this study, Davoren et al. [61] processed 20 femur samples using the ICMP silica protocol (QIAGEN Blood Maxi Kit) with modifications and compared it to traditional phenol/chloroform method to determine DNA quantity and quality of samples recovered from mass gravesites. The ICMP silica method begins with the bone powder being incubated for 18 hours at 56ºC in 15 mL of ATL extraction buffer with 10 mg of Proteinase K and 300 µL of 1 M DTT. Following the overnight incubation, a second digestion is performed by adding 14 mL of AL buffer and incubating at 70ºC for 1 hour. Ethanol (22 mL of 96%) is added to the lysate and bound to the Blood Maxi column. The column is then washed with QIAGEN buffers (AW1, AW2) and the DNA is eluted twice in 3 mL of AE buffer at 72ºC. The 6 mL of eluted DNA was then concentrated using a Centriplus YM-100 column for a volume of 50 µL and then it was washed with 2 mL of water and

centrifuged until the retentate was again at 50 µL. The Centricon membrane was washed with water, added to the extract, and concentrated again for a final volume of 100 µL. The organic protocol started with an overnight digestion, extraction with PCIA, and then purification and concentration of the DNA extract with a Centriplus YM-100 exactly as with the ICMP laboratory protocol. Following qPCR of the samples, the ICMP method demonstrated a lower cycle number (27 cycles) to reach the early log phase than the organic protocol (30 cycles), which means samples extracted with the organic protocol exhibited higher levels of inhibition. qPCR also demonstrated that the ICMP method produced DNA quantities three times higher than the organic protocol. The silica method produced full profiles for all 20 bones samples, whereas the organic method failed to produce 6 profiles [61]. Overall, the ICMP silica-based method performed more optimally, but the process was very laborious and time consuming.

In 2012, Amory et al. [10] reported success when automating the extraction protocol developed by Davoren et al. [61]. The method developed by Davoren et al. was used to identify many of the Yugoslavian MP remains; however, this process required high amounts of bone (either 5.6 g or 9.8 g) and reagents. Amory et al. developed a more efficient protocol that used less starting material (0.5 g) and produced higher quality STR results. In this study, ICMP's original protocol [61] was compared to an automated full demineralization (FD) protocol using 40 bone samples. Each sample was extracted five times, once with 2 g of bone powder and the original silica method (Maxi2g), once with 0.5 g of bone powder using the original silica method, twice with the full demineralization method (QIAGEN QIAquick kit) using 0.5 g, and once with the QIAquick Kit (0.5 g) automated on the QIAcube. The full demineralization protocol includes an incubation in

15 mL of lysis buffer (0.5 M EDTA and 1% N-lauroylsarcosinate) overnight at 56ºC. The lysate was concentrated to 300 µL using a 100K Amicon filter, mixed with five volumes of PB1 buffer, and filtered through a QIAquick column. The column was washed three times with PE buffer and the DNA was eluted in 50 µL of EB buffer. An additional QIAquick purification step was performed if the extract was severely inhibited as indicated during DNA quantification. The automated extraction processes began after the addition of PB1 to the lysate and additional purifications were performed on the QIAcube platform, if needed. DNA quantification results indicated that regardless of the protocol used, inhibitors were still present at low levels. However, decreasing the amount of starting material also decreased the amount of inhibitors. The Maxi2g protocol uses four times the amount of bone powder than the FD protocol. The results suggest that on a per gram basis, the Maxi2g never produced better results than the FD protocol, but the FD protocol produced higher yields for 52.5% of samples. It was determined that extra purification yielded Ct IPC values <30, no inhibitors were reported, and that additional purifications do not always remove all inhibitors. Overall, samples extracted using the FD protocol produced more profiles that were of sufficient quality for submission to the DNA database compared to the original silica-based protocol (62.5% vs. 47.5%). The automated FD protocol gave similar results to the manual FD protocol, is of comparable cost to the manual method, and reduces the risk of human error and contamination [10].

Marshall et al. developed a high-volume silica extraction protocol for bone samples that combines ultrafiltration and purification while allowing for sample extraction in up to 20 mL of buffer. Bone samples (0.5 g of bone powder) were pulverized, processed using a complete demineralization protocol, and extracted with either Hi-Flow® silica columns or

by phenol/chloroform organic extraction. DNA purity, recovery, and extraction efficiency were measured for both extraction protocols. Both methods were reported as being equally efficient in recovering DNA from bone, as both methods yielded similar DNA quantities. Any differences in results were thought to be due to inconsistencies in the bone itself and/or stochastic differences during amplification. Inhibitors were most likely present in samples after organic extraction but may have been removed in samples using the Hi-Flow protocol resulting in a DNA extract of higher purity. While both methods recovered similar quantities of DNA, more alleles were reported for samples extracted using the Hi-Flow protocol. In addition, the Hi-Flow method reduced hands on time (a difference of 4 hours) and also eliminated the use of hazardous materials [64].

Lee et al. used human genomic DNA spiked with hematin and humic acid to evaluate various purification methods [26]. In addition, bone samples from the Korean War were also used to compare DNA extraction methods for PCR inhibitor removal from old skeletal remains. DNA purification was performed using modifications of the QIAamp® DNA Mini and Maxi kits, QIAquick® PCR Purification kit, and the QIAamp® Mini spin columns coupled with buffers from the QIAquick® PCR Purification kit. Degraded and intact DNA from bones without inhibitors present showed little difference between the three purification methods in terms of DNA yield recovery. Furthermore, the three purification methods used to process the spiked inhibited samples displayed suitable inhibitor removal, with the exception of an increased C_T value (Δ 28.9) at the highest humic acid concentration (30 µg) using the QIAamp® Mini kit. When comparing bone DNA extraction methods, the full demineralization protocol in conjunction with the QIAamp®

Blood Maxi spin columns and QIAquick® PCR Purification buffers, produced a higher DNA yield and more efficiently removed PCR inhibitors than the other two methods [26].

Hu et al. also investigated the effects of various concentrations of inhibitors when mixed with control DNA to simulate challenging biological samples [50]. To remove inhibitors, four DNA extraction methods were evaluated including two silica-based methods, PowerClean® DNA Clean-Up kit and DNA IQ™ System, phenol-chloroform, and Chelex®-100. The PowerClean® DNA Clean-Up kit successfully removed all of the various PCR inhibitors except for indigo at higher concentrations (>1.998 µg/µL). The DNA IQ™ System is a widely used system in forensics laboratories for routine DNA extraction. However, there is some contention in the literature regarding the efficiency of this system to completely remove all types of PCR inhibitors [50]. In this comparative study [50], the phenol-chloroform and Chelex®-100 methods were much less effective at removing inhibitors than the two commercial silica-based kits examined. There are advantages and disadvantages to each method, but overall, the silica-based commercial kits removed the majority of inhibitors most effectively [50].

In another study, the ability of synchronous coefficient of drag alteration (SCODA) technology to purify DNA samples containing common PCR inhibitors was compared to the performance of the silica-based QIAquick® PCR Purification Kit [48]. Spiked DNA samples were purified with both kits and genotyped using the AmpF*l*STR® Identifiler® Plus PCR Amplification Kit. The samples purified using the SCODA technique showed no inhibition in downstream STR profiles, but the samples purified with the QIAquick® kit displayed internal PCR control (IPC) amplification failure for all melanin and humic acid samples. A difference in the color of the DNA extracts after purification with SCODA was

observed as being clear, whereas the QIAquick® extracts were still darkly colored, suggesting that these extracts most likely still contained inhibitors. Full STR profiles were produced from all samples purified using SCODA, while samples purified using the QIAquick® columns yielded mixed results (majority with full profiles, one partial profile, and three failed amplifications) [48]. However, this is a very niche and labor-intensive method requiring specialized equipment, and is not amenable for routine use in forensic laboratories.

In a study conducted by Kuś et al. [79], three different extraction methods were compared using fragments of bones and teeth in various conditions ranging from a few months to 70 years after death; half of the samples were either from a medicolegal autopsy or from a criminal case and the remaining half were exhumed from graves. The three DNA extraction methods evaluated were organic (phenol/chloroform), PrepFiler® Forensic DNA Extraction Kit, and QIAamp® DNA Investigator Kit. The organic extraction required an overnight digestion and was purified using the QIAquick PCR Purification Kit. The results were divided into two categories: medicolegal (fresh samples or preserved tissue fragments with a maximum age of 5 years) and exhumation (skeletonized for 70 years) samples. As expected, the average DNA concentrations for medicolegal samples was much higher than from the aged bone samples. The DNA Investigator kit produced no results for aged samples and the organic extraction method produced the highest DNA concentrations for both sample types. However, the organic method required more bone powder than the other methods. For medicolegal cases, all methods produced at least a partial profile, but the DNA Investigator kit resulted in less alleles. For aged bone samples, PrepFiler® and organic extraction showed a similar number of alleles, whereas DNA Investigator produced

almost no alleles for all samples. Overall, the organic method and PrepFiler® were comparable, but the organic method is more time consuming and uses toxic materials. The DNA Investigator kit was considered unsuitable for aged bone samples in this study [79].

DNA extraction methods have been continuously optimized over the past 30 years. Extraction methods need constant improvement in order to be faster, cheaper, use less starting material, yield higher concentrations of DNA, reduce inhibitors, and become automatable. The most common DNA extraction methods include the traditional phenol/chloroform method and silica columns or silica-coated magnetic beads. The phenol/chloroform method is well established producing high DNA yields but does not effectively remove all PCR inhibitors. In fact, phenol is a PCR inhibitor, and a poorly executed extraction may cause inhibition. Silica columns such as those from the QIAamp® DNA Investigator kit have demonstrated clean extractions with little evidence of PCR inhibition; however, DNA yield is consistently lower than other methods tested. DNA IQ™ and PrepFiler® systems use silica-coated paramagnetic beads for extraction purification. PrepFiler® uses small amounts of starting material, has demonstrated effectiveness in removing a variety of PCR inhibitors, and produces DNA yields comparable to organic extraction. The DNA IQ™ method performs a simultaneous extraction and purification making it a popular method for forensic samples, however, it may not be the most effective kit in removing PCR inhibitors. Overall, all methods discussed previously have some advantages and disadvantages, but many circumstances must be evaluated in order to choose an optimal method for hard tissue extraction. However, it has been continuously demonstrated that total demineralization protocols are the most effective methods at breaking down organic bone material to release DNA for extraction.

DNA Markers for Human Identification

The primary goal of routine forensic DNA analysis is to obtain a DNA profile from a biological sample recovered from a crime scene or set of human remains, which is compared to reference profiles to determine a "match" [80].

Short Tandem Repeats (STRs)

Short tandem repeats (STRs) or microsatellites are the most common genetic marker used in DNA analysis for the identification of human remains [80,81]. STRs are short sequences of DNA consisting of 2-6 repeating nucleotide units [82,83]. They are used because of their relatively small amplicon size (75 – 450 bp), highly polymorphic nature, and high power of discrimination (PD) [84-88]. STRs are amplified using PCR and can therefore be retrieved from very small amounts of biological material (0.1 ng), and due to their small size, highly fragmented DNA is more likely to be amplified. STRs can be multiplexed allowing for more genetic markers to be analyzed simultaneously, and currently, the most common method used to detect STR markers is by size separation using capillary electrophoresis (CE) [87,89]. The amplicons are labeled with fluorescent dyes and separated by length. While CE is relatively simple and cost-effective with existing commercial STR kits, the complete sequence of each amplicon is not determined, only the differences in length [81]. STR typing is limited by the number of markers which can be multiplexed, separated, and detected using CE-based methods due to restraints in the number of dye channels and the space in each channel (<500 bp) [81]. The current megaplex STR kits (GlobalFiler® (Thermo Fisher Scientific), PowerPlex® Fusion System (Promega), Investigator® 24plex QS (QIAGEN)) contain 21-24 loci [24,90-92], including the 20 core CODIS loci and amelogenin. The more STR markers included in a multiplex

kit, the more discriminatory the system is, and therefore a better chance of identification with fewer chances of adventitious matches [93]. However, no more than 25-30 STR loci can be typed simultaneously because of the current spectral capabilities of the genetic analyzers used for CE [86,87]. Therefore, in order to increase the discriminatory power of STR systems for the identification of highly challenging samples, other types of DNA markers may be interrogated using alternate technologies such as MPS.

Single Nucleotide Polymorphisms (SNPs)

Single nucleotide polymorphisms (SNPs) are also genetic markers that can be used for human identification purposes. As the name suggests, a SNP is a single base change within a DNA sequence [85]. Therefore, it is possible to amplify small targets (~60 bp) making SNPs ideal for use with highly degraded samples [18]. SNPs also have a lower mutation rate (10^{-8}) [95, 96] than STRs (10^{-3} to 10^{-5}) [97] making them genetically useful markers [98]. Conversely, there are also some disadvantages to using SNPs. They are largely bi-allelic markers, making them relatively uninformative by individual locus. However, they do become informative when a panel of 50-100 SNPs is used for identification, producing a PD (10^{-16}) similar to the original 13 core CODIS loci [99]. Due to the adoption of 20 core CODIS loci, HID SNP panels may need to be notably larger to match the power of discrimination of the latest STR kits. The bi-allelic nature of SNPs also makes them poor candidates for mixture deconvolution [100], but in combination with STRs, they may be more informative. Ideally, SNPs used for identification would have high heterozygosity and low Fst values, which increases the SNP panel's efficiency, meaning less SNPs for higher PDs or lower match probabilities. Kidd et al. [100] identified universal SNPs that show minimal allele frequency variation among populations and are

highly informative for human identification. However, SNPs can also be informative for other forensically relevant purposes.

SNPs can be classified into several categories according to their use in forensic identification. Identity informative SNPs (iiSNPs) are used for human identification, ancestry informative SNPs (aiSNPs) for identifying biogeographical ancestry (BGA), lineage informative SNPs (liSNPs) may be used to deduce genealogies and family pedigrees, and phenotype informative SNPs (piSNPs) are used for predicting externally visible characteristics (EVCs) such as hair, skin, and eye color [101]. Although the primary question to be answered in missing persons cases is individualization, the additional information that can be provided by SNP analysis may be useful for forensic intelligence purposes when an identification is not possible. Currently, the most commonly used method for SNP genotyping is single base extension (SBE), which incorporates fluorescently labeled dNTPs one at a time to the 3' end of a primer directly adjacent to the SNP. SBE is the method used in a commercial SNP-typing kit (SNaPshot®) and relies on CE for detection [102]. The SNaPshot assay allows multiplexing between 30 – 40 SNPs with sensitivity down to 31 pg of DNA. In addition, the more SNP assays that are performed to genotype enough SNPs, the more DNA is consumed, depleting samples very quickly [103].

Well defined assays for the prediction of EVCs have been reported in literature and have been used for forensic intelligence purposes. IrisPlex and HIrisPlex developed by Walsh et al. are SNP panels that are highly predictive of eye and hair color. IrisPlex was created using six of the most informative SNPs for blue and brown eye color [104-106] with an accuracy prediction of >90% [107]. Walsh et al. [108] further expanded IrisPlex

into HIrisPlex for the prediction of hair and eye color simultaneously. HIrisPlex includes 24 hair and eye color predictive variants (23 SNPs and one Indel) including the six original SNPs used in IrisPlex [108]. The development of these assays may be helpful in determining EVCs in the deceased from MFIs and missing persons' investigations [109]. BGA is another area of focus for continual improvement by the forensic community using SNP markers. The first ancestry informative marker (AIM) assay for use in forensics was developed in 2003 and contained 178 SNPs. Because this was prior to MPS, the 178 SNPs were still being sequenced using the SNPstream system (single base extension chemistry) with several multiplexes [110]. Two more AIM SNP panels were then developed, a 34-plex SNP assay [111, 112] and a 47 SNP marker panel [113]. Since the implementation of MPS in forensic research, many BGA assays have been developed including the Precision ID Ancestry panel from Thermo Fisher Scientific [114], Primer Panel B from Verogen used with the ForenSeq™ DNA Signature Prep kit [115], and many other custom assays for various regions worldwide [116-119].

Microhaplotypes

Microhaplotypes are one of the newest marker types being investigated in the forensic field for the use in human identification, ancestry, kinship, and mixture deconvolution [124, 125]. They are essentially two or more SNP markers located within 200 bps of each other [98]. Kidd et al [120] proposed criteria for a viable microhaplotype locus including at least three haplotypes (alleles) within a 200 bp non-recombinant hot spot region. Like individual SNPs, microhaps are small enough to be used for degraded DNA, but they also provide a better potential for mixture deconvolution because of their multiallelic nature [24]. When closely linked bi-allelic and tri-allelic SNPs are genotyped

together they become multiallelic haplotypes [18]. Many different haplotypes exist because of rare recombinations and demographic dispersion (migration, isolation, admixture, random genetic drift, and/or selection), which may be used to determine BGA of an individual [120, 121]. There are three potentially forensically relevant markers that consist of closely linked SNPs including haploblocks [122], mini-haplotypes [123], and microhaplotypes (microhaps) [98, 120].

Mitochondrial DNA (mtDNA)

When DNA typing methods for nuclear DNA fail due to extensive DNA damage and degradation, mitochondrial DNA (mtDNA) is often targeted for HID [16, 102, 126, 127]. mtDNA is a circular, double-stranded genome located within the mitochondria organelle. Cells contain hundreds to thousands of copies of mtDNA, compared to two copies of nuclear DNA per cell [16, 126, 127]. Because of the high abundance, it is more likely that mtDNA will be recovered from very old, highly compromised samples (such as skeletal remains) when very little or no nuclear DNA is present. Therefore, mtDNA analysis is important for missing persons' and mass disaster investigations [16, 126]. Mitochondrial DNA is inherited maternally as a haplotype-block (non-recombinant) and therefore has a low PD for individualization purposes [126, 127]. However, it has a very high mutation rate, about ten times higher than STRs, so high sequence variation may be an advantage for forensic DNA typing [128]. The non-coding or hypervariable regions (HVI and HVII) have traditionally been sequenced using CE-based methods to link relatives or trace lineage within the same maternal line [126]. The utility of mtDNA for kinship analysis may be more useful than STRs in certain situations, particularly when

relying on distant relatives as references. Because mtDNA is maternally inherited, any person with the same maternal line can provide a reference sample for comparison.

Even though the hypervariable regions of mtDNA show the most variation in sequence between individuals, complete mitochondrial genome (mtGenome) sequencing can be very useful. The coding region of mtDNA has a mutation rate 10X less than the non-coding control regions (or hypervariable regions) that are used in forensic analysis. However, there are several positions in the mtGenome that act as mutational hot spots and some sites that are more prone to mutations than others. It has also been determined that some haplogroups cannot be completely defined based on control region data, so whole mitochondrial sequencing is relevant and may be more informative [129]. Haplogroup identification can be useful in MFIs and the identification of human remains because they can provide a general origin of a person. It is also easier to sequence the whole mitochondrial genome than it is to sequence the entire human genome because the mtGenome is only 16,569 bp in length. Sequencing the whole mtGenome may increase powers of discrimination for lineage markers [130, 131] as ~75% of variation in the mtGenome is in the coding region [130]. Whole mtGenome sequencing can be accomplished by Sanger sequencing, but now is more commonly sequencing by MPS. Thermo Fisher Scientific has developed a sequencing panel for whole mtGenome sequencing consisting of 81 primer pairs of ≤175 bp, which is beneficial for degraded DNA such as skeletal remains [132-135].

Sequencing Technology

Watson and Crick identified the basic structure of DNA in 1953, but the capability to sequence or “read” DNA at the nucleotide level was not developed until about 15 years

later by Robert Holley [136]. One of the first sequencing methods used was created in the 1970s by Alan Coulson and Frederick Sanger called the plus and minus system. The ‘plus and minus method’ was used by Sanger to sequence the first genome, a bacteriophage [137]. Towards the end of the 1970s, Frederick Sanger developed a new method that would be used for sequencing for the next 25 years called Sanger or ‘chain-terminating’ sequencing. Sanger sequencing was easier and more efficient than the ‘plus and minus method’ and used radioactive or fluorescently labeled dideoxynucleotides (dNTPs) to terminate the nucleotide chain [138]. Ultimately, newer methods of sequencing such as pyrosequencing and eventually the use of paramagnetic beads, emulsion PCR, and microprocessor chips began to emerge [136].

Since the introduction of Sanger sequencing in 1977, much progress has been made in the fields of molecular biology and genetics. The Human Genome Project and many other species genome projects have since been completed [136]. Sanger sequencing has been used in forensics for about 45 years. However, Sanger sequencing has several disadvantages including low throughput, high cost, labor intensive, and being technically difficult [139-141]. Therefore, the forensic community has shifted its focus to a newer form of DNA sequencing that provides more potential to increase multiplex size, deconvolute mixtures, and obtain more information from each sample.

Massively Parallel Sequencing (MPS)

Massively parallel sequencing (MPS), also referred to as next generation sequencing (NGS), or second-generation sequencing (SGS) has offered an alternative method to Sanger sequencing for DNA typing. The forensic DNA community has embraced this newer technology to sequence the entire mitochondrial genome, investigate

larger and more discriminatory SNP panels for HID, ancestry, and phenotypic information [112, 116, 134, 142]. The more recent development of MPS has transformed genomic analyses, allowing high-throughput sequencing to generate more genetic information from each sample while reducing the cost, time, and risk of contamination compared to previous sequencing methods [85, 87, 143-146].

MPS provides technical improvements over previously used sequencing technologies such as no longer requiring bacterial cloning of DNA fragments or electrophoresis and sequencing multiple reactions simultaneously [90]. In addition, MPS allows scientists to obtain larger amounts of data from each sample including, but not limited to, sequence variations or SNPs within STRs, degree of accuracy, read length, strand bias, and coverage [85, 87, 103, 143, 144]. MPS can provide more comprehensive sequence information about conventional STR markers, can sequence the entire mitochondrial genome, and even combine markers that are not routinely typed with current CE-based methods [142].

When deep sequencing first began there were three platforms released for use by laboratories including the Genome Sequencer from 454 Life Sciences (later Roche) in 2005, followed by the Genome Analyzer by Solexa (later developed by Illumina) in 2006, and finally the SOLiD system by Applied Biosystems (now Thermo Fisher Scientific) in 2007. Illumina sequencing technology involves adapter-ligated DNA being amplified on a flow cell covered with oligonucleotides that are complementary to the adaptors. The DNA fragments are hybridized to the flow cell and then amplified by bridge amplification. Following amplification, the bridges are made linear to form clusters (~1000 copies for each cluster) of the original DNA molecule. However, 454 and SOLiD sequencing

technology follow a different approach. Instead of being hybridized to a flow cell, adapter-ligated DNA is hybridized to complementary oligonucleotide covered beads for amplification in emulsion PCR. After emulsion PCR, thousands to millions of copies of the original DNA molecule are now coating the beads ready for sequencing [145]. Illumina and 454 technologies use sequencing-by-synthesis, whereas SOLiD uses sequencing-by-ligation. Sequencing-by-synthesis (Illumina) uses a DNA polymerase to extend a sequencing primer by incorporating nucleotides that produce a sequence complementary to the DNA molecule. Fluorescent reversible dye terminators are used to incorporate one base per cycle to the DNA template. After the dye terminators are incorporated into the template DNA, the fluorescent terminators are removed prior to the next cycle [147]. For sequencing-by-synthesis using 454, the principle of pyrosequencing technology is used. Non-terminating deoxynucleotides are added in sequential order to the DNA template and a pyrophosphate is released and converted into ATP. The ATP is used as a light substitution and captured by a charge-coupled device (CCD) camera [148]. Conversely, sequencing-by-ligation takes a slightly different approach. SOLiD technology uses a DNA ligase to add a fluorescently labeled eight-base probe to an oligonucleotide chain, five bases are template specific and three bases are universal for hybridization to the template. The three universal bases and the fluorophore are cleaved off and a new set of probes take their place. After one round of sequencing is completed, the new DNA strand is “melted” off and a new primer is added. This process is repeated multiple times to incorporate different primers so that each base is sequenced twice and the colors become nucleotides [149]. These systems have the capability of real high throughput sequencing. To lower costs and throughput to achieve large samples sizes (but not requiring the whole genome; i.e. targeted

and small amplicon sequencing), three ‘personal sequencers’ were developed. In 2010, Roche 454 unveiled the Genome Sequencer Junior, shortly after in 2011, Illumina released the MiSeq, and Life Technologies acquired Ion Torrent, using a similar sequencing technology to 454 [145].

Despite the many advantages of MPS, there are also some disadvantages. MPS techniques are not routinely used in forensic labs, nor is the instrumentation currently available to many forensic casework laboratories. These platforms have yet to be validated for forensic casework, and implementation would require a substantial investment financially, and also in terms of time, analyst training, and validation. NGS technologies have the capability of increasing sample throughput, quickly generating high quality sequences [153], and improving overall efficiency of DNA sequencing, while reducing the cost and time it takes to obtain genetic information [143, 154]. However, to date MPS is still too expensive and time consuming for routine casework, but rather used for niche applications like missing persons, war remains, forensic intelligence, ancestry, mitochondrial analysis, and EVCs, etc.

To date, there are four MPS platforms available for forensic applications: the Ion Torrent Personal Genome Machine (PGM™), its successor the Ion S5™ System by Thermo Fisher Scientific, the Illumina MiSeq FGx™ System, and the QIAGEN GeneReader platform.

Ion Torrent™ Systems

The PGM was introduced in 2012, while the Ion S5™ was introduced in 2015 and both are MPS platforms that use semiconductor-sequencing technology. These instruments are high-throughput DNA sequencers that release hydrogen ions, which changes the pH of

the solution and is detected when a nucleotide is integrated into the DNA strand. When the pH changes, the chemical information produced is translated into digital information or data (Fig. 1.1) [85, 144]. The PGM™ and S5™ use massively parallel sequencing chemistry with respectable accuracy and user-friendly instrumentation [144]. These platforms do not require the use of a camera to detect light or fluorescence, which makes them a less expensive option for sequencing. Other advantages include up to a 400 bp read length (required for STR typing), relatively short sequencing time (~8 hours), cheaper sequencing costs compared to traditional Sanger sequencing, and the ability to use three different chip sizes depending on how many samples are being sequenced and the desired coverage per sample. These sequencers provide vast amounts of data regarding coverage, reads, and information from each sample. The Ion Torrent platforms are compatible with sample-tagging barcodes, which allow many samples to be personalized with specific identification tags and then sequenced in the same run [144]. Because MPS does not depend on size to complete sequencing, it may be able to overcome some of the issues with CE-based methods. These issues include the limited number of markers that can be multiplexed at one time and the number of non-overlapping fluorophores to determine numerous amplicons in one run [103, 150, 151].

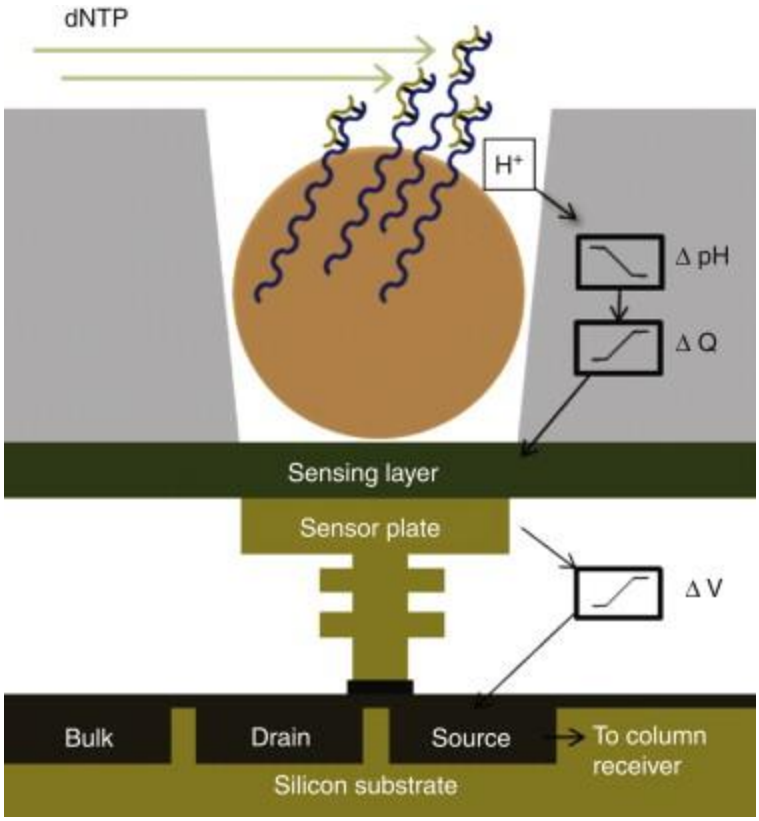


Figure 1.1 The process by which chemical information is translated to digital information via MPS using the Ion Torrent™ systems. Life Technologies http://www3.appliedbiosystems.com/cms/groups/applied_markets_marketing/documents/generaldocuments/cms_094273.pdf [191]

Illumina MiSeq FGx™

The Illumina MiSeq FGx™ System is another sequencing platform currently on the market and is specifically designed for forensic purposes. The MiSeq uses reversible dye terminator dNTPs for sequencing. When DNA is passed over a flow-cell of complementary oligonucleotides, a solid phase PCR produces clusters of cloned oligonucleotides. This process of creating clusters is called bridge amplification because the DNA replicates the strands forming an arch, or bridge, between each other. After bridge amplification reversible dye terminator dNTPs incorporate a fluorophore on the 3' hydroxyl, it is cleaved so that polymerization can continue, and sequencing can start. The nucleotides being

incorporated are monitored by a CCD to ensure proper placement of the nucleotides (Figure 2) [136, 139].

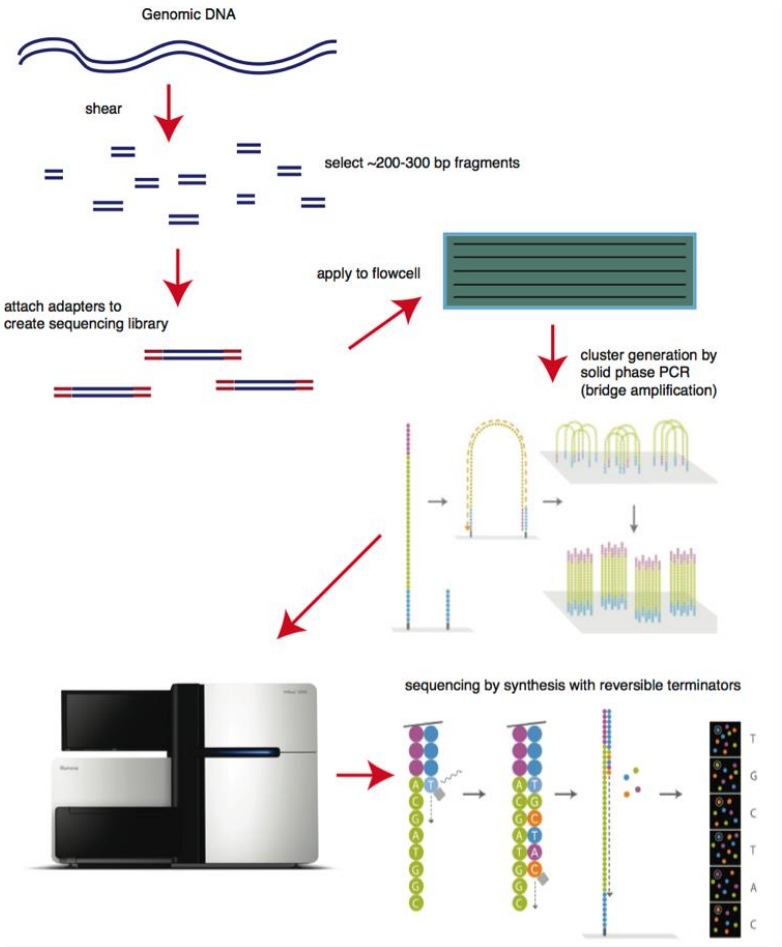


Figure 1.2 Illumina's bridge amplification and reversible dye terminator sequencing using the MiSeq FGx™. BiteSize Bio https://bitesizebio.com/13546/sequencing-by-synthesis-explaining-the-illumina-sequencing-technology/ [18]

QIAGEN GeneReader

The GeneReader NGS platform is the newest MPS platform on the market for forensic applications launched by QIAGEN in 2015. However, the GeneReader is quite different from the other platforms on the market because it is an all-in-one workflow, beginning at sample preparation and ending at data analysis. The GeneReader platform works in tandem with the QIAcube extraction robot for sample preparation. This platform

also sequences with sequencing-by-synthesis technology, like Illumina, but with minor modifications. Instead of incorporating a fluorescent dNTP for each template, the GeneReader incorporates only enough dNTPs to make an identification [152].

Massively Parallel Sequencing Forensic Panels and Chemistries

There are several sequencing panels available for use in forensic casework such as DVI and missing persons. Many SNP multiplexes have been developed that include at least 140 SNPs for forensic and human identification [142, 155, 156]. Additionally, the ForenSeq™ DNA Signature Prep Kit includes two separate primer panels with both STRs and SNPs. Primer Mix A is simply the identity panel consisting of 27 autosomal STRs, 24 Y-STRs, 7 X-STRs, and 94 iiSNPs. However, Primer Mix B consists of all markers in Primer Mix A with the addition of 22 piSNPs and 56 BGA SNPs [157, 158]. Furthermore, Thermo Fisher has a Precision ID Identity panel made up of 90 autosomal SNPs and 34 upper Y-Clade SNPs. In addition, they also provide the Precision ID GlobalFiler® NGS STR Panel, which consists of 32 STR markers, 1 Y-indel, and two amelogenin sex markers. Most recently, QIAGEN released a MPs-Plex SNP panel designed in collaboration with ICMP to specifically identify missing persons, and it is comprised of over 1400 identification SNPs [12], the largest SNP panel yet.

In the case of degraded samples such as forensic casework, mass disasters, and missing persons' cases, MPS may be able to provide more probative information (SNPs, mtDNA, STRs, microhaps) for these challenging samples than CE-based STR typing can provide. However, MPS within forensic biology is still in its infancy.

MPS Data Analysis Pipelines

Due to the large amount of data generated using MPS-based sequencing, the analysis pipeline is an extremely important part of the process and is also the main area of concern within the forensic community. After DNA samples have been sequenced the data is retrieved in a way specific to each platform. On the Ion S5™ or semi-conductor sequencing instruments, the raw data from the samples will automatically go through signal processing and base calling on the Torrent Suite Server (TSS). Once this is completed, the data can be visualized on the server and various plugins can be run. Plugins are a type of data tool that analyzes and presents the data in different formats such as CSV files, FASTQ and FASTA files, PDFs, BEF files, and diagrams. CSV files can be downloaded and the data can be manipulated and viewed in Excel. Run reports show data and quality metrics and can be downloaded as PDFs. BEF files may be obtained and uploaded into a data analysis software tool called Converge™ (Thermo Fisher Scientific).

Converge™ is a data analysis software tool created by Thermo Fisher Scientific for the analysis of STRs, SNPs, and microhaplotypes. Converge™ integrates DNA data management and analysis and can be used for casework, research, kinship, and paternity. It has many valuable features including the ability to upload external information (images, PDF documents, and CE profiles), merging profiles and creating a consensus profile (CE and MPS data from multiple kits), and family tree construction within kinship analysis. Converge™ allows the user to add all forensic STR kits, as well as determine stutter and balance thresholds for individual loci. Samples can be associated with a particular case and all case information, including attachments, are accessible when a case is selected.

However, there are many other third-party data analysis tools that may also be used to analyze MPS data.

Integrative Genomics Viewer (IGV) is an online tool/application that allows the interaction and visualization of many types of data including MPS and array-based data generated from any MPS instrument [159]. This tool has many useful features in order to view sequencing data easily [160]. The user has access to many features and is able to look at the data on a large scale (whole genome) or a small scale (base pairs). Sequencing instruments will provide the user with Binary Alignment MAP (BAM) and Binary Alignment Index (BAI) files, and BED files among others. These files hold all of the DNA sequences for each sample that were sequenced on the instrumentation. For forensic MPS purposes, IGV uses BAM and BAI indexed files as its input so the user can see subsets of data. If multiple file types are uploaded simultaneously, each file will appear in a separate panel and can be merged if the user desires.

The Illumina ForenSeq™ Universal Analysis Software (UAS) is a tool designed specifically for the MiSeq FGx™ Forensic Genomics System for use with the ForenSeq™ DNA Signature Prep Kit. This software analyzes sequenced DNA samples for human identification and can perform run setup, sample management, analysis, and report generation. The ForenSeq™ UAS provides population statistics and automated sample comparison, as well as an estimation of BGA, and EVCs from various population datasets [161].

The STR allele identification tool – Razor (STRait Razor) is a bioinformatics tool that uses Perl script on a Linux/Unix based system and created to effectively detect STRs from massively parallel sequencing platforms [162, 163]. However, it can now be used on

a Windows platform as of the latest update [164]. STRait Razor uses a FASTQ file from the sequencing instrument to analyze complex and simple STR repeat motifs [162]. The program consists of two components: the coding script and a complementary Excel workbook. The script recognizes all haplotypes and the Excel workbook collects the haplotype information and formats it so the user can easily interpret the data [164]. Alleles are detected by matching the flanking regions around the DNA sequence, which allow for all extraneous nucleotides to be removed from the sequence leaving only the repeats. The final step includes filtering the reads that are not STR sequences. Alleles are called by comparing the repeat region to known alleles of the same length [163], similar to how CE compares peaks to the allelic ladder. STRait Razor also highlights variations in DNA sequences [162, 163]. Although STRait Razor is primarily used for STRs, SNPs and insertion-deletions (InDels) can now be detected with the latest version [164]. In addition, this software has been used in the forensics field for many applications including MPS multiplexes [87, 88, 165], sequence variation and length-based analysis of population data [166, 167], and the characterization of InDels [168].

With MPS technology continually evolving, easier and faster data analysis capabilities are necessary for the massive amounts of data that MPS produces. Many data analysis software tools are being used and created for massively parallel sequencing including mitoSAVE [169] and AFDIL-QIAGEN mtDNA Expert (AQME) [170] for mtDNA, the Allele Frequency Database (ALFRED) [171, 172] for SNPs, and SEQ Mapper [173] and My-Forensic-Loci-queries (MyFLq) [174] for STRs.

Purpose of this Study

The DNA obtained from human remains recovered from mass disaster and missing persons' investigations range from pristine to highly damaged, degraded, and/or inhibited. These insults may have arisen due to bodies being buried, severely burnt, or exposed to harsh environmental conditions. The objective of this study was to assess the tolerance of STR and MPS chemistries designed for HID purposes to common inhibitors and their relative performance with these types of challenging samples. This investigation is important because MPS based chemistries have yet to be fully tested with inhibited samples in order to define the tolerance of these systems. Additionally, very few studies have evaluated the current MPS sequencing chemistries and primer panels with challenging remains and other difficult, forensically relevant samples. If MPS methods are being considered as the future of forensic DNA analysis, it is necessary that these sequencing chemistries be extensively tested before implementation into crime laboratories.

Inhibitor studies are important for forensic casework because they inform the forensic DNA community on how to maximize the removal of PCR inhibitors from various types of challenging samples in order to generate more complete DNA profiles and more probative information. This study was divided into four phases. The first and second phases were to test the tolerance of conventional STR typing chemistry and MPS chemistry to various concentrations of forensically relevant PCR inhibitors (hematin, collagen, calcium, humic acid, and melanin) that may be found in routine forensic casework, as well as missing persons' and mass disaster cases (skeletal samples in particular). Various concentrations of PCR inhibitors were added to DNA extracts at moderate (1 ng) and low amounts (0.1 ng) of DNA template. The samples were STR-typed using the GlobalFiler®

PCR Amplification and Investigator® 24plex QS kits, and sequenced on the Ion PGM™ using the HID-Ion AmpliSeq™ Library kit and the MiSeq FGx™ using the ForenSeq™ DNA Signature Prep Kit (Primer Mix A) to determine the baseline tolerance of each system to inhibited samples. The third phase had multiple goals including 1) to evaluate the efficiency of various DNA extraction methods to remove high amounts of PCR inhibitors from challenging samples prior to MPS, and 2) compare the quality of STR/SNP analysis using an early access panel for degraded samples with Precision ID chemistry on the Ion S5™, the ForenSeq™ DNA Signature Prep Kit (using Primer Mix A) on the MiSeq FGx™, and traditional CE-based STR typing for the identification of human remains. Blood, hair, and bone samples were collected and spiked with high amounts of the relevant inhibitor (humic acid, melanin, hematin, or calcium). Blood and hair samples were extracted using the three most commonly used commercial kits in forensic laboratories: PrepFiler® BTA (Life Technologies™, Carlsbad, CA), DNA IQ™ (Promega, Madison, WI), and QIAamp® DNA Investigator (QIAGEN, Hilden, Germany), and an organic method. Bone samples were extracted using the same three commercial extraction kits and two different total demineralization protocols. Finally, the fourth phase of this project was to evaluate two MPS chemistries and platforms using environmentally challenged human remains such as may be encountered in missing persons' cases. Bone and tooth samples (exposed to cremation, embalming, thermal degradation, fire, and decomposition) were extracted using a total demineralization protocol, quantified, STR-typed via CE, and then sequenced using both a custom AmpliSeq™ STR and iiSNP panel designed for degraded samples with Precision ID chemistry on the Ion S5™ System and the ForenSeq™ DNA Signature Prep Kit (using Primer Mix A) on the MiSeq FGx™.

References

1. Kovras I, Robins S (2016) Death as the border: Managing missing migrants and unidentified bodies at the EU's Mediterranean frontier. Polit Geogr 55:40-49. doi:10.1016/j.polgeo.2016.05.003
2. Salado Puerto M, Tuller H (2017) Large-scale forensic investigations into the missing: Challenges and considerations. Forensic Sci Int 279:219-228. doi:10.1016/j.forsciint.2017.08.025
3. Alvarez-Cubero MJ, Saiz M, Martinez-Gonzalez LJ, Alvarez JC, Eisenberg AJ, Budowle B, Lorente JA (2012) Genetic identification of missing persons: DNA analysis of human remains and compromised samples. Pathobiology 79 (5):228-238. doi:10.1159/000334982
4. (FBI) FBI Combined DNA Index System (CODIS). FBI, U.S. government, U.S. Department of Justice https://www.fbi.gov/services/laboratory/biometric-analysis/codis.
5. (FBI) FBI CODIS and NDIS Fact Sheet - FBI. FBI, U.S. government, U.S. Department of Justice https://www.fbi.gov/services/laboratory/biometric-analysis/codis/codis-and-ndis-fact-sheet.
6. NIJ (2018) NamUs 2.0. UNTHSC, National Institute of Justice, U.S. Department of Justice https://www.namus.gov/.
7. (UNTCHI) UNTCHI UNT Center for Human Identification UNT Center for Human Identification http://untchi.org/.
8. Vallee J (2018) Department of Defense DNA Operations. Armed Forces Medical Examiner System, Dover AFB, DE

9. Houck MM (2013) The National Missing and Unidentified Persons System (NamUs). In: Enc Forensic Sci. Academic Press, Waltham, p 375. doi:10.1016/b978-0-12-382165-2.00057-x

10. Amory S, Huel R, Bilic A, Loreille O, Parsons TJ (2012) Automatable full demineralization DNA extraction procedure from degraded skeletal remains. Forensic Sci Int Genet 6 (3):398-406. doi:10.1016/j.fsigen.2011.08.004

11. ICMP (2018) International Commission on Missing Persons (ICMP). ICMP. https://www.icmp.int/about-us/history/.

12. ICMP (2017) QIAGEN and ICMP launch next-generation DNA testing solutions for identifying missing persons ICMP,

13. Alonso A, Martin P, Albarran C, Garcia P, Fernandez de Simon L, Jesus Iturralde M, Fernandez-Rodriguez A, Atienza I, Capilla J, Garcia-Hirschfeld J, Martinez P, Vallejo G, Garcia O, Garcia E, Real P, Alvarez D, Leon A, Sancho M (2005) Challenges of DNA profiling in mass disaster investigations. Croat Med J 46 (4):540-548

14. Goodwin WH, Simmons T (2013) Disaster Victim Identification. In: Enc Forensic Sci. Academic Press, Waltham, pp 332-338. doi:10.1016/b978-0-12-382165-2.00061-1

15. Brough AL, Morgan B, Rutty GN (2015) The basics of disaster victim identification. Journal of Forensic Radiology and Imaging 3 (1):29-37. doi:10.1016/j.jofri.2015.01.002

16. Budowle B, Bieber FR, Eisenberg AJ (2005) Forensic aspects of mass disasters: strategic considerations for DNA-based human identification. Leg Med (Tokyo) 7 (4):230-243. doi:10.1016/j.legalmed.2005.01.001

17. Prinz M, Carracedo A, Mayr WR, Morling N, Parsons TJ, Sajantila A, Scheithauer R, Schmitter H, Schneider PM, International Society for Forensic G (2007) DNA Commission of the International Society for Forensic Genetics (ISFG): recommendations regarding the role of forensic genetics for disaster victim identification (DVI). Forensic Sci Int Genet 1 (1):3-12. doi:10.1016/j.fsigen.2006.10.003

18. Zietkiewicz E, Witt M, Daca P, Zebracka-Gala J, Goniewicz M, Jarzab B, Witt M (2012) Current genetic methodologies in the identification of disaster victims and in forensic analysis. J Appl Genet 53 (1):41-60. doi:10.1007/s13353-011-0068-7

19. Peschel O, Eisenmenger W (2013) Airplane Crashes and Other Mass Disasters. In: Enc Forensic Sci. Academic Press, Waltham, pp 188-192. doi:10.1016/b978-0-12-382165-2.00182-3

20. Karija Vlahovic M, Kubat M (2012) DNA extraction method from bones using Maxwell(R) 16. Leg Med (Tokyo) 14 (5):272-275. doi:10.1016/j.legalmed.2012.04.004

21. Andelinovic S, Sutlovic D, Erceg Ivkosic I, Skaro V, Ivkosic A, Paic F, Rezic B, Definis-Gojanovic M, Primorac D (2005) Twelve-year experience in identification of skeletal remains from mass graves. Croat Med J 46 (4):530-539

22. INTERPOL (2009) Disaster Victim Identification Guide. INTERPOL,

23. Noor MSM, Khoo LS, Alias WZZ, Hasmi AH, Ibrahim MA, Mahmood MS (2017) The clandestine multiple graves in Malaysia: The first mass identification operation of human skeletal remains. Forensic Sci Int 278:410. e411-410. e419. doi:PMID 410.e1 10.1016/j.forsciint.2017.05.014

24. Watherston J, McNevin D, Gahan ME, Bruce D, Ward J (2018) Current and emerging tools for the recovery of genetic information from post mortem samples: New directions for disaster victim identification. Forensic Sci Int Genet 37:270-282. doi:10.1016/j.fsigen.2018.08.016

25. Montelius K, Lindblom B (2012) DNA analysis in Disaster Victim Identification. Forensic Sci Med Pathol 8 (2):140-147. doi:10.1007/s12024-011-9276-z

26. Lee HY, Park MJ, Kim NY, Sim JE, Yang WI, Shin KJ (2010) Simple and highly effective DNA extraction methods from old skeletal remains using silica columns. Forensic Sci Int Genet 4 (5):275-280. doi:10.1016/j.fsigen.2009.10.014

27. Kitayama T, Ogawa Y, Fujii K, Nakahara H, Mizuno N, Sekiguchi K, Kasai K, Yurino N, Yokoi T, Fukuma Y, Yamamoto K, Oki T, Asamura H, Fukushima H (2010) Evaluation of a new experimental kit for the extraction of DNA from bones and teeth using a non-powder method. Leg Med (Tokyo) 12 (2):84-89. doi:10.1016/j.legalmed.2009.12.004

28. Mundorff AZ, Bartelink EJ, Mar-Cash E (2009) DNA preservation in skeletal elements from the World Trade Center disaster: recommendations for mass fatality management. J Forensic Sci 54 (4):739-745. doi:10.1111/j.1556-4029.2009.01045.x

29. Mundorff A, Davoren JM (2014) Examination of DNA yield rates for different skeletal elements at increasing post mortem intervals. Forensic Sci Int Genet 8 (1):55-63. doi:10.1016/j.fsigen.2013.08.001

30. Ferreira STG, Garrido RG, Paula KA, Nogueira RC, Oliveira ES, Moraes AV (2013) Cartilage and phalanges from hallux: alternative sources of samples for DNA typing in

disaster victim identification (DVI). A comparative study. Forensic Sci Int Genet Suppl Ser 4 (1):e366-e367

31. Calacal GC, Apaga DLT, Salvador JM, Jimenez JAD, Lagat LJ, Villacorta RPF, Lim MCF, Fortun RDR, Datar FA, De Ungria MCA (2015) Comparing different post-mortem human samples as DNA sources for downstream genotyping and identification. Forensic Sci Int Genet 19:212-220. doi:10.1016/j.fsigen.2015.07.017
32. Allouche M, Hamdoum M, Mangin P, Castella V (2008) Genetic identification of decomposed cadavers using nails as DNA source. Forensic Sci Int Genet 3 (1):46-49. doi:10.1016/j.fsigen.2008.07.008
33. Habib M, Pierre-Noel A, Fogt F, Budimlija Z, Prinz M (2017) Direct amplification of biological evidence and DVI samples using the Qiagen Investigator 24plex GO! Kit. Forensic Sci Int Genet Suppl Ser 6:E208-E210. doi:10.1016/j.fsigss.2017.09.079
34. Pfeiffer H, Hühne J, Ortmann C, Waterkamp K, Brinkmann B (1999) Mitochondrial DNA typing from human axillary, pubic and head hair shafts–success rates and sequence comparisons. Int J Leg Med 112 (5):287-290
35. Dhanardhono T, Wulandari N, Bhima SKL, Ahmad HJ, Widodo PT (2013) DNA profiling of disaster victim identification in Trenggalek shipwreck case. FForensic Sci Int Genet Suppl Ser 4 (1):e5-e6. doi:10.1016/j.fsigss.2013.10.002
36. Bonnet J, Colotte M, Coudy D, Couallier V, Portier J, Morin B, Tuffet S (2010) Chain and conformation stability of solid-state DNA: implications for room temperature storage. Nucleic Acids Res 38 (5):1531-1546. doi:10.1093/nar/gkp1060
37. Hofreiter M, Serre D, Poinar HN, Kuch M, Pääbo S (2001) ancient DNA. Nat Rev Genet 2 (5):353

38. Alaeddini R, Walsh SJ, Abbas A (2010) Forensic implications of genetic analyses from degraded DNA--a review. Forensic Sci Int Genet 4 (3):148-157. doi:10.1016/j.fsigen.2009.09.007

39. Sancar A, Lindsey-Boltz LA, Unsal-Kacmaz K, Linn S (2004) Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. Annu Rev Biochem 73 (1):39-85. doi:10.1146/annurev.biochem.73.011303.073723

40. Freese E, Cashel M (1964) Crosslinking of Deoxyribonucleic Acid by Exposure to Low Ph. Biochim Biophys Acta 91 (1):67-77

41. Goffin C, Verly WG (1983) Interstrand DNA crosslinks due to AP (apurinic/apyrimidinic) sites. FEBS letters 161 (1):140-144

42. Bertinchamps AJ, Hüttermann J, Köhnlein W, Teoule R (1978) Effects of ionizing radiation on DNA.

43. Hall A, Ballantyne J (2004) Characterization of UVC-induced DNA damage in bloodstains: forensic implications. Anal Bioanal Chem 380 (1):72-83. doi:10.1007/s00216-004-2681-3

44. Thompson RE, Duncan G, McCord BR (2014) An Investigation of PCR Inhibition Using Plexor®-Based Quantitative PCR and Short Tandem Repeat Amplification. J Forensic Sci 59 (6):1517-1529

45. Kasu M, Shires K (2015) The validation of forensic DNA extraction systems to utilize soil contaminated biological evidence. Leg Med (Tokyo) 17 (4):232-238. doi:10.1016/j.legalmed.2015.01.004

46. Hebda LM, Foran DR (2015) Assessing the Utility of Soil DNA Extraction Kits for Increasing DNA Yields and Eliminating PCR Inhibitors from Buried Skeletal Remains. J Forensic Sci 60 (5):1322-1330. doi:10.1111/1556-4029.12878

47. Kemp BM, Monroe C, Smith DG (2006) Repeat silica extraction: a simple technique for the removal of PCR inhibitors from DNA extracts. J Archaeo Sci 33 (12):1680-1689. doi:10.1016/j.jas.2006.02.015

48. Schmedes S, Marshall P, King JL, Budowle B (2013) Effective removal of co-purified inhibitors from extracted DNA samples using synchronous coefficient of drag alteration (SCODA) technology. Int J Leg Med 127 (4):749-755. doi:10.1007/s00414-012-0810-7

49. Faber KL, Person EC, Hudlow WR (2013) PCR inhibitor removal using the NucleoSpin(R) DNA Clean-Up XS kit. Forensic Sci Int Genet 7 (1):209-213. doi:10.1016/j.fsigen.2012.06.013

50. Hu Q, Liu Y, Yi S, Huang D (2015) A comparison of four methods for PCR inhibitor removal. Forensic Sci Int Genet 16:94-97. doi:10.1016/j.fsigen.2014.12.001

51. Alaeddini R (2012) Forensic implications of PCR inhibition--A review. Forensic Sci Int Genet 6 (3):297-305. doi:10.1016/j.fsigen.2011.08.006

52. Monroe C, Grier C, Kemp BM (2013) Evaluating the efficacy of various thermo-stable polymerases against co-extracted PCR inhibitors in ancient DNA samples. Forensic Sci Int 228 (1-3):142-153. doi:10.1016/j.forsciint.2013.02.029

53. Opel KL, Chung D, McCord BR (2010) A study of PCR inhibition mechanisms using real time PCR. J Forensic Sci 55 (1):25-33. doi:10.1111/j.1556-4029.2009.01245.x

54. Pionzio AM, McCord BR (2014) The effect of internal control sequence and length on the response to PCR inhibition in real-time PCR quantitation. Forensic Sci Int Genet 9:55-60. doi:10.1016/j.fsigen.2013.10.010

55. Besecker J, Cornell KA, Hampikian G (2013) Dynamic passivation with BSA overcomes LTCC mediated inhibition of PCR. Sensors and Actuators B-Chemical 176:118-123. doi:10.1016/j.snb.2012.09.110

56. Butler JM (2005) Constructing STR multiplex assays. In: Forensic DNA Typing Protocols. Springer, pp 53-65

57. Mulero JJ, Chang CW, Calandro LM, Green RL, Li Y, Johnson CL, Hennessy LK (2006) Development and validation of the AmpFℓSTR® Yfiler™ PCR amplification kit: a male specific, single amplification 17 Y-STR multiplex system. J Forensic Sci 51 (1):64-75

58. Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM, Smerick JB, Budowle B (2001) Validation of short tandem repeats (STRs) for forensic usage: performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples. J Forensic Sci 46 (3):647-660

59. Bessetti J (2007) An introduction to PCR inhibitors. J Microbiol Meth 28:159-167

60. Davoren J, Vanek D, Konjhodzic R, Crews J, Huffine E, Parsons TJ (2007) Highly effective DNA extraction method for nuclear short tandem repeat testing of skeletal remains from mass graves. Croat Med J 48 (4):478-485

61. Yang DY, Eng B, Waye JS, Dudar JC, Saunders SR (1998) Improved DNA extraction from ancient bones using silica-based spin columns. American Journal of Physical

Anthropology: The Official Publication of the American Association of Physical Anthropologists 105 (4):539-543

62. Huynen L, Lambert DM (2015) A Concentrated Hydrochloric Acid-based Method for Complete Recovery of DNA from Bone. J Forensic Sci 60 (6):1553-1557. doi:10.1111/1556-4029.12846
63. Marshall PL, Stoljarova M, Schmedes SE, King JL, Budowle B (2014) A high volume extraction and purification method for recovering DNA from human bone. Forensic Sci Int Genet 12:155-160. doi:10.1016/j.fsigen.2014.06.011
64. Ozdemir-Kaynak E, Yesil-Celiktas O (2015) Microwave-assisted digestion combined with silica-based spin column for DNA isolation from human bones. Anal Biochem 486:44-50. doi:10.1016/j.ab.2015.06.034
65. Greenspoon SA, Scarpetta MA, Drayton ML, Turek SA (1998) QIAamp spin columns as a method of DNA isolation for forensic casework. J Forensic Sci 43 (5):1024-1030
66. Phillips K, McCallum N, Welch L (2012) A comparison of methods for forensic DNA extraction: Chelex-100(R) and the QIAGEN DNA Investigator Kit (manual and automated). Forensic Sci Int Genet 6 (2):282-285. doi:10.1016/j.fsigen.2011.04.018
67. Ip SC, Lin SW, Lai KM (2015) An evaluation of the performance of five extraction methods: Chelex(R) 100, QIAamp(R) DNA Blood Mini Kit, QIAamp(R) DNA Investigator Kit, QIAsymphony(R) DNA Investigator(R) Kit and DNA IQ. Sci Justice 55 (3):200-208. doi:10.1016/j.scijus.2015.01.005
68. Brevnov MG, Pawar HS, Mundt J, Calandro LM, Furtado MR, Shewale JG (2009) Developmental validation of the PrepFiler™ forensic DNA extraction kit for extraction of genomic DNA from biological samples. J Forensic Sci 54 (3):599-607

69. Zgonjanin D, Antov M, Alghafri R, Petkovic S, Vukovic R, Stojiljkovic G, Toljic D (2017) DNA analysis from human skeletal remains in forensic casework. Forensic Sci Int Genet Suppl Ser 6:E342-E345. doi:10.1016/j.fsigss.2017.09.117

70. Bogas V, Balsa F, Carvalho M, Anjos MJ, Pinheiro MF, Corte-Real F (2011) Comparison of four DNA extraction methods for forensic application. Forensic Sci Int Genet Suppl Ser 3 (1):e194-e195. doi:10.1016/j.fsigss.2011.08.097

71. Joël J, Glanzmann B, Germann U, Cossu C (2015) DNA extraction of forensic adhesive tapes—A comparison of two different methods. Forensic Sci Int Genet Suppl Ser 5:e579-e581. doi:10.1016/j.fsigss.2015.09.229

72. Kallassy H, El Khoury LY, Eid M, Chalhoub M, Mansour I (2018) Comparison of four DNA extraction methods to extract DNA from cigarette butts collected in Lebanon. Sci Justice. doi:10.1016/j.scijus.2018.09.003

73. Iyavoo S, Hadi S, Goodwin W (2013) Evaluation of five DNA extraction systems for recovery of DNA from bone. Forensic Sci Int Genet Suppl Ser 4 (1):e174-e175. doi:10.1016/j.fsigss.2013.10.090

74. Weiner S, Price PA (1986) Disaggregation of Bone into Crystals. Calcified Tissue International 39 (6):365-375. doi:Doi 10.1007/Bf02555173

75. DeNiro MJ, Weiner S (1988) Chemical, enzymatic and spectroscopic characterization of "collagen" and other organic fractions from prehistoric bones. Geochi Cosmo Acta 52 (9):2197-2206

76. Loreille OM, Diegoli TM, Irwin JA, Coble MD, Parsons TJ (2007) High efficiency DNA extraction from bone by total demineralization. Forensic Sci Int Genet 1 (2):191-195. doi:10.1016/j.fsigen.2007.02.006

77. Edson SM, Ross J, Coble M, Parson T, Barritt S (2004) Naming the dead-confronting the realities of the rapid identification of degraded skeletal remains. Forensic Sci Rev 16 (1):63-88

78. Kus M, Ossowski A, Zielinska G (2016) Comparison of three different DNA extraction methods from a highly degraded biological material. J Forensic Leg Med 40:47-53. doi:10.1016/j.jflm.2016.03.002

79. Barbisin M, Fang R, O'Shea CE, Calandro LM, Furtado MR, Shewale JG (2009) Developmental validation of the Quantifiler® Duo DNA quantification kit for simultaneous quantification of total human and human male DNA and detection of PCR inhibitors in biological samples. J Forensic Sci 54 (2):305-319

80. Chen CM, Sio CP, Lu YL, Chang HT, Hu CH, Pai TW (2014) Identification of conserved and polymorphic STRs for personal genomes. BMC Genomics 15 Suppl 10 (10):S3. doi:10.1186/1471-2164-15-S10-S3

81. Bornman DM, Hester ME, Schuetter JM, Kasoji MD, Minard-Smith A, Barden CA, Nelson SC, Godbold GD, Baker CH, Yang B, Walther JE, Tornes IE, Yan PS, Rodriguez B, Bundschuh R, Dickens ML, Young BA, Faith SA (2012) Short-read, high-throughput sequencing technology for STR genotyping. Biotech Rapid Dispatches 2012:1-6

82. Edwards A, Civitello A, Hammond HA, Caskey CT (1991) DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet 49 (4):746-756

83. Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nat Rev Genet 5 (6):435-445. doi:10.1038/nrg1348

84. Churchill JD, Chang J, Ge J, Rajagopalan N, Wootton SC, Chang C-W, Lagacé R, King JL, Budowle B (2015) Blind study evaluation illustrates utility of the Ion PGM™ system for use in human identity DNA typing. Croat Med J 56 (3):218-229

85. Seo SB, King JL, Warshauer DH, Davis CP, Ge J, Budowle B (2013) Single nucleotide polymorphism typing with massively parallel sequencing for human identification. Int J Leg Med 127 (6):1079-1086. doi:10.1007/s00414-013-0879-7

86. Zeng X, King JL, Stoljarova M, Warshauer DH, LaRue BL, Sajantila A, Patel J, Storts DR, Budowle B (2015) High sensitivity multiplex short tandem repeat loci analyses with massively parallel sequencing. Forensic Sci Int Genet 16:38-47. doi:10.1016/j.fsigen.2014.11.022

87. Zeng X, King J, Hermanson S, Patel J, Storts DR, Budowle B (2015) An evaluation of the PowerSeq Auto System: A multiplex short tandem repeat marker kit compatible with massively parallel sequencing. Forensic Sci Int Genet 19:172-179. doi:10.1016/j.fsigen.2015.07.015

88. Wang DY, Chang CW, Lagacé RE, Calandro LM, Hennessy LK (2012) Developmental validation of the AmpFℓSTR® Identifiler® Plus PCR Amplification Kit: an established multiplex assay with improved performance. J Forensic Sci 57 (2):453-465

89. Yang Y, Xie B, Yan J (2014) Application of next-generation sequencing technology in forensic science. Genomics Proteomics Bioinformatics 12 (5):190-197. doi:10.1016/j.gpb.2014.09.001

90. Kraemer M, Prochnow A, Bussmann M, Scherer M, Peist R, Steffen C (2017) Developmental validation of QIAGEN Investigator® 24plex QS Kit and Investigator®

24plex GO! Kit: Two 6-dye multiplex assays for the extended CODIS core loci. Forensic Sci Int Genet 29:9-20

91. Oostdik K, Lenz K, Nye J, Schelling K, Yet D, Bruski S, Strong J, Buchanan C, Sutton J, Linner J (2014) Developmental validation of the PowerPlex® Fusion System for analysis of casework and reference samples: A 24-locus multiplex for new database standards. Forensic Sci Int Genet 12:69-76
92. Wang DY, Gopinath S, Lagacé RE, Norona W, Hennessy LK, Short ML, Mulero JJ (2015) Developmental validation of the GlobalFiler® Express PCR Amplification Kit: a 6-dye multiplex assay for the direct amplification of reference samples. Forensic Sci Int Genet 19:148-155
93. Ensenberger MG, Thompson J, Hill B, Homick K, Kearney V, Mayntz-Press KA, Mazur P, McGuckian A, Myers J, Raley K, Raley SG, Rothove R, Wilson J, Wieczorek D, Fulmer PM, Storts DR, Krenke BE (2010) Developmental validation of the PowerPlex 16 HS System: an improved 16-locus fluorescent STR multiplex. Forensic Sci Int Genet 4 (4):257-264. doi:10.1016/j.fsigen.2009.10.007
94. Reich DE, Schaffner SF, Daly MJ, McVean G, Mullikin JC, Higgins JM, Richter DJ, Lander ES, Altshuler D (2002) Human genome sequence variation and the influence of gene history, mutation and recombination. Nat Genet 32 (1):135-142. doi:10.1038/ng947
95. Nachman MW, Crowell SL (2000) Estimate of the mutation rate per nucleotide in humans. Genetics 156 (1):297-304

96. Huang QY, Xu FH, Shen H, Deng HY, Liu YJ, Liu YZ, Li JL, Recker RR, Deng HW (2002) Mutation patterns at dinucleotide microsatellite loci in humans. Am J Hum Genet 70 (3):625-634. doi:Doi 10.1086/338997

97. Kidd KK, Pakstis AJ, Speed WC, Lagace R, Chang J, Wootton S, Ihuegbu N (2013) Microhaplotype loci are a powerful new type of forensic marker. Forensic Sci Int Genet Suppl Ser 4 (1):e123-e124. doi:10.1016/j.fsigss.2013.10.063

98. Gill P (2001) An assessment of the utility of single nucleotide polymorphisms (SNPs) for forensic purposes. Int J Leg Med 114 (4-5):204-210

99. Kidd KK, Pakstis AJ, Speed WC, Grigorenko EL, Kajuna SL, Karoma NJ, Kungulilo S, Kim JJ, Lu RB, Odunsi A, Okonofua F, Parnas J, Schulz LO, Zhukova OV, Kidd JR (2006) Developing a SNP panel for forensic identification of individuals. Forensic Sci Int 164 (1):20-32. doi:10.1016/j.forsciint.2005.11.017

100. Mehta B, Daniel R, Phillips C, McNevin D (2017) Forensically relevant SNaPshot® assays for human DNA SNP analysis: a review. Int J Leg Med 131 (1):21-37

101. Budowle B, van Daal A (2008) Forensically relevant SNP classes. Biotechniques 44 (5):603-608, 610. doi:10.2144/000112806

102. Weiler NE, de Vries G, Sijen T (2016) Development of a control region-based mtDNA SNaPshot selection tool, integrated into a mini amplicon sequencing method. Sci Justice 56 (2):96-103. doi:10.1016/j.scijus.2015.11.003

103. Daniel R, Santos C, Phillips C, Fondevila M, van Oorschot RA, Carracedo A, Lareu MV, McNevin D (2015) A SNaPshot of next generation sequencing for forensic SNP analysis. Forensic Sci Int Genet 14:50-60. doi:10.1016/j.fsigen.2014.08.013

104. Walsh S, Liu F, Ballantyne KN, van Oven M, Lao O, Kayser M (2011) IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. Forensic Sci Int Genet 5 (3):170-180. doi:10.1016/j.fsigen.2010.02.004

105. Walsh S, Lindenbergh A, Zuniga SB, Sijen T, de Knijff P, Kayser M, Ballantyne KN (2011) Developmental validation of the IrisPlex system: determination of blue and brown iris colour for forensic intelligence. Forensic Sci Int Genet 5 (5):464-471. doi:10.1016/j.fsigen.2010.09.008

106. Walsh S, Wollstein A, Liu F, Chakravarthy U, Rahu M, Seland JH, Soubrane G, Tomazzoli L, Topouzis F, Vingerling JR, Vioque J, Fletcher AE, Ballantyne KN, Kayser M (2012) DNA-based eye colour prediction across Europe with the IrisPlex system. Forensic Sci Int Genet 6 (3):330-340. doi:10.1016/j.fsigen.2011.07.009

107. Liu F, van Duijn K, Vingerling JR, Hofman A, Uitterlinden AG, Janssens AC, Kayser M (2009) Eye color and the prediction of complex phenotypes from genotypes. Curr Biol 19 (5):R192-193. doi:10.1016/j.cub.2009.01.027

108. Walsh S, Liu F, Wollstein A, Kovatsi L, Ralf A, Kosiniak-Kamysz A, Branicki W, Kayser M (2013) The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. Forensic Sci Int Genet 7 (1):98-115. doi:10.1016/j.fsigen.2012.07.005

109. Chaitanya L, Pajnic IZ, Walsh S, Balazic J, Zupanc T, Kayser M (2017) Bringing colour back after 70 years: Predicting eye and hair colour from skeletal remains of World War II victims using the HIrisPlex system. Forensic Sci Int Genet 26:48-57. doi:10.1016/j.fsigen.2016.10.004

110. Phillips C (2015) Forensic genetic analysis of bio-geographical ancestry. Forensic Sci Int Genet 18:49-65. doi:10.1016/j.fsigen.2015.05.012

111. Fondevila M, Phillips C, Santos C, Freire Aradas A, Vallone PM, Butler JM, Lareu MV, Carracedo A (2013) Revision of the SNPforID 34-plex forensic ancestry test: Assay enhancements, standard reference sample genotypes and extended population studies. Forensic Sci Int Genet 7 (1):63-74. doi:10.1016/j.fsigen.2012.06.007

112. Phillips C, Salas A, Sanchez JJ, Fondevila M, Gomez-Tato A, Alvarez-Dios J, Calaza M, de Cal MC, Ballard D, Lareu MV, Carracedo A, Consortium SN (2007) Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs. Forensic Sci Int Genet 1 (3-4):273-280. doi:10.1016/j.fsigen.2007.06.008

113. Kersbergen P, van Duijn K, Kloosterman AD, den Dunnen JT, Kayser M, de Knijff P (2009) Developing a set of ancestry-sensitive DNA markers reflecting continental origins of humans. BMC Genet 10 (1):69. doi:10.1186/1471-2156-10-69

114. Thermo Fisher Scientific (2017) Precision ID Panels with the Ion S5 System: Application Guide. MAN0015831 Rev. B.0 edn. Thermo Fisher Scientific,

115. Verogen (2018) ForenSeq DNA Signature Prep Reference Guide. VD2018005 Rev. A edn. Verogen,

116. Gettings KB, Lai R, Johnson JL, Peck MA, Hart JA, Gordish-Dressman H, Schanfield MS, Podini DS (2014) A 50-SNP assay for biogeographic ancestry and phenotype prediction in the U.S. population. Forensic Sci Int Genet 8 (1):101-108. doi:10.1016/j.fsigen.2013.07.010

117. Phillips C, Freire Aradas A, Kriegel AK, Fondevila M, Bulbul O, Santos C, Serrulla Rech F, Perez Carceles MD, Carracedo A, Schneider PM, Lareu MV (2013)

Eurasiaplex: a forensic SNP assay for differentiating European and South Asian ancestries. Forensic Sci Int Genet 7 (3):359-366. doi:10.1016/j.fsigen.2013.02.010

118. Li CX, Pakstis AJ, Jiang L, Wei YL, Sun QF, Wu H, Bulbul O, Wang P, Kang LL, Kidd JR, Kidd KK (2016) A panel of 74 AISNPs: Improved ancestry inference within Eastern Asia. Forensic Sci Int Genet 23:101-110. doi:10.1016/j.fsigen.2016.04.002

119. Santos C, Phillips C, Fondevila M, Daniel R, van Oorschot RAH, Burchard EG, Schanfield MS, Souto L, Uacyisrael J, Via M, Carracedo A, Lareu MV (2016) Pacifiplex: an ancestry-informative SNP panel centred on Australia and the Pacific region. Forensic Sci Int Genet 20:71-80. doi:10.1016/j.fsigen.2015.10.003

120. Kidd KK, Pakstis AJ, Speed WC, Lagace R, Chang J, Wootton S, Haigh E, Kidd JR (2014) Current sequencing technology makes microhaplotypes a powerful new type of genetic marker for forensics. Forensic Sci Int Genet 12:215-224. doi:10.1016/j.fsigen.2014.06.014

121. Tishkoff SA, Verrelli BC (2003) Patterns of human genetic diversity: implications for human evolutionary history and disease. Annu Rev Genomics Hum Genet 4 (1):293-340. doi:10.1146/annurev.genom.4.070802.110226

122. Ge J, Budowle B, Planz JV, Chakraborty R (2010) Haplotype block: a new type of forensic DNA markers. Int J Leg Med 124 (5):353-361. doi:10.1007/s00414-009-0400-5

123. Pakstis AJ, Fang R, Furtado MR, Kidd JR, Kidd KK (2012) Mini-haplotypes as lineage informative SNPs and ancestry inference SNPs. Eur J Hum Genet 20 (11):1148-1154. doi:10.1038/ejhg.2012.69

124. dos Santos Pereira J, Fontes FL, de Medeiros SR, de Almeida Freitas R, de Souza LB, da Costa Miguel MC (2016) Association of the XPD and XRCC3 gene polymorphisms with oral squamous cell carcinoma in a Northeastern Brazilian population: A pilot study. Arch Oral Biol 64:19-23. doi:10.1016/j.archoralbio.2015.12.004

125. Wang H, Zhu J, Zhou N, Jiang YJ, Wang L, He W, Peng D, Su Q, Mao J, Chan D, Liang WB, Zhang L (2015) NGS technology makes microhaplotype a potential forensic marker. Forensic Sci Int Genet Suppl Ser 5:E233-E234. doi:10.1016/j.fsigss.2015.09.093

126. Cortellini V, Verzeletti A, Cerri N, Cisana S, De Ferrari F (2015) Evaluation of mtDNA stability across the maternal line: A study on three generations in two families. Forensic Sci Int Genet Suppl Ser 5:E19-E20. doi:10.1016/j.fsigss.2015.09.008

127. Hu CT, Yan JW, Chen F, Zhang QX, Wang HD, Yin CY, Fan HT, Hu LL, Shen CM, Meng HT, Zhang YD, Wang H, Zhu BF (2016) Genetic analysis of 15 mtDNA SNP loci in Chinese Yi ethnic group using SNaPshot minisequencing. Gene 576 (1 Pt 1):105-108. doi:10.1016/j.gene.2015.09.071

128. Parson W, Brandstatter A, Alonso A, Brandt N, Brinkmann B, Carracedo A, Corach D, Froment O, Furac I, Grzybowski T, Hedberg K, Keyser-Tracqui C, Kupiec T, Lutz-Bonengel S, Mevag B, Ploski R, Schmitter H, Schneider P, Syndercombe-Court D, Sorensen E, Thew H, Tully G, Scheithauer R (2004) The EDNAP mitochondrial DNA population database (EMPOP) collaborative exercises: organisation, results and perspectives. Forensic Sci Int 139 (2-3):215-226. doi:10.1016/j.forsciint.2003.11.008

129. van Oven M, Kayser M (2009) Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum Mutat 30 (2):E386-394. doi:10.1002/humu.20921

130. King JL, LaRue BL, Novroski NM, Stoljarova M, Seo SB, Zeng X, Warshauer DH, Davis CP, Parson W, Sajantila A, Budowle B (2014) High-quality and high-throughput massively parallel sequencing of the human mitochondrial genome using the Illumina MiSeq. Forensic Sci Int Genet 12:128-135. doi:10.1016/j.fsigen.2014.06.001

131. Parsons TJ, Coble MD (2001) Increasing the forensic discrimination of mitochondrial DNA testing through analysis of the entire mitochondrial DNA genome. Croat Med J 42 (3):304-309

132. Churchill JD, Peters D, Capt C, Strobl C, Parson W, Budowle B (2017) Working towards implementation of whole genome mitochondrial DNA sequencing into routine casework. Forensic Sci Int Genet Suppl Ser 6:E388-E389. doi:10.1016/j.fsigss.2017.09.167

133. Gouveia N, Brito P, Bogas V, Bento AM, Balsa F, Serra A, Lopes V, Sampaio L, São Bento M, Cunha P, Porto MJ (2017) Massively parallel sequencing of forensic samples using precision ID mtDNA whole genome panel on the ion S5™ system. Forensic Sci Int Genet Suppl Ser 6:e167-e168. doi:10.1016/j.fsigss.2017.09.057

134. Strobl C, Eduardoff M, Bus MM, Allen M, Parson W (2018) Evaluation of the precision ID whole MtDNA genome panel for forensic analyses. Forensic Sci Int Genet 35:21-25. doi:10.1016/j.fsigen.2018.03.013

135. Woerner AE, Ambers A, Wendt FR, King JL, Moura-Neto RS, Silva R, Budowle B (2018) Evaluation of the precision ID mtDNA whole genome panel on two massively

parallel sequencing systems. Forensic Sci Int Genet 36:213-224. doi:10.1016/j.fsigen.2018.07.015

136. Heather JM, Chain B (2016) The sequence of sequencers: The history of sequencing DNA. Genomics 107 (1):1-8. doi:10.1016/j.ygeno.2015.11.003

137. Sanger F, Coulson AR (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Mol Biol 94 (3):441-448

138. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74 (12):5463-5467

139. McElhoe JA, Holland MM, Makova KD, Su MS, Paul IM, Baker CH, Faith SA, Young B (2014) Development and assessment of an optimized next-generation DNA sequencing approach for the mtgenome using the Illumina MiSeq. Forensic Sci Int Genet 13:20-29. doi:10.1016/j.fsigen.2014.05.007

140. Morey M, Fernandez-Marmiesse A, Castineiras D, Fraga JM, Couce ML, Cocho JA (2013) A glimpse into past, present, and future DNA sequencing. Mol Genet Metab 110 (1-2):3-24. doi:10.1016/j.ymgme.2013.04.024

141. Mardis ERJN (2011) A decade's perspective on DNA sequencing technology. 470 (7333):198

142. Borsting C, Fordyce SL, Olofsson J, Mogensen HS, Morling N (2014) Evaluation of the Ion Torrent HID SNP 169-plex: A SNP typing assay developed for human identification by second generation sequencing. Forensic Sci Int Genet 12:144-154. doi:10.1016/j.fsigen.2014.06.004

143. Bottino CG, Chang CW, Wootton S, Rajagopalan N, Langit R, Lagace RE, Silva R, Moura-Neto RS (2015) STR genotyping using ion torrent PGM and STR 24-plex

system: Performance and data interpretation. Forensic Sci Int Genet Suppl Ser 5:E325-E326. doi:10.1016/j.fsigss.2015.09.129

144. Eduardoff M, Santos C, de la Puente M, Gross TE, Fondevila M, Strobl C, Sobrino B, Ballard D, Schneider PM, Carracedo A, Lareu MV, Parson W, Phillips C (2015) Inter-laboratory evaluation of SNP-based forensic identification by massively parallel sequencing using the Ion PGM. Forensic Sci Int Genet 17:110-121. doi:10.1016/j.fsigen.2015.04.007

145. Berglund EC, Kiialainen A, Syvanen AC (2011) Next-generation sequencing technologies and applications for human genetic history and forensics. Investig Genet 2 (1):23. doi:10.1186/2041-2223-2-23

146. Flores S, Sun J, King J, Budowle B (2014) Internal validation of the GlobalFiler Express PCR Amplification Kit for the direct amplification of reference DNA samples on a high-throughput automated workflow. Forensic Sci Int Genet 10:33-39. doi:10.1016/j.fsigen.2014.01.005

147. Bentley DR, Balasubramanian S, Swerdlow HP, et al. (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456 (7218):53-59. doi:10.1038/nature07517

148. Ronaghi M, Karamohamed S, Pettersson B, Uhlen M, Nyren P (1996) Real-time DNA sequencing using detection of pyrophosphate release. Anal Biochem 242 (1):84-89. doi:10.1006/abio.1996.0432

149. McKernan KJ, Peckham HE, Costa GL, et al. (2009) Sequence and structural variation in a human genome uncovered by short-read, massively parallel ligation sequencing using two-base encoding. Genome Res 19 (9):1527-1541. doi:10.1101/gr.091868.109

150. Life Technologies.

151. Holland MM, McQuillan MR, O'Hanlon KA (2011) Second generation sequencing allows for mtDNA mixture deconvolution and high resolution detection of heteroplasmy. Croat Med J 52 (3):299-313

152. Van Neste C, Van Nieuwerburgh F, Van Hoofstat D, Deforce D (2012) Forensic STR analysis using massive parallel sequencing. Forensic Sci Int Genet 6 (6):810-818

153. BiteSize Bio.

154. Goodwin S, McPherson JD, McCombie WR (2016) Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet 17 (6):333-351. doi:10.1038/nrg.2016.49

155. Davis C, Peters D, Warshauer D, King J, Budowle B (2015) Sequencing the hypervariable regions of human mitochondrial DNA using massively parallel sequencing: Enhanced data acquisition for DNA samples encountered in forensic testing. Leg Med (Tokyo) 17 (2):123-127. doi:10.1016/j.legalmed.2014.10.004

156. Parson W, Strobl C, Huber G, Zimmermann B, Gomes SM, Souto L, Fendt L, Delport R, Langit R, Wootton S, Lagace R, Irwin J (2013) Reprint of: Evaluation of next generation mtGenome sequencing using the Ion Torrent Personal Genome Machine (PGM). Forensic Sci Int Genet 7 (6):632-639. doi:10.1016/j.fsigen.2013.09.007

157. Zhang S, Bian Y, Chen A, Zheng H, Gao Y, Hou Y, Li C (2017) Developmental validation of a custom panel including 273 SNPs for forensic application using Ion Torrent PGM. Forensic Sci Int Genet 27:50-57. doi:10.1016/j.fsigen.2016.12.003

158. de la Puente M, Phillips C, Santos C, Fondevila M, Carracedo A, Lareu MV (2017) Evaluation of the Qiagen 140-SNP forensic identification multiplex for massively parallel sequencing. Forensic Sci Int Genet 28:35-43. doi:10.1016/j.fsigen.2017.01.012

159. Illumina (2015) ForenSeq DNA Signature Prep Reference Guide. vol 15049528, v01 edn., San Diego, CA

160. Illumina (2016) ForenSeq™ DNA Signature Prep Kit Data Sheet. www.illumina.com. Illumina, Illumina

161. Jim Robinson HT, Jill Mesirov (2018) Integrative Genomics Viewer. Broad Institute, Regents of the University of California. http://software.broadinstitute.org/software/igv/home.

162. Thorvaldsdottir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 14 (2):178-192. doi:10.1093/bib/bbs017

163. Illumina (2016) ForenSeq Universal Analysis Software Guide 15053876 edn. Illumina,

164. Warshauer DH, Lin D, Hari K, Jain R, Davis C, Larue B, King JL, Budowle B (2013) STRait Razor: a length-based forensic STR allele-calling tool for use with second generation sequencing data. Forensic Sci Int Genet 7 (4):409-417. doi:10.1016/j.fsigen.2013.04.005

165. Warshauer DH, King JL, Budowle B (2015) STRait razor v2. 0: the improved STR allele identification tool–razor. Forensic Sci Int Genet Suppl Ser 14:182-186

166. Woerner AE, King JL, Budowle B (2017) Fast STR allele identification with STRait Razor 3.0. Forensic Sci Int Genet 30:18-23. doi:10.1016/j.fsigen.2017.05.008

167. Elwick K, Zeng X, King J, Budowle B, Hughes-Stamm S (2018) Comparative tolerance of two massively parallel sequencing systems to common PCR inhibitors. Int J Legal Med 132 (4):983-995. doi:10.1007/s00414-017-1693-4

168. Gettings KB, Kiesler KM, Faith SA, Montano E, Baker CH, Young BA, Guerrieri RA, Vallone PM (2016) Sequence variation of 22 autosomal STR loci detected by next generation sequencing. Forensic Sci Int Genet 21:15-21. doi:10.1016/j.fsigen.2015.11.005

169. Novroski NMM, King JL, Churchill JD, Seah LH, Budowle B (2016) Characterization of genetic sequence variation of 58 STR loci in four major population groups. Forensic Sci Int Genet 25:214-226. doi:10.1016/j.fsigen.2016.09.007

170. Wendt FR, Warshauer DH, Zeng X, Churchill JD, Novroski NMM, Song B, King JL, LaRue BL, Budowle B (2016) Massively parallel sequencing of 68 insertion/deletion markers identifies novel microhaplotypes for utility in human identity testing. Forensic Sci Int Genet 25:198-209. doi:10.1016/j.fsigen.2016.09.005

171. King JL, Sajantila A, Budowle B (2014) mitoSAVE: mitochondrial sequence analysis of variants in Excel. Forensic Sci Int Genet 12:122-125. doi:10.1016/j.fsigen.2014.05.013

172. Sturk-Andreaggi K, Peck MA, Boysen C, Dekker P, McMahon TP, Marshall CK (2017) AQME: A forensic mitochondrial DNA analysis tool for next-generation sequencing data. Forensic Sci Int Genet 31:189-197. doi:10.1016/j.fsigen.2017.09.010

173. Rajeevan H, Soundararajan U, Kidd JR, Pakstis AJ, Kidd KK (2012) ALFRED: an allele frequency resource for research and teaching. Nucleic Acids Res 40 (Database issue):D1010-1015. doi:10.1093/nar/gkr924

174. Osier MV, Cheung K-H, Kidd JR, Pakstis AJ, Miller PL, Kidd KK (2001) ALFRED: an allele frequency database for diverse populations and DNA polymorphisms—an update. Nucleic Acids Res 29 (1):317-319

175. Lee JC, Tseng B, Chang LK, Linacre A (2017) SEQ Mapper: A DNA sequence searching tool for massively parallel sequencing data. Forensic Sci Int Genet 26:66-69. doi:10.1016/j.fsigen.2016.10.006

176. Van Neste C, Vandewoestyne M, Van Criekinge W, Deforce D, Van Nieuwerburgh F (2014) My-Forensic-Loci-queries (MyFLq) framework for analysis of forensic STR data generated by massive parallel sequencing. Forensic Sci Int Genet 9:1-8. doi:10.1016/j.fsigen.2013.10.012

CHAPTER II

Comparative sensitivity and inhibitor tolerance of GlobalFiler® PCR amplification and Investigator® 24plex QS kits for challenging samples[1]

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

[1] Elwick K, Mayes C, Hughes-Stamm S (2018) Comparative sensitivity and inhibitor tolerance of GlobalFiler® PCR amplification and Investigator® 24plex QS kits for challenging samples. Leg Med (Tokyo, Japan) https://doi.org/10.1016/j.legalmed.2018.01.006

Abstract

In cases such as mass disasters or missing persons, human remains are challenging to identify as they may be fragmented, burnt, been buried, decomposed, and/or contain inhibitory substances. This study compares the performance of a relatively new STR kit in the US market (Investigator® 24plex QS kit; QIAGEN) with the GlobalFiler® PCR Amplification kit (Thermo Fisher Scientific) when genotyping highly inhibited and low level DNA samples.

In this study, DNA samples (N = 3 in triplicate) ranging from 1 ng to 7.8 pg were amplified to define the sensitivity of two systems. In addition, DNA (1 ng and 0.1 ng input amounts) was spiked with various concentrations of five inhibitors common to human remains (humic acid, melanin, hematin, collagen, calcium). Furthermore, bone (N = 5) and tissue samples from decomposed human remains (N = 6) were used as mock casework samples for comparative analysis with both STR kits.

The data suggests that the GlobalFiler® kit may be slightly more sensitive than the Investigator® kit. On average STR profiles appeared to be more balanced and average peak heights were higher when using the GlobalFiler® kit. However, the data also shows that the Investigator® kit may be more tolerant to PCR inhibitors. While both STR kits showed a decrease in alleles as the inhibitor concentration increased, more complete profiles were obtained when the Investigator® kit was used.

Of the 11 bone and decomposed tissue samples tested, 8 resulted in more complete and balanced STR profiles when amplified with the GlobalFiler® kit.

Keywords: DNA typing, short tandem repeats, human identification, PCR inhibitors, GlobalFiler® PCR Amplification Kit, Investigator® 24plex QS Kit

Introduction

Forensic scientists may be tasked with identifying human remains in circumstances such as missing person cases, mass disasters, migrant deaths, and forensic cases.

Short tandem repeat (STR) analysis via capillary electrophoresis (CE) is the standard technology in forensic laboratories for human identification purposes [1-5]. STRs produce relatively short amplicons (<400 bp) [6-8] and are highly polymorphic [5,8,9] providing a high power of discrimination [8,10]. However, many factors can make STR typing of human remains more difficult by compromising the quantity and/or quality of the DNA for analysis.

Environmental insults such as UV exposure, humidity, and microbial infestation can result in severe damage and degradation of DNA in biological samples [11-14]. Other environmental conditions can introduce added complications for DNA typing in the form of PCR inhibitors, such as humic and fumic acid in buried samples [15,16]. In addition, biological tissues such as bone, hair, teeth, and blood contain various PCR inhibitors that may be co-extracted with the DNA [17-19]. Commonly co-extracted inhibitors include humic acid, hematin, collagen, calcium, melanin, indigo, bile salt, and urea, and each have different mechanisms by which they inhibit DNA amplification [17,18, 20-22].

PCR inhibition is the most common cause of PCR failure when adequate amounts of DNA are present [22]. Inhibitors can cause total failure of the PCR resulting in no amplified products or may simply reduce the efficiency of the PCR. In this way, inhibited samples may mimic low template samples, as the amount of DNA available for amplification can be greatly reduced. Negative downstream effects such as allele dropout,

lower peak heights, peak height imbalance, stutter, locus-specific dropout, and poor sensitivity may also be observed [20,23].

This study evaluated the sensitivity and performance of the GlobalFiler® PCR Amplification and Investigator® 24plex QS kits with low template and challenging samples. In addition, we also assessed the comparative tolerance of each kit to PCR inhibitors commonly associated with human remains. The two kits being tested are both relatively new 6-dye multiplex kits each with 24 markers. The Investigator® kit also contains two internal PCR controls (Quality Sensor QS1 (74 bp) and QS2 (435 bp)), which are designed to detect PCR inhibition or confirm DNA degradation and amplification success in general. Developmental validation studies have been performed for each of these kits [24,25], and various other studies [25-27] have described the utility of these papers with various types of samples. However, this study reports the comparative performance of these two STR kits with a much wider range of inhibitors, and when the amount of DNA template is both relatively high (1 ng) and low (0.1 ng).

Materials and Methods

Sample Preparation

Three sources of DNA were used for the sensitivity and inhibitor studies including NIST standard 2372 Component A Male (National Institute of Standards and Technology, Gaithersburg, MD) and two male donors. Donor samples (semen) were obtained in accordance with Sam Houston State University (SHSU) Institutional Review Board guidelines (# 2015-12-26123) and extracted using the AllPrep® DNA/RNA Micro Kit (QIAGEN Inc., Hilden, Germany) [28]. For the sensitivity study, all three DNA sources

were amplified in triplicate using template input amounts of 1 ng, 0.8 ng, 0.5 ng, 0.25 ng, 0.125 ng, 0.0625 ng, 0.0313 ng, 0.0156 ng, and 0.0078 ng.

The bone (N = 5) and decomposed tissue (N = 6) samples were obtained from bodies willed to the Southeast Texas Applied Forensic Science Facility (STAFS). Bone samples were extracted using a complete demineralization protocol [29] or the PrepFiler Express™ BTA kit [30]. DNA was purified from decomposed muscle samples using the QIAamp® DNA Investigator® kit [31]. All samples were quantified using the Quantifiler® Trio DNA Quantification kit (Thermo Fisher Scientific, Waltham, MA) on the 7500 Real-Time PCR System as per manufacturer's instructions [32].

Inhibitor Preparation

Five inhibitors (humic acid, melanin, hematin, collagen, and calcium) were tested in this study. A range of inhibitor concentrations was used to test the tolerance of both PCR amplification kits to high amounts of PCR inhibition (Table 2.1). All inhibitor stocks were prepared according to guidelines established in Opel et al. [23]. All inhibitors were prepared in 10 mL volumes and any dilutions made were prepared with deionized water.

Table 2.1 Final concentrations of the five PCR inhibitors (in 25 µL reaction)

Inhibitor	Units	Inhibitor Concentrations					
		0	**1**	**2**	**3**	**4**	**5**
Humic Acid	ng/µL	0	50	100	200	225	250
Melanin	ng/µL	0	25	35	40	45	50
Hematin	µM	0	300	500	1000	1050	1100
Collagen	ng/µL	0	50	100	112.5	130	160
Calcium	µM	0	250	350	500	650	850

STR Amplification

STR typing was performed using the GlobalFiler® PCR Amplification kit (Thermo Fisher Scientific) and the Investigator® 24plex QS kit (QIAGEN) in 25 µL reaction volumes as per manufacturer's instructions [24,33]. Bone and tissue samples were amplified with 0.8 ng of DNA (or maximum sample volume (15 µL) if DNA was less than 0.053 ng/µL). Inhibited samples were amplified with 10 µL of each inhibitor at the required concentration (Table 2.1) and 5 µL of DNA (0.2 ng/µL or 0.02 ng/µL). Inhibitor controls were performed using sterilized deionized water in lieu of the inhibitor. PCR amplification was performed on a GeneAmp® PCR System 9700 (Applied Biosystems, Carlsbad, CA) using the cycling parameters recommended for each amplification kit [24,33].

Capillary Electrophoresis and Data Interpretation

PCR products were separated and detected via capillary electrophoresis according to the respective manufacturer protocols using a 3500 Genetic Analyzer with a 36 cm capillary array and POP-4 (Thermo Fisher Scientific). A 5 s injection time at 15 kV was used for GlobalFiler®, and a 30 s injection time at 13 kV was used for the Investigator® 24plex QS kit (as recommended for each kit). STRs were analyzed using GeneMapper ID-X v. 1.4 (Applied Biosystems). Based on internal validation data, for GlobalFiler®, allele peaks were assigned using an analytical threshold of 150 RFUs and a stochastic threshold of 600 RFUs and for Investigator® 24plex QS, peaks were assigned using an analytical threshold of 100 RFUs and a stochastic threshold of 200 RFUs. When the RFU value of a heterozygote peak was below the analytical threshold it was considered dropout. However, if a homozygote peak was below the stochastic threshold, one allele was considered dropped out. Average peak height for each sample was calculated by adding the peak

heights of each allele and then dividing by the total number of alleles possible. Average heterozygote peak height ratios (PHRs) were calculated for each sample by averaging the values of the peak with the smaller RFU value divided by the peak with the larger RFU value at each heterozygous locus. The QS1 and QS2 peaks were used to assess the level of PCR inhibition. A Q/S ratio was calculated by dividing the peak height of the QS1 peak by the height of the QS2 peak.

Results and Discussion

Sensitivity Study

To test the sensitivity of the GlobalFiler® PCR Amplification and Investigator® 24plex QS kits, three DNA sources were diluted from 1 ng to 7.8 pg and tested in triplicate. Overall, as the DNA input decreased, STR profile quality and completeness also decreased. Complete STR profiles were obtained from all samples using both kits down to 250 pg. At 125 pg, GlobalFiler® recovered 100% of alleles, while the Investigator® kit recovered 91% of alleles. Both kits detected less than ~55% of alleles at 31.25 pg of DNA (**Appendix 2.1**), with GlobalFiler® continuing to detect more alleles than the Investigator® kit down to 7.8 pg (Fig. 2.1). A previous study [25] reported that the Investigator® kit produced full STR profiles with 125 pg and 50% of expected alleles with 8 pg of DNA. However, in our study, 91% of alleles were reported at 125 pg and 1% of alleles were reportable at 8 pg, which is consistent with another study by Martín et al. [26].

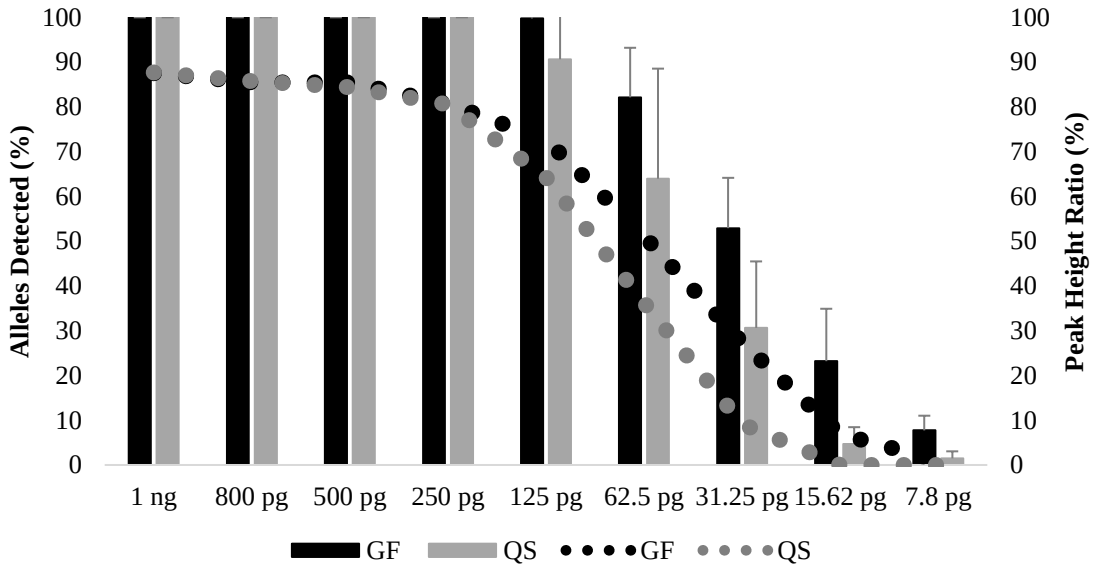


Figure 2.1 Sensitivity study comparing profile completeness (Bar; percentage of alleles detected) and profile balance (Line; PHR) when amplification is performed using the GlobalFiler® PCR Amplification and Investigator® 24plex QS kits. Error bars represent mean ± SD (N = 3 in triplicate).

Average PHRs were comparable until 125 pg, where the average PHRs decreased from approximately 88% to 73% with the GlobalFiler® kit and 88% to 64% with the Investigator® kit. The average PHRs of the Investigator® kit remained lower than those of the GlobalFiler® for all template amounts <250 pg (Fig. 2.1). Amplification with the GlobalFiler® kit also resulted in higher average peak heights than the Investigator® kit, producing peaks at approximately twice the height of the Investigator® kit at all template amounts tested (Fig. 2.2).

In general, these data suggest that the GlobalFiler® PCR Amplification kit may be slightly more sensitive than the Investigator® 24plex QS kit.

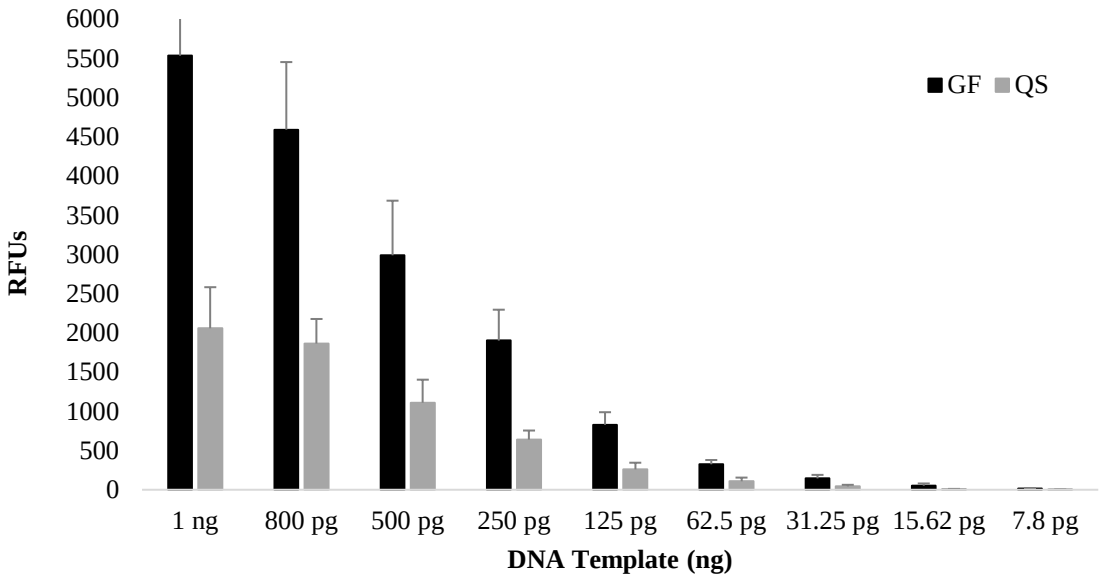


Figure 2.2 APH of the GlobalFiler® PCR Amplification and Investigator® 24plex QS kits when DNA input ranged from 1 ng to 7.8 pg (N = 3 in triplicate). Error bars represent mean ± SD.

Inhibitor Tolerance

To test the inhibitor tolerance of the GlobalFiler® PCR Amplification kit and the Investigator® 24plex QS kit, three DNA sources were amplified once.

Reportable Alleles

Overall, STR typing was more successful when inhibited samples were amplified using the Investigator® 24plex QS kit than with the GlobalFiler® kit. As expected, the amplification of 1 ng of DNA resulted in more complete STR profiles for each inhibitor concentration tested compared to 0.1 ng of DNA with both kits (Fig. 2.3). With 1 ng DNA in the PCR, more DNA was available for amplification despite the presence of inhibitory agents. Complete STR profiles were obtained from both 1 ng and 0.1 ng (uninhibited control DNA).

All inhibitors resulted in a decreasing number of reportable alleles when both 0.1 ng and 1 ng of template DNA was amplified using the GlobalFiler® kit (Fig. 2.3A & B). With 1 ng DNA, the GlobalFiler® kit appeared to be more tolerant to collagen, calcium, and melanin than humic acid and hematin. Almost 50% of alleles were reported at the highest concentration for collagen, calcium, and melanin, while humic acid and hematin resulted in 34% and 9% of reportable alleles (Fig. 2.3A). With less DNA (0.1 ng) the same pattern was observed but with a higher degree of allele and locus dropout (Fig. 2.3B).

When 1 ng of DNA was amplified using the Investigator® 24plex QS kit, complete (or near complete) profiles were generated at all inhibitor concentrations with all inhibitors (Fig. 2.3C). However, when less DNA (0.1 ng) was amplified, all inhibitors except calcium showed a decreasing number of reportable alleles (Fig. 2.3D). At the lower DNA input value (0.1 ng), over 80% of alleles were called with samples spiked with calcium and humic acid, while samples spiked with melanin, hematin, and collagen reported an average of less than 60% of alleles at the highest inhibitor concentrations (Fig. 2.3D). This decrease in reportable alleles suggests that the Investigator® 24plex QS kit is more susceptible to these inhibitors. However, overall, the kits performed comparably when samples were spiked with the lowest concentration of inhibitors.

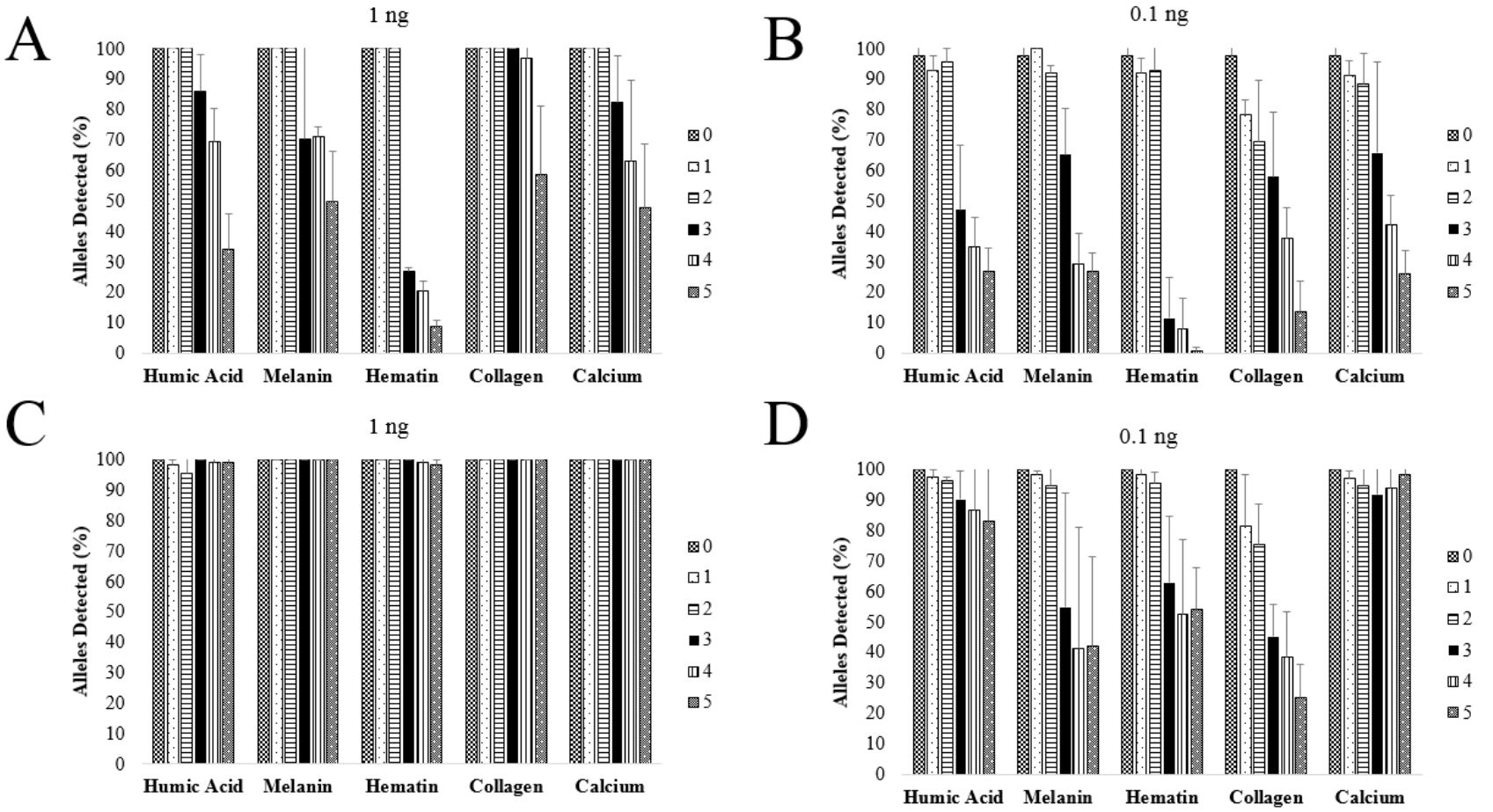


Figure 2.3 Percentage of alleles detected with increasing concentrations of five inhibitors (0, 1, 2, 3, 4, 5 correlate with the concentrations for each inhibitor in Table 1; 0 indicates no inhibitor added) A) with 1 ng of amplified DNA using GlobalFiler®, B) with 0.1 ng of amplified DNA using GlobalFiler®, C) with 1 ng of amplified DNA using Investigator® 24plex QS, D) with 0.1 ng of amplified DNA using Investigator® 24plex QS (N =3).

Allele Dropout

Average allele dropout rates were calculated for each locus when 0.1 and 1 ng of DNA was amplified with increasing concentrations of inhibitor (humic acid, melanin, hematin, collagen, or calcium). As expected, higher allele dropout rates were observed with 0.1 ng DNA compared to 1 ng regardless of the STR kit used (Fig. 2.4). Overall, the GlobalFiler® kit appeared more susceptible to dropout than the Investigator® kit when highly inhibited samples were amplified (Fig. 2.4A & B). The GlobalFiler® dropout rate ranged from 10% to 73% for 0.1 ng of DNA and 0% to 48% for 1 ng of DNA for all inhibitors combined (Fig. 2.4A). The Investigator® kit showed much lower dropout rates ranging from 3% to 42% for 0.1 ng of DNA and 0% to 3% for 1 ng for all inhibitors combined (Fig. 2.4B).

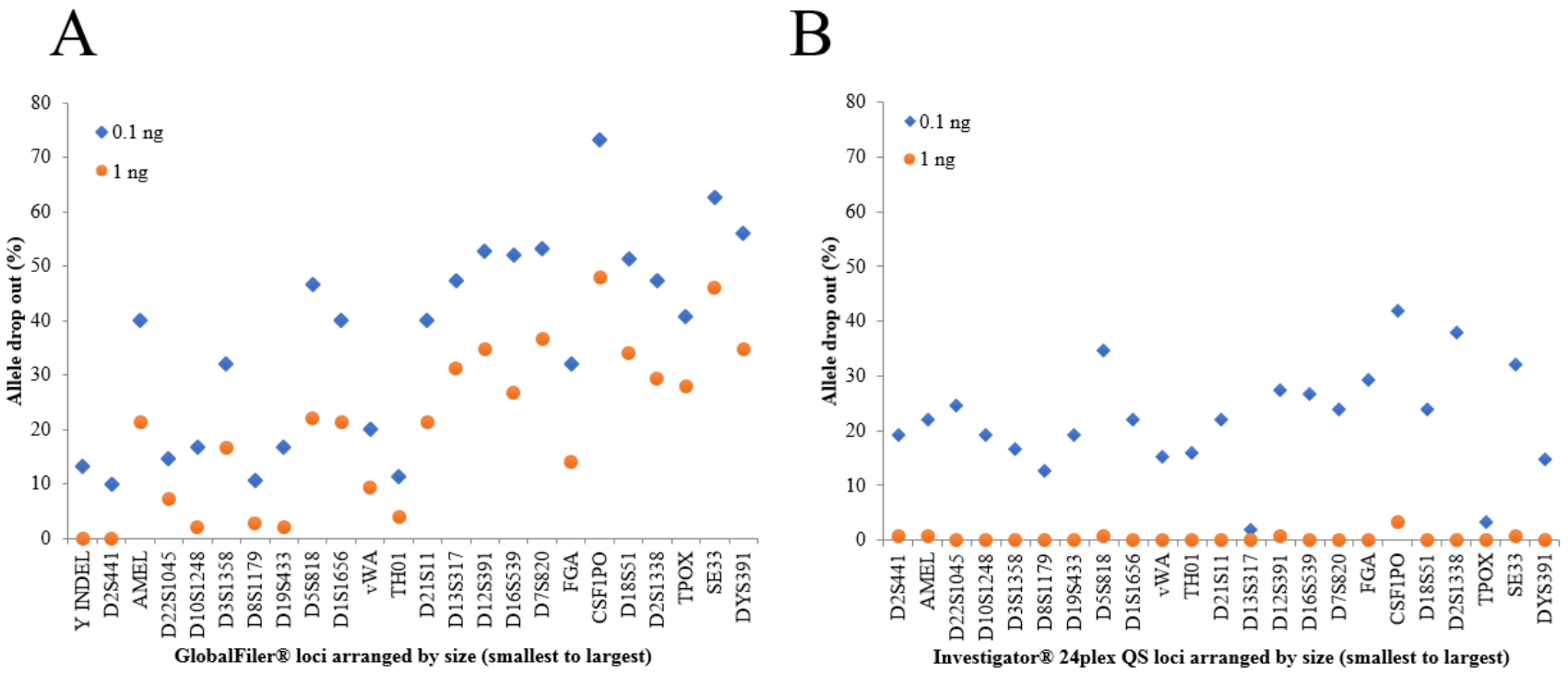


Figure 2.4 Allele drop out rates averaged across all inhibitors with 0.1 ng and 1 ng DNA input for A) GlobalFiler® PCR Amplification kit and B) Investigator® 24plex QS kit (N =3).

For samples amplified with GlobalFiler®, the longer STR markers (D7S820, CSF1PO, and SE33) appeared to be slightly more susceptible to PCR inhibition at both 0.1 and 1 ng of DNA than the other loci (Fig. 2.4A). These data suggest that the larger loci may be more susceptible to dropout in the presence of PCR inhibitors. In addition, the same loci that displayed allele dropout with 1 ng also showed (greater) dropout at 0.1 ng suggesting that dropout is not simply related to the amount of DNA in the PCR reaction, but also the locus itself (or size of amplicon). Pionzio et al. [34] reported that amplicon size has a large impact on the amount of inhibition in real-time PCR and also state that inhibitors often cause large STR loci to drop out first while the smaller loci amplify well in the presence of inhibitors. Dropout of alleles at larger loci is commonly observed with inhibitors because smaller loci are generally more resistant to inhibition [20].

Very little allele dropout was observed with 1 ng of DNA when samples were amplified with the Investigator® 24plex QS kit (Fig. 2.4B). However, with 0.1 ng of DNA, the dropout rate increased, but showed no relationship to locus size. These data may suggest that the Investigator® kit is more tolerant to inhibition, and unlike the trend observed with the GlobalFiler® kit, does not preferentially affect the larger amplicons. Regardless of the DNA input, CSF1PO dropped out most frequently, followed by SE33 (Fig. 2.4). Kraemer et al. [25] also observed dropout of CSF1PO, D2S1338, and SE33 at high inhibitor concentrations with the same inhibitors used in this study (humic acid, hematin, calcium, and collagen) when using the Investigator® kit.

Average Peak Height and Intra-Locus Peak Height Balance

As expected, the average peak height (APH) decreased as the concentration of each inhibitor increased (Fig. 2.5). This trend was observed with both 1 ng and 0.1 ng of DNA

input. When the GlobalFiler® kit was used to amplify 1 ng and 0.1 ng of non-inhibited DNA (control), the APHs were 15,500 RFUs and 950 RFUs respectively. When 1 ng of DNA was spiked with increasing concentrations of inhibitors, the APHs dropped to ~100-700 RFUs and with 0.1 ng decreased to ~60-900 RFUs. However, the same decreasing APH trend was not clearly observed when inhibited samples were amplified with the Investigator® kit (Fig. 2.5A & C). The APHs of STR profiles generated from uninhibited DNA (controls) with 1 ng and 0.1 ng were ~3100 RFUs and ~550 RFUs, respectively (Fig. 4C & D). When 1 ng of DNA was spiked with increasing concentrations of inhibitors, the APHs remained ~2000-4000 RFUs (Fig. 2.5C). However, when 0.1 ng of DNA was amplified the APHs did decrease as the melanin, hematin, and collagen inhibitor concentrations increased (Fig. 2.5D). The average peak height of each inhibitor ranged from ~60 RFUs to ~550 RFUs.

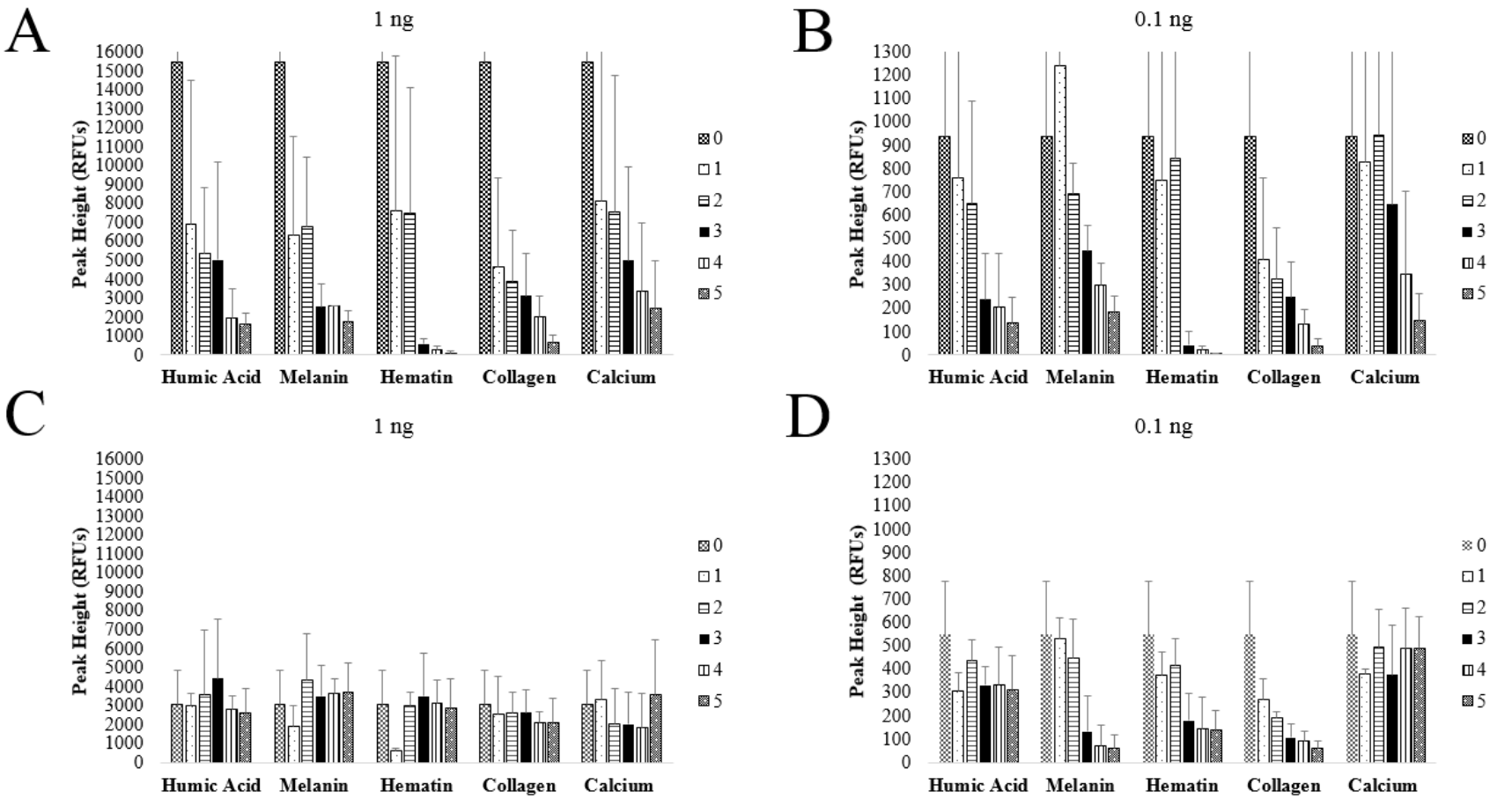


Figure 2.5 Average peak height with increasing concentrations of five inhibitors A) with 1 ng of amplified DNA using GlobalFiler®, B) with 0.1 ng of amplified DNA using GlobalFiler®, C) with 1 ng of amplified DNA using Investigator® 24plex QS, D) with 0.1 ng of amplified DNA using Investigator® 24plex QS (N = 3).

In general, the Investigator® kit seemed to produce slightly more balanced STR profiles than the GlobalFiler® kit for all inhibitors tested (Fig. 2.6). When 1 ng of inhibited DNA was amplified with the GlobalFiler® kit, the average PHRs ranged from ~4% to 88% (compared to 76% in the uninhibited control sample). A previous study [26] reported similar average PHRs above 80% when GlobalFiler® was used to amplify casework-type samples such as saliva, blood, semen, tissues, bones, and teeth. With 0.1 ng of DNA, all inhibitor PHRs decreased as the concentration of inhibitors increased (Fig. 2.6B). When 0.1 ng of inhibited DNA was amplified with the GlobalFiler® kit, the average PHRs ranged from 5% to 73% (compared to 85% in the uninhibited control sample). Complete PCR failure was observed at the highest concentration of hematin.

The average PHRs of STR profiles generated with the Investigator® 24plex QS kit from 1 ng and 0.1 ng of uninhibited DNA (controls) were 78% and 71%, respectively (Fig. 2.6C & D). When 0.1 ng of inhibited DNA was amplified, peak height balance was reduced by an average of ~10% (Fig. 2.6D). Humic acid, melanin, hematin, and collagen PHRs decreased as the inhibitor concentrations increased. However, samples spiked with calcium remained well balanced as the inhibitor concentrations increased.

These data suggest that although the Investigator® kit may be more resistant to inhibitors than the GlobalFiler® kit (regardless of the input DNA template), both kits generated reasonably balanced DNA profiles when low levels of inhibitor were encountered.

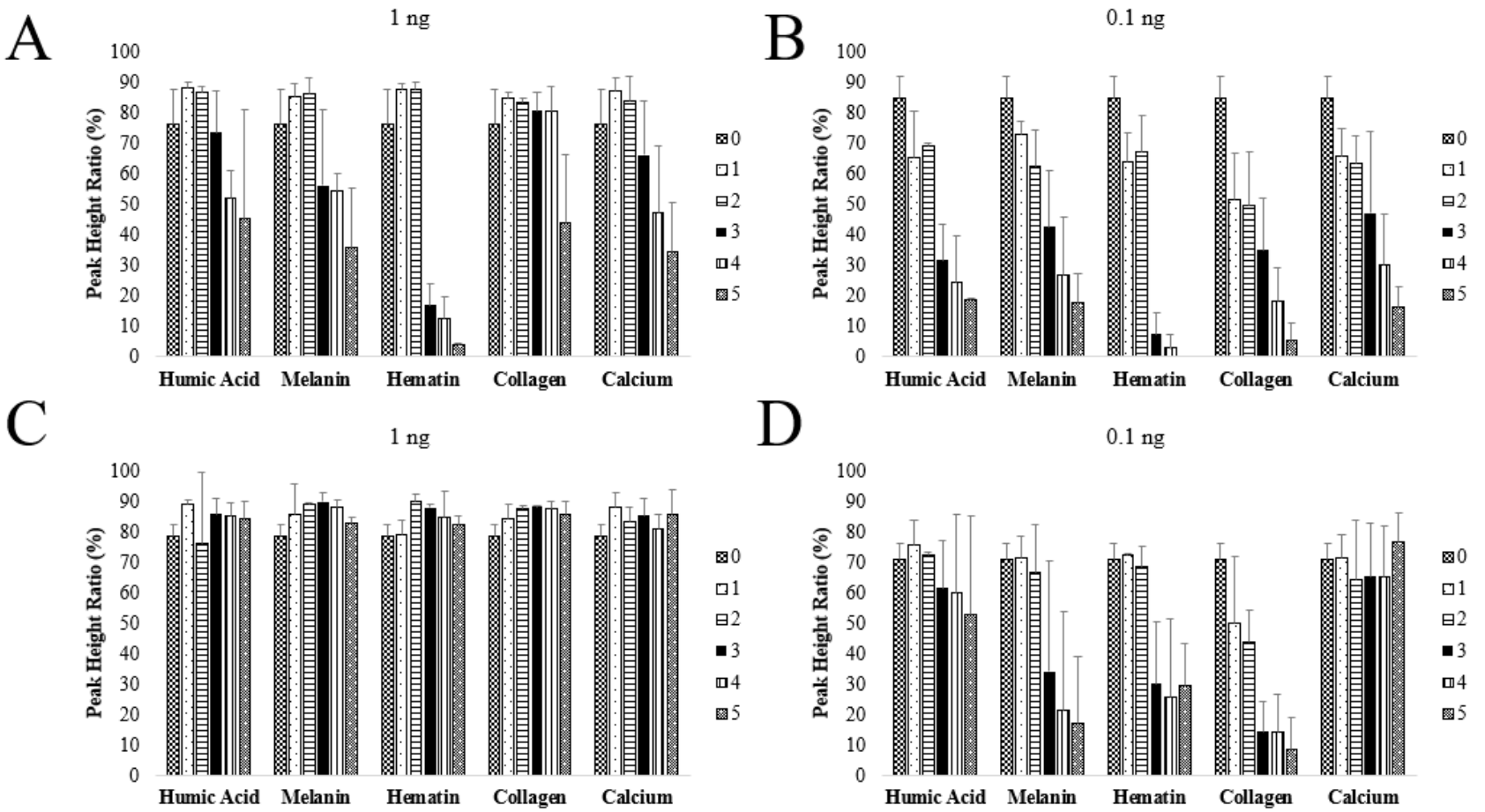


Figure 2.6 Heterozygote peak height ratios with increasing concentrations of five inhibitors A) with 1 ng of amplified DNA using GlobalFiler®, B) with 0.1 ng of amplified DNA using GlobalFiler®, C) with 1 ng of amplified DNA using Investigator® 24plex QS, D) with 0.1 ng of amplified DNA using Investigator® 24plex QS (N = 3).

Quality Sensors

Although the QS system was designed to qualitatively assess sample quality and serve as an internal PCR control, we have used the ratio of the Q/S peak heights in this study as a quantitative measure. The Q/S ratio is a means of detecting PCR inhibition in individual samples visualized within the STR electropherogram. In this data set, complete drop out of the QS2 peak was observed when the Q/S ratio was ≥ 3.84. Therefore, any sample where the QS2 dropped out was given an arbitrary Q/S value of 5.

The Q/S ratio of the control samples (no inhibitor) was 0.68 and 0.59 when 1 ng and 0.1 ng DNA was amplified, respectively. With 1 ng of DNA input, the QS1 and QS2 peaks were present in all samples with the exception of one replicate in hematin at the highest concentration. As the concentration of the inhibitor increased, the Q/S ratio seemed to increase slightly in samples spiked with humic acid, melanin, hematin, and collagen, while not at all for samples spiked with calcium (Fig. 2.7A). These data suggest that the Investigator® chemistry is tolerant to all inhibitors tested. However, with 0.1 ng DNA input, the QS2 sensor was most aggressively affected by hematin and collagen, while the Q/S ratio was comparable for calcium, melanin, and humic acid (Fig. 2.7A). As collagen and hematin are reported to bind to DNA [23], this potentially explains why the Q/S ratio was much lower when more DNA template was included; more inhibitor is binding to the template DNA and therefore reducing the amount of inhibitor available to interfere with the amplification of the quality sensors. Overall, the performance of the QS1 sensor is consistent with Scherer et al. [27], reporting that the QS system operates as it is designed to with the QS1 sensor being very stable in the presence of extremely high inhibitor

concentrations, in contrast to the QS2 sensor, which only indicated inhibition with high levels of hematin and collagen.

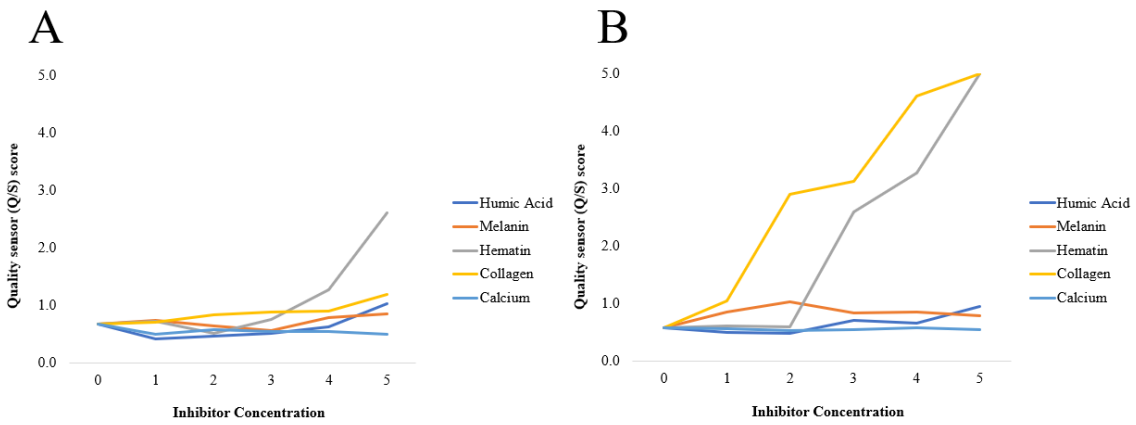


Figure 2.7 Quality sensor (Q/S) scores across all concentrations of 5 inhibitors at A) 1 ng and B) 0.1 ng DNA input.

Bone and Decomposed Tissue Samples

A total of 11 mock casework samples (5 bone and 6 decomposed muscle tissues) were amplified and genotyped once using both GlobalFiler® and Investigator® kits (Table 2.2). All genotypes for these samples were known. All bone samples resulted in more complete (12 - 32% more alleles) and more balanced STR profiles when amplified with the GlobalFiler® kit (Fig. 2.8). However, the results of the tissue samples were variable. Half of the samples resulted in higher number of reportable alleles with the GlobalFiler® kit (samples 8, 10, and 11) while the other half (samples 6, 7, and 9) showed more complete profiles with the Investigator® kit (Fig. 2.8). Of the 11 samples amplified, 8 samples showed better results when using the GlobalFiler® kit. In almost all samples, QS peaks were present and Q/S ratios did not indicate significant levels of PCR inhibition. The two kits produced concordant STR profiles for all samples. Due to the small sample size, no

definitive conclusion can be made as to which amplification kit may produce the most successful STR typing results with such challenging samples.

Table 2.2 Sample information for bone and decomposed tissue samples amplified with GlobalFiler® PCR Amplification and Investigator® 24plex QS kits. The quantitation value, Degradation Index (DI), and Inhibition (ΔCT) were determined using Quantifiler® Trio.

Sample No.	Sample Type	Insult	Quant	DI	ΔCT
1	Bone	Burned	0.5877	2.03	-0.46
2		Buried	0.0177	6.06	-0.22
3		Buried	0.0052	16.28	-0.35
4		Skeletonized	0.0130	2.10	-0.34
5		Embalmed	0.2041	34.65	-0.41
6	Muscle	Decomp	0.0007	2.04	-0.56
7		Decomp	1.0432	12.73	-0.33
8		Decomp	0.0053	4.01	0.18
9		Decomp	0.0021	7.69	0.01
10		Decomp	0.1649	46.8	-0.59
11		Decomp	0.0019	4.36	-0.06

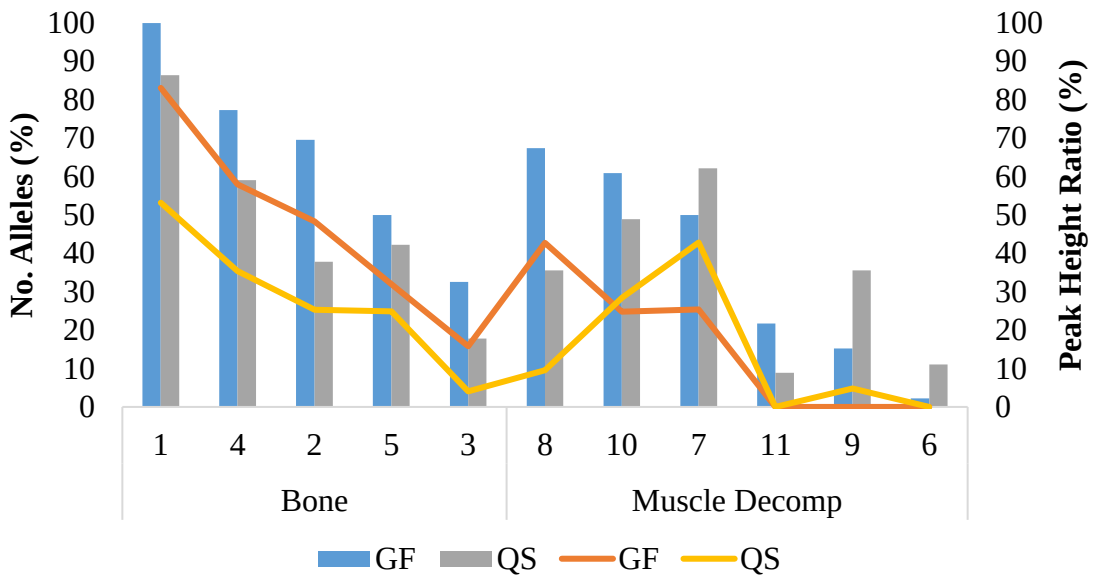


Figure 2.8 Comparing profile completeness and APHR using GlobalFiler® PCR Amplification and Investigator® 24plex QS kits with bone and decomposing tissue mock casework samples.

Conclusions

In this study, we examined the comparative sensitivity and performance of two commercial STR kits (GlobalFiler® PCR Amplification kit and Investigator® 24plex QS kit) with low template, highly inhibited, and challenging samples. We assessed the tolerance of both STR kits to five PCR inhibitors (humic acid, melanin, hematin, collagen, and calcium) to aid in the analysis of human remains from forensic, missing persons, and mass disaster cases.

The results of this research suggest that the GlobalFiler® kit is slightly more sensitive than the Investigator® 24plex QS kit, producing more complete and balanced STR profiles with peak heights at least 2-fold greater. However, the Investigator® kit was more tolerant than the GlobalFiler® kit to all of the PCR inhibitors tested in this study when both 1 ng and 0.1 ng of DNA was amplified. It should also be noted that this study

was designed to test the upper limits of inhibitor tolerance, and therefore both kits may be expected to perform more comparably with samples that contain substantially lower concentrations of these inhibitors.

Acknowledgements

The authors would like to thank the individuals who provided the biological samples used in this study. This project was partially supported by NIJ Award Number 2015-DN-BX-K066 awarded to SHS. We would also like to thank Rachel Houston for her comments and review of the manuscript and Esiri Tasker for providing assistance.

References

1. Zubakov D, Kokmeijer I, Ralf A, Rajagopalan N, Calandro L, Wootton S, Langit R, Chang C, Lagace R, Kayser M (2015) Towards simultaneous individual and tissue identification: A proof-of-principle study on parallel sequencing of STRs, amelogenin, and mRNAs with the Ion Torrent PGM. Forensic Sci Int Genet 17:122-128. doi:10.1016/j.fsigen.2015.04.002
2. Van Neste C, Van Nieuwerburgh F, Van Hoofstat D, Deforce D (2012) Forensic STR analysis using massive parallel sequencing. Forensic Sci Int Genet 6 (6):810-818
3. Templeton JEL, Brotherton PM, Llamas B, Soubrier J, Haak W, Cooper A, Austin JJ (2013) DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification. Investig Genet 4 (1):26
4. Alvarez-Cubero MJ, Saiz M, Martinez-Gonzalez LJ, Alvarez JC, Eisenberg AJ, Budowle B, Lorente JA (2012) Genetic identification of missing persons: DNA analysis of human remains and compromised samples. Pathobiology 79 (5):228-238. doi:10.1159/000334982
5. Bornman DM, Hester ME, Schuetter JM, Kasoji MD, Minard-Smith A, Barden CA, Nelson SC, Godbold GD, Baker CH, Yang B, Walther JE, Tornes IE, Yan PS, Rodriguez B, Bundschuh R, Dickens ML, Young BA, Faith SA (2012) Short-read, high-throughput sequencing technology for STR genotyping. Biotech Rapid Dispatches 2012:1-6

6. Seo SB, King JL, Warshauer DH, Davis CP, Ge J, Budowle B (2013) Single nucleotide polymorphism typing with massively parallel sequencing for human identification. Int J Leg Med 127 (6):1079-1086. doi:10.1007/s00414-013-0879-7
7. Luce C, Montpetit S, Gangitano D, O'Donnell P (2009) Validation of the AMPFℓSTR® MiniFilerTM PCR Amplification Kit for use in forensic casework. J Forensic Sci 54 (5):1046-1054
8. Green RL, Lagace RE, Oldroyd NJ, Hennessy LK, Mulero JJ (2013) Developmental validation of the AmpFlSTR(R) NGM SElect PCR Amplification Kit: A next-generation STR multiplex with the SE33 locus. Forensic Sci Int Genet 7 (1):41-51. doi:10.1016/j.fsigen.2012.05.012
9. Chen CM, Sio CP, Lu YL, Chang HT, Hu CH, Pai TW (2014) Identification of conserved and polymorphic STRs for personal genomes. BMC Genomics 15 Suppl 10 (10):S3. doi:10.1186/1471-2164-15-S10-S3
10. Elena S, Alessandro A, Ignazio C, Sharon W, Luigi R, Andrea B (2016) Revealing the challenges of low template DNA analysis with the prototype Ion AmpliSeq Identity panel v2.3 on the PGM Sequencer. Forensic Sci Int Genet 22:25-36. doi:10.1016/j.fsigen.2015.07.011
11. Hall A, Sims LM, Ballantyne J (2014) Assessment of DNA damage induced by terrestrial UV irradiation of dried bloodstains: forensic implications. Forensic Sci Int Genet 8 (1):24-32. doi:10.1016/j.fsigen.2013.06.010
12. Diegoli TM, Farr M, Cromartie C, Coble MD, Bille TW (2012) An optimized protocol for forensic application of the PreCR Repair Mix to multiplex STR amplification of

UV-damaged DNA. Forensic Sci Int Genet 6 (4):498-503. doi:10.1016/j.fsigen.2011.09.003

13. Fondevila M, Phillips C, Naverán N, Cerezo M, Rodríguez A, Calvo R, Fernández LM, Carracedo Á, Lareu MV (2008) Challenging DNA: Assessment of a range of genotyping approaches for highly degraded forensic samples. Forensic Sci Int Genet Suppl Ser 1 (1):26-28. doi:10.1016/j.fsigss.2007.10.057
14. Vernarecci S, Ottaviani E, Agostino A, Mei E, Calandro L, Montagna P (2015) Quantifiler® Trio Kit and forensic samples management: a matter of degradation. Forensic Sci Int Genet 16:77-85
15. Kasu M, Shires K (2015) The validation of forensic DNA extraction systems to utilize soil contaminated biological evidence. Leg Med (Tokyo) 17 (4):232-238. doi:10.1016/j.legalmed.2015.01.004
16. Hebda LM, Foran DR (2015) Assessing the Utility of Soil DNA Extraction Kits for Increasing DNA Yields and Eliminating PCR Inhibitors from Buried Skeletal Remains. J Forensic Sci 60 (5):1322-1330. doi:10.1111/1556-4029.12878
17. Kemp BM, Monroe C, Smith DG (2006) Repeat silica extraction: a simple technique for the removal of PCR inhibitors from DNA extracts. J Archaeol Sci 33 (12):1680-1689. doi:10.1016/j.jas.2006.02.015
18. Schmedes S, Marshall P, King JL, Budowle B (2013) Effective removal of co-purified inhibitors from extracted DNA samples using synchronous coefficient of drag alteration (SCODA) technology. Int J Leg Med 127 (4):749-755. doi:10.1007/s00414-012-0810-7

19. Faber KL, Person EC, Hudlow WR (2013) PCR inhibitor removal using the NucleoSpin(R) DNA Clean-Up XS kit. Forensic Sci Int Genet 7 (1):209-213. doi:10.1016/j.fsigen.2012.06.013
20. Thompson RE, Duncan G, McCord BR (2014) An Investigation of PCR Inhibition Using Plexor®-Based Quantitative PCR and Short Tandem Repeat Amplification. J Forensic Sci 59 (6):1517-1529
21. Hu Q, Liu Y, Yi S, Huang D (2015) A comparison of four methods for PCR inhibitor removal. Forensic Sci Int Genet 16:94-97. doi:10.1016/j.fsigen.2014.12.001
22. Alaeddini R (2012) Forensic implications of PCR inhibition--A review. Forensic Sci Int Genet 6 (3):297-305. doi:10.1016/j.fsigen.2011.08.006
23. Opel KL, Chung D, McCord BR (2010) A study of PCR inhibition mechanisms using real time PCR. J Forensic Sci 55 (1):25-33. doi:10.1111/j.1556-4029.2009.01245.x
24. Scientific TF (2015) GlobalFiler™ PCR Amplification Kit. Thermo Fisher Scientific, Thermo Fisher Scientific
25. Kraemer M, Prochnow A, Bussmann M, Scherer M, Peist R, Steffen C (2017) Developmental validation of QIAGEN Investigator® 24plex QS Kit and Investigator® 24plex GO! Kit: Two 6-dye multiplex assays for the extended CODIS core loci. Forensic Sci Int Genet 29:9-20
26. Martin P, de Simon LF, Luque G, Farfan MJ, Alonso A (2014) Improving DNA data exchange: validation studies on a single 6 dye STR kit with 24 loci. Forensic Sci Int Genet 13:68-78. doi:10.1016/j.fsigen.2014.07.002
27. Scherer M, König M, Breitbach M, Cornelius S, Bussmann M, Prochnow A, Peist R (2015) Improving sample analysis and interpretation using QIAGEN's latest

Investigator® STR multiplex PCR assays with a novel quality sensor. Forensic Sci Int Genet Suppl Ser 5:e308-e309

28. QIAGEN (2007-2014) AllPrep ® DNA/RNA Micro Handbook. QIAGEN, Hilden, Germany
29. Loreille OM, Diegoli TM, Irwin JA, Coble MD, Parsons TJ (2007) High efficiency DNA extraction from bone by total demineralization. Forensic Sci Int Genet 1 (2):191-195. doi:10.1016/j.fsigen.2007.02.006
30. Corporation LT (2010) PrepFiler Express™ and PrepFiler Express BTA™ Forensic DNA Extraction Kits. Life Technologies Corporation,
31. QIAGEN (2012) QIAamp DNA Investigator Handbook. QIAGEN,
32. Scientific TF (2017) Quantifiler™ HP and Trio DNA Quantification Kits: User Guide. F edn. Thermo Fisher Scientific,
33. QIAGEN (2016) Investigator® 24plex QS Handbook. QIAGEN,
34. Pionzio AM, McCord BR (2014) The effect of internal control sequence and length on the response to PCR inhibition in real-time PCR quantitation. Forensic Sci Int Genet 9:55-60. doi:10.1016/j.fsigen.2013.10.010

CHAPTER III

Comparative tolerance of two massively parallel sequencing systems to common PCR inhibitors[1]

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

[1]Elwick K, Zeng X, King JL, Budowle B, Hughes-Stamm S (2017) Comparative tolerance of two massively parallel sequencing systems to common PCR inhibitors. Int J Leg Med 132:983-995. https://doi.org/10.1007/s00414-017-1693-4

Abstract

Human remains can be severely affected by the environment, and the DNA may be damaged, degraded, and/or inhibited. In this study, a DNA sample (at 1 ng DNA target input in triplicate) was spiked with five concentrations of five inhibitors (humic acid, melanin, hematin, collagen, and calcium) and sequenced with both the HID-Ion AmpliSeq™ Library Kit and ID panel on the Ion PGM™ System and the ForenSeq™ DNA Signature Prep Kit on the MiSeq FGx™. The objective of this study was to compare the baseline tolerance of the two sequencing chemistries and platforms to common inhibitors encountered in human remains recovered from missing person cases.

The two chemistries generally were comparable but not always susceptible to the same inhibitors or at the same capacity. The HID-Ion AmpliSeq™ Library Kit and ID panel and the ForenSeq™ DNA Signature Prep Kit both were susceptible to humic acid, melanin, and collagen; however, the ForenSeq™ kit showed greater inhibition to melanin and collagen than the AmpliSeq™ kit. In contrast, the ForenSeq™ kit was resistant to the effects of hematin and calcium, whereas the AmpliSeq™ kit was highly inhibited by hematin. STRs and SNPs showed the same trend among inhibitors when using the ForenSeq™ kit. Generally, locus read depth, heterozygote allele balance, and the numbers of alleles typed were inversely correlated with increasing inhibitor concentration. The larger STR loci were affected more so by the presence of inhibitors compared to smaller STR amplicons and SNP loci. Additionally, it does not appear that sequence noise is affected by the inhibitors. The noise percentage however does increase as the inhibitor concentration increases, due to the decrease in locus read depth and not likely because of chemistry effects.

Keywords: HID-Ion AmpliSeq™ Library Kit, ForenSeq™ DNA Signature Prep Kit, STR, SNP, PCR inhibitors, Noise

Introduction

More than 40,000 unidentified human remains are stored and waiting to be identified in the United States [1]. Human remains cases can include forensic cases and identification of missing migrants, refugees [2] , and victims of mass disaster [3]. Victim identification can be challenging when the remains are exposed to harsh environmental conditions causing DNA degradation and/or inhibition of downstream typing [4,5]. PCR inhibitors co-extract with the DNA and often interfere with downstream DNA typing success. Inhibitors can either affect *Taq* polymerase efficiency or bind to the DNA. When the *Taq* is affected, generally, the larger loci are lost. However, when inhibitors bind the DNA, alleles may be lost regardless of amplicon size, presumably based on where in the template the inhibitor binds [6]. PCR inhibitors often associated with human remains include humic acid in soil, melanin in hair and skin, hematin in red blood cells, collagen in soft tissue and bone, and calcium in bone [6,7].

STRs are the most common genetic marker used in DNA analysis for the identification of human remains due to their relatively small amplicon size (75 – 450 bp), ability to be multiplexed, and high power of discrimination (PD) [8-13]. However, traditional capillary electrophoresis (CE)–based STR typing is limited by the number of markers which can be multiplexed (typically between 25-30 markers) due to constraints in the number of dye channels and the resolving space in each channel [12]. Single nucleotide polymorphisms (SNPs) are alternative markers for human identification (HID) [14-15], which also can be used to determine bio-ancestry and phenotypic information such as hair, eye, and skin color [11,14]. SNPs are single base changes in the genome and therefore can be contained within amplicons, in theory, as small as 50-60 bp [15]. This feature makes

SNPs suitable for typing highly degraded and challenging samples. To reach an equivalent PD of commercial STR kits, approximately 60 well-balanced SNPs must be analyzed [16,17]. With massively parallel sequencing (MPS), DNA molecules are sequenced in parallel to increase throughput and provide more genetic information by assessing the primary sequence of an amplicon [8,11,18-20]. MPS can provide comprehensive sequence information on conventional STR markers, allow sequencing of the entire mitochondrial genome, and enable simultaneous analysis of different marker systems [21]. With the development of this technology it is now possible to multiplex large numbers of STRs and SNPs, and if desired, both marker systems in one analysis.

The goal of this study was to evaluate in a system approach the tolerance of various known PCR inhibitors commonly encountered in forensic and missing person casework with two MPS sequencing chemistries. Although effects of inhibitors are likely to impact the PCR more so than other aspects of the analytical system, the sample preparation and the sequencing chemistry are intertwined. Therefore, this study sought to determine if samples salted with inhibitors could be typed using MPS systems and to determine if MPS systems (comparatively) were affected negatively due to the presence of an inhibitor in a sample. Studies that describe the performance of HID systems with highly inhibited samples will provide data to improve the utility and robustness of each system and to support the overall validity of MPS. The two chemistries evaluated in this study were the HID-Ion AmpliSeq™ Library kit and ID panel on the Ion PGM™ System and the ForenSeq™ DNA Signature Prep Kit using Primer Mix A on the MiSeq FGx™. The HID-Ion AmpliSeq™ Identity Panel consists of 90 autosomal HID SNPs and 34 upper clade Y-SNPs [22]. The ForenSeq™ DNA Signature Prep Kit (Primer Mix A) incorporates

Amelogenin, 27 autosomal STRs, 24 Y-STRs, 7 X-STRs, and 94 HID SNPs [23]. The effects of each inhibitor were evaluated using locus read depth, allele calls, heterozygote allele balance, loci most refractory to individual inhibitors, and generation of noise.

Materials and Methods

Sample and Inhibitor Preparation

Semen was obtained from a single anonymous donor in accordance with Sam Houston State University Institutional Review Board Guidelines (# 2015-12-26123). The semen sample was selected only because it was a convenient sample within the laboratory with substantial amounts of DNA. The DNA was extracted with the AllPrep® DNA/RNA Micro kit and quantified with Quantifiler® Human DNA Quantification Kit.

The inhibitor concentrations used in this study were based on the PGM and Ion AmpliSeq™ chemistry. The PGM platform and AmpliSeq™ kit was tested to its limits with the inhibitor concentrations and then the ForenSeq™ kit was tested with the same concentrations. Initial inhibitor concentrations were based on previous studies with CE-based systems [6,39-41]. The five concentrations of humic acid, melanin, hematin, collagen, and calcium were listed in Table 3.1. Calcium hydrogen phosphate (100 mM) (Sigma Aldrich, Milwaukee, WI) was prepared in 0.5 N hydrochloric acid (Fisher Scientific, Waltham, MA) in a total volume of 10 mL. Humic acid (1 mg/mL) (Alfa Aesar, Ward Hill, MA) was prepared in deionized water in a total volume of 10 mL. Collagen from calf skin (1 mg/mL) (Sigma Aldrich, St. Louis, MO) was prepared in 0.1 N acetic acid (Fisher Scientific) in a total volume of 10 mL. Hematin (100 mM) (ICN Biomedicals, Aurora, OH) was prepared in 0.1 N sodium hydroxide (Fisher Scientific) in a total volume of 10 mL. Melanin (1 mg/mL) (Sigma) was prepared in 0.5 N ammonium hydroxide

(Fisher Scientific) in a total volume of 10 mL. All subsequent working solutions were prepared with deionized water. All inhibitors were added to the MPS library preparation prior to the initial PCR to achieve the desired final inhibitor concentration (Table 3.1).

Table 3.1 The concentrations of the five PCR inhibitors tested in this study

Inhibitor	Inhibitor Concentration				
	1	**2**	**3**	**4**	**5**
Humic Acid (ng/µL)	5	7	10	17	25
Melanin (ng/µL)	4	5	7	10	12
Hematin (µM)	1	3	5	7	10
Collagen (ng/µL)	180	250	300	350	400
Calcium (µM)	350	500	650	850	1100

Ion PGM™ Sequencing

All sequencing reactions were performed with 1 ng of input DNA. Various concentrations of humic acid, melanin, hematin, collagen, or calcium were added to the DNA (Table 3.1). Each sample was amplified in triplicate using the Ion AmpliSeq™ Library Kit 2.0 and ID panel according to manufacturer's specifications [22]. After amplification, samples were purified using Agencourt® AMPure® XP Reagent (Beckman Coulter, Indianapolis, IN) and ethanol. Following library purification, the samples were quantified using the Ion Library TaqMan® Quantitation Assay (Thermo Fisher Scientific). The libraries with humic acid, melanin, and hematin were approximately 10 pM which were lower in concentration than the desired 20-50 pM input. The libraries were normalized to 10 pM and pooled to 100 µL for loading onto the Ion Chef™ System

(Thermo Fisher Scientific). The libraries with collagen and calcium were diluted to 25 pM and pooled to 100 µL. Pooled libraries were batched according to concentration, added to the Ion Chef™ and loaded onto 316 barcoded semiconductor chips. Sequencing was performed using the Ion PGM™ System. Positive reference samples, and negative controls were included in each sequencing run. Data analysis was conducted using Torrent Suite v4.6, the HID_SNP_Genotyper plugin v4.3.1, and an in-house workbook created at UNTHSC.

MiSeq FGx™ Sequencing

Various concentrations of five inhibitors (humic acid, melanin, hematin, collagen, and calcium) were added to DNA samples (1 ng) (Table 3.1). Each of these samples was amplified in triplicate using ForenSeq™ DNA Signature Prep Kit (using Primer Mix A) according to manufacturer's specifications [38]. Sequencing was performed using the Illumina FGx™ system (10 µL pooled libraries were used). Three sequencing runs were performed. Three reference samples, a positive control and a negative control were included in each sequence run. Data analysis was conducted using STRaitRazor v2s [28] and R software [42].

Results and Discussion

A minimum read depth threshold of 2X was used in this study. The average read depth was calculated for each inhibitor concentration and reference sample (three replicates). Using the Ion PGM platform, Ion AmpliSeq™ library kit, and Identity panel, three reference samples with no inhibitor added resulted in an average locus read depth of 2587X. The cluster densities of the three MiSeq FGx™ runs were 539 k/mm^2, 1312 k/mm^2, and 1397 k/mm^2. The cluster density of the first run was substantially lower than the other

two runs. Therefore, the average read depth of the reference samples was not calculated from the combined three runs. The average read depth of three replicates of the reference sample in the first run was computed separately from the six reference samples in the other two runs. The average STR locus read depth was 900X for the first run, and 790X for the other two runs. However, the average SNP locus read depth was 280X for the first run and 234X for the second and third runs, respectively.

Humic Acid

For the AmpliSeq^{TM} Library Kit and ID panel, compared with the average read depth of the reference samples (2587X), five concentrations of humic acid showed decreasing SNP locus read depth (**Appendix 3.1**). SNP typing success decreased as the concentration of humic acid increased. The percentage of SNPs reported dropped from 99% (without inhibitor) to 9% (17 ng/µL of humic acid) (Fig. 3.1). The SNP success increased slightly from 17 ng/µL to 25 ng/µL (9% to 12%). SNP heterozygote allele balance also decreased as the concentration of the inhibitor increased, with the exception of a slight increase from 5% at 17 ng/µL to 12% at 25 ng/µL (Fig. 3.2). Out of all 124 SNPs, rs873196 was identified as the SNP most resistant to humic acid (Fig. 3.3).

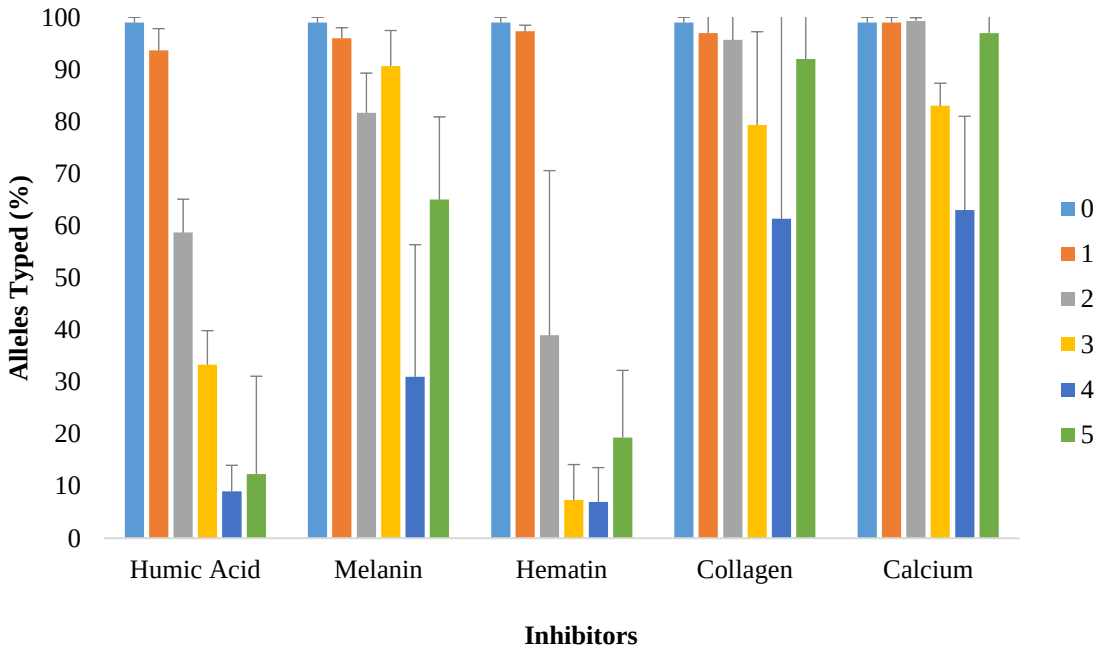


Figure 3.1 Percentage of SNP alleles (Ion Ampliseq™ Library Kit and ID panel) reported with 1 ng of DNA input with five concentrations of five inhibitors. Concentration 0 represents no inhibitor added. The inhibitor concentrations are listed in Table 3.1. Data presented as averaged + standard deviation (N = 3).

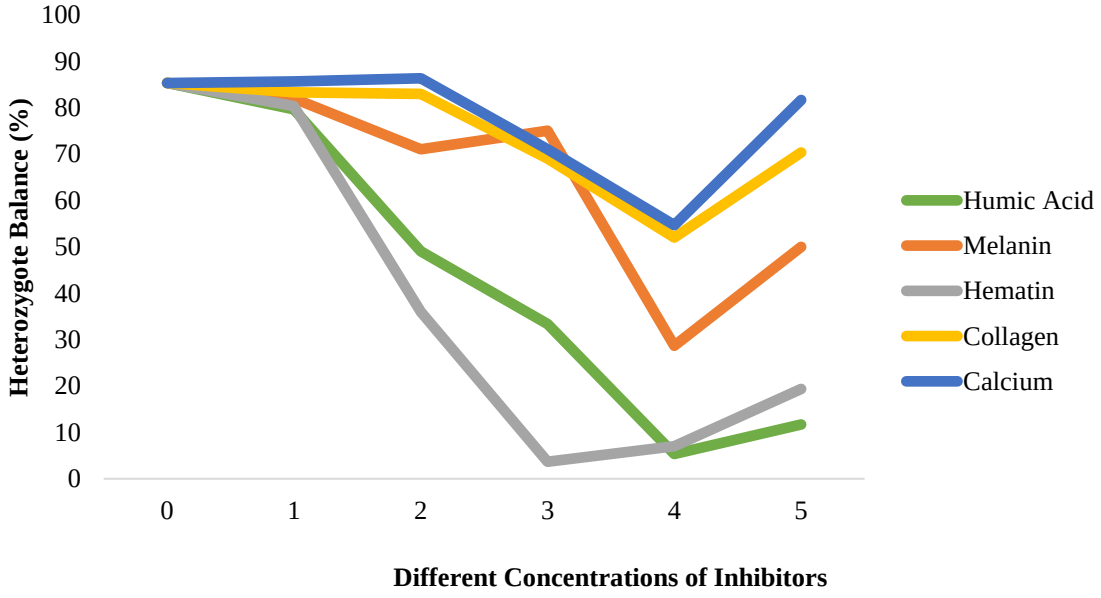


Figure 3.2 Average heterozygote SNP allele balance (Ion Ampliseq™ Library Kit and ID panel) with 1 ng of DNA input with five concentrations of five inhibitors tested. Concentration 0 means no inhibitor added. The inhibitor concentrations are listed in Table 3.1.

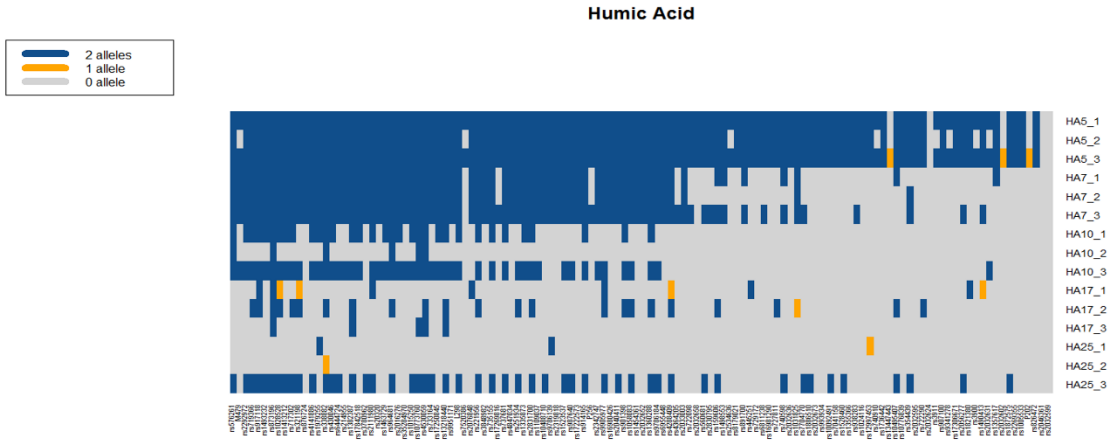


Figure 3.3 The heatmap of alleles typed at each SNP locus with five concentrations of humic acid (Ion Ampliseq™ Library Kit and ID panel).

For the ForenSeq™ DNA Signature Prep Kit, STR loci average read depth was calculated. Compared with the average read depth of three reference samples in the first run (900X), three concentrations of humic acid (5 ng/µL, 7 ng/µL, and 10 ng/µL) had a slightly less STR locus read depth (**Appendix 3.2**). However, at greater concentrations of humic acid (17 ng/µL and 25 ng/µL) the STR loci average read depth decreased to 55X and 2X, respectively. The SNP loci average read depth showed the same trend as that of the STR loci. Humic acid decreased the average read depth substantially to 2X, compared with 280X for the reference samples (**Appendix 3.3**). Consistent with read depth, STR and SNP typing success decreased as the concentration of humic acid increased (Figs. 3.4-3.5) for the ForenSeq™ DNA Signature Prep Kit. The percentages of SNPs typed were higher than that of STRs when humic acid was added (Figs. 3.4-3.5).

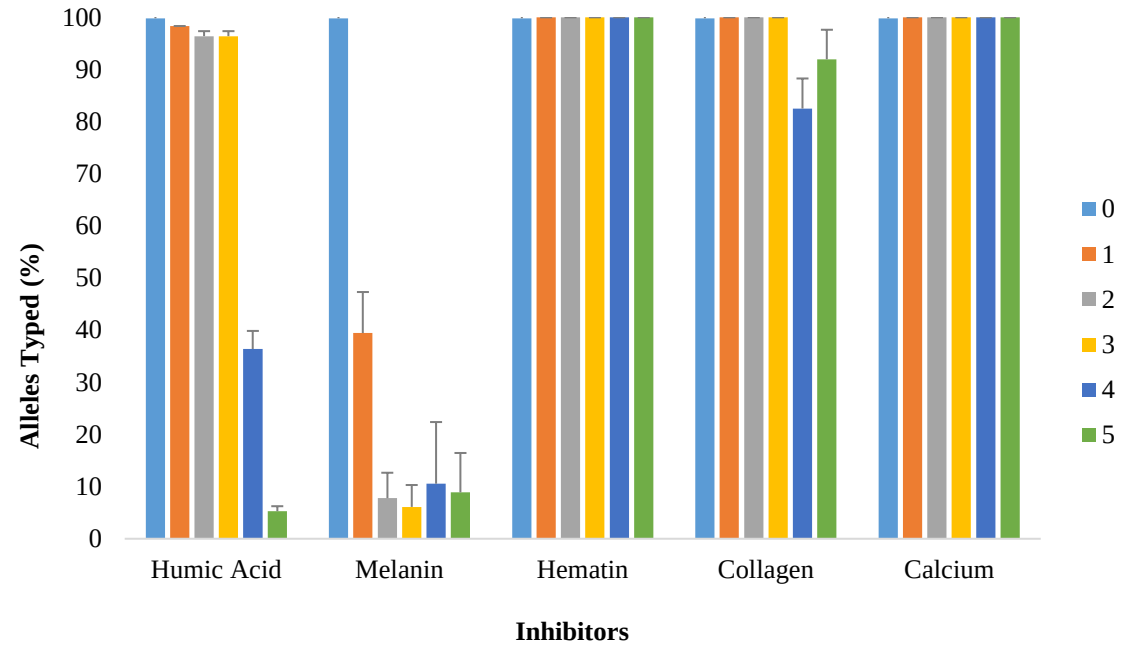


Figure 3.4 Percentage of SNP alleles (ForenSeq™ DNA Signature Prep Kit) reported with 1 ng of DNA input with five concentrations of five inhibitors. Concentration 0 represents no inhibitor added. The inhibitor concentrations are listed in Table 3.1. Data presented as averaged + standard deviation (N = 3).

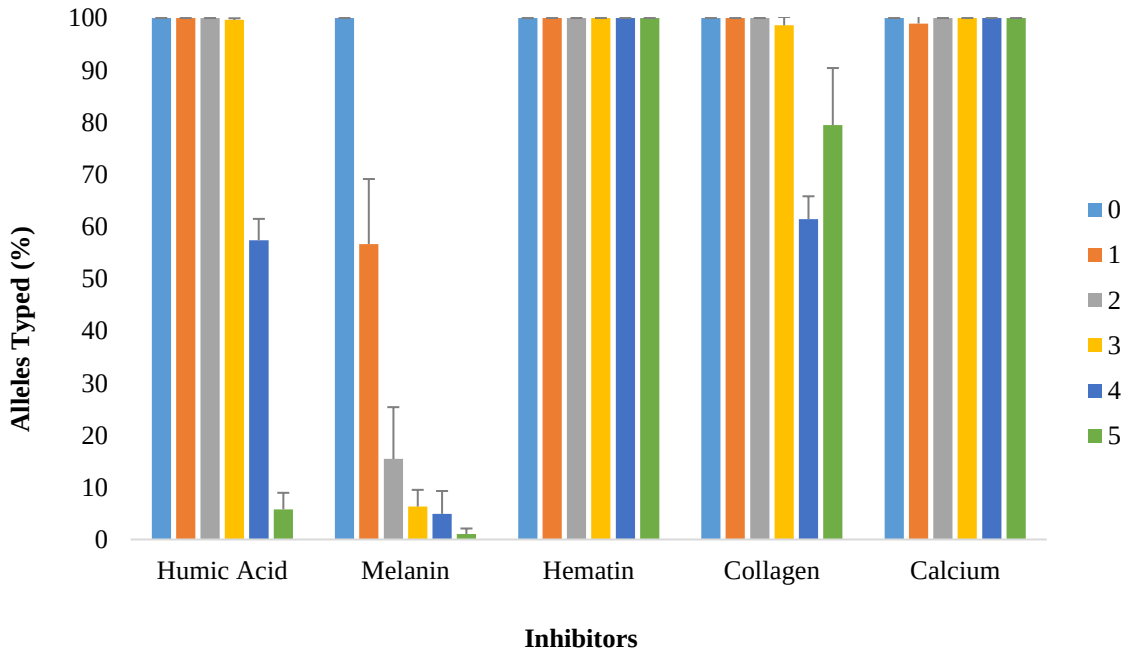


Figure 3.5 Average heterozygote STR allele balance (ForenSeq™ DNA signature Prep Kit) with 1 ng of DNA input with five concentrations with five inhibitors tested.

Concentration 0 means no inhibitor added. The inhibitor concentrations are listed in Table 3.1.

This observation indicates that SNPs in ForenSeq™ DNA Signature Prep Kit may be more resilient in the presence of this inhibitor than STRs. With the increased concentrations of humic acid, one or both alleles of all of the heterozygote STR or SNP loci dropped out, whereby the average heterozygote allele balance decreased to zero (Figs. 3.6-3.7).

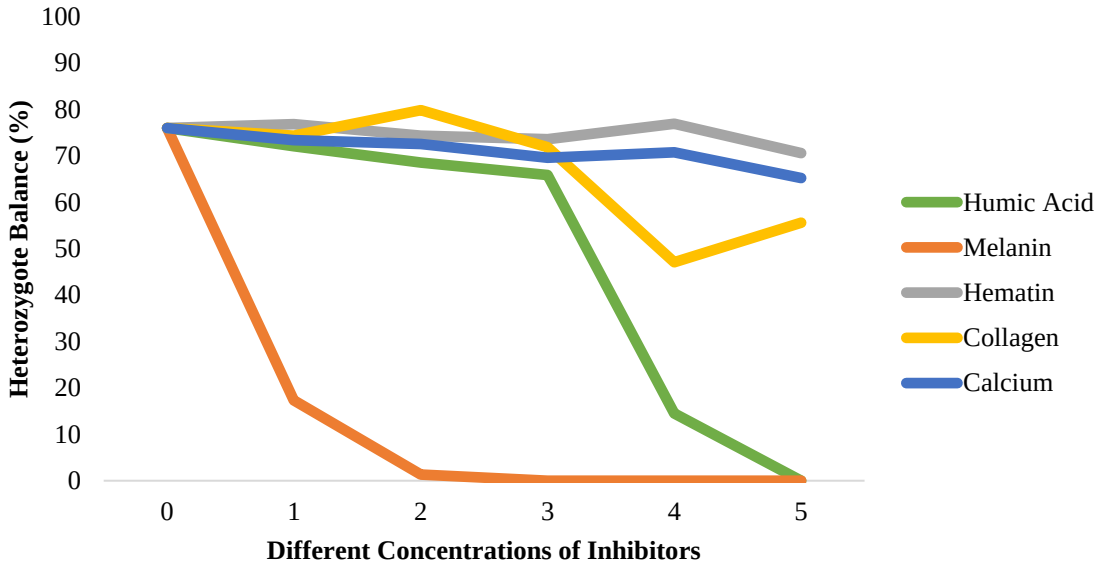


Figure 3.6 Average heterozygote STR allele balance (ForenSeq™ DNA Signature Prep Kit) with 1 ng of DNA input with five concentrations with five inhibitors tested. Concentration 0 means no inhibitor added. The inhibitor concentrations are listed in Table 3.1.

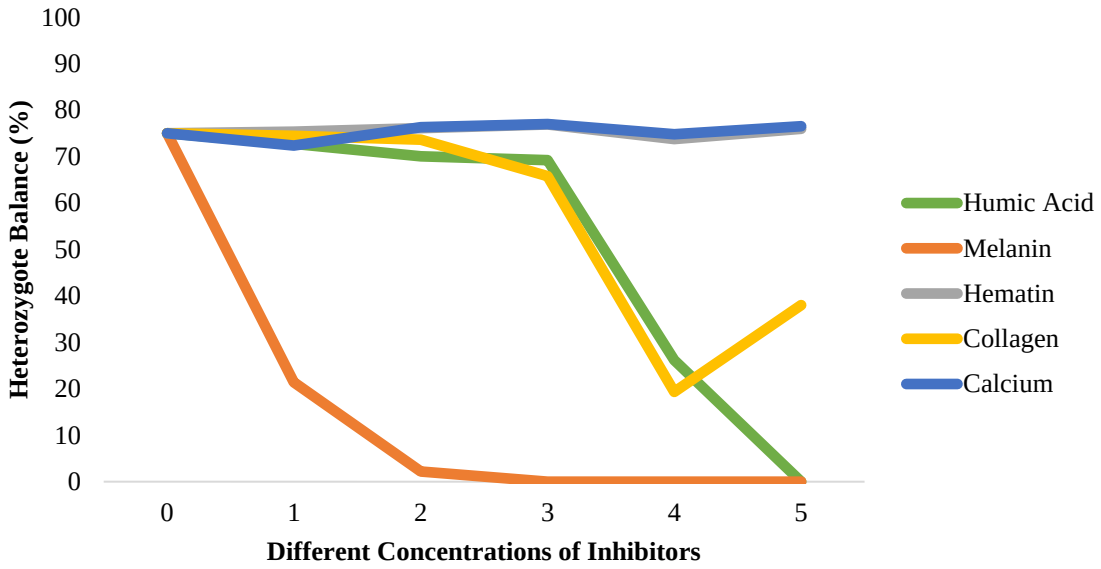


Figure 3.7 Average heterozygote SNP allele balance (ForenSeq™ DNA Signature Prep Kit) with 1 ng of DNA input with five concentrations with five inhibitors tested. Concentration 0 means no inhibitor added. The inhibitor concentrations are listed in Table 3.1.

The STR and SNP loci most refractory to humic acid were identified. The heatmap of alleles typed at each STR locus with five concentrations of humic acid is shown in Fig. 3.8. With the increment of humic acid concentrations (from 5 ng/µL to 25 ng/µL), allele drop out was observed. The largest STR loci (amplicon size ≥ 200 bp) were less resistant to humic acid, and allele drop out was observed starting from 5 ng/µL of humic acid. In contrast, the 15 smallest STR loci were more resistant to humic acid (Fig. 3.8). Forty-three SNP loci were deemed relatively resistant to humic acid because allele drop out was observed only at 25 ng/µL of humic acid (Fig. 3.9). These results are consistent with those of Jäger et al. [43] who observed inhibitory effects at high concentrations of humic acid (i.e., 133.3 µM (30.3 ng/µL)).

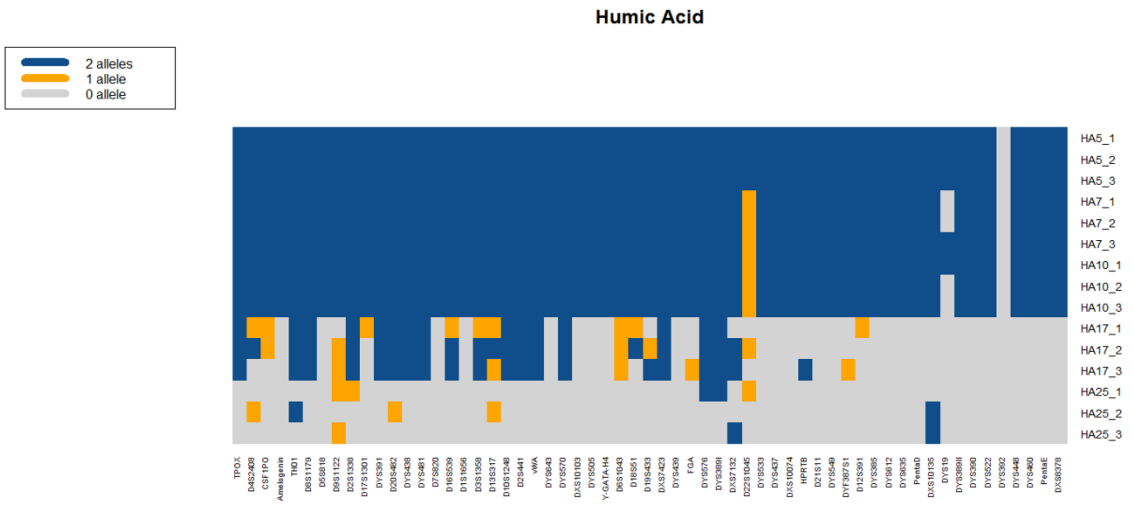


Figure 3.8 The heatmap of alleles typed at each STR locus with five concentrations of humic acid (ForenSeq™ DNA Signature Prep Kit).

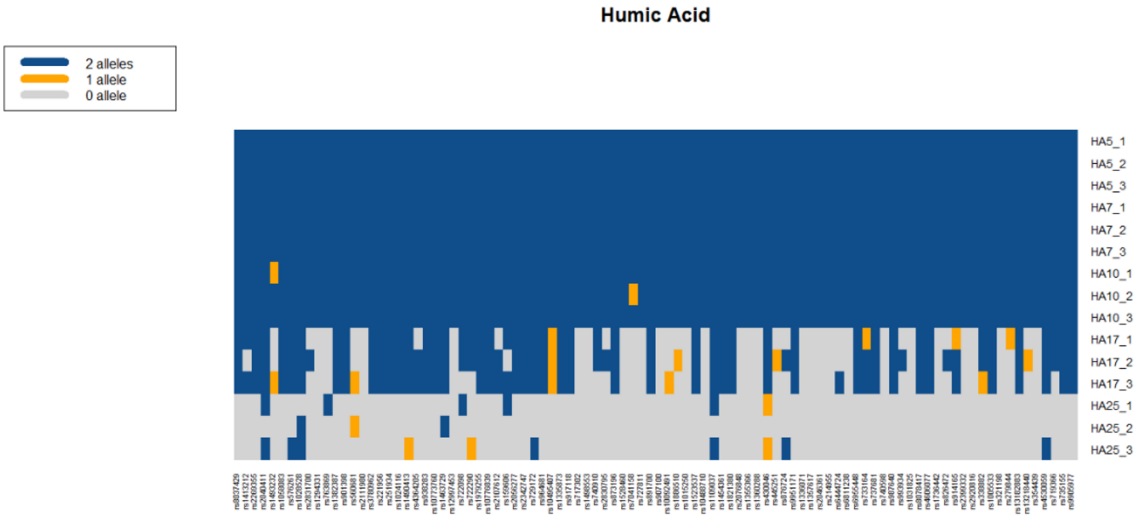


Figure 3.9 The heatmap of alleles typed at each SNP locus with five concentrations of humic acid (ForenSeq™ DNA Signature Prep Kit).

Overall, both the AmpliSeq™ Library Kit and ID panel and the ForenSeq™ DNA Signature Prep Kit were susceptible to inhibition by humic acid. Locus read depth, number of alleles typed, and heterozygote allele balance were inversely correlated with increasing

concentrations of humic acid. The STRs and SNPs with small amplicon sizes were more resistant to the effects of humic acid.

Melanin

For the AmpliSeq™ Library Kit and ID panel, compared with the average read depth of the reference samples (2587X), five concentrations of melanin (Table 3.1) showed decreasing SNP locus read depth (**Appendix 3.1**). SNP typing success also decreased as the concentration of melanin increased (Fig. 3.1). The percentage of SNPs reported decreased from 99% (no inhibitor) to 65% (12 ng/µL), although the percentage of SNPs reported at 10 ng/µL fell to 31% (Fig. 3.1). The heterozygote SNP allele balance was more variable for melanin. The general trend was a decrease in allele balance as the concentration of melanin increased, with the exception of increasing values at two concentrations, 7 ng/µL and 12 ng/µL, of 75% and 50%, respectively (Fig. 3.2). Of all 124 SNPs, there were 14 SNPs resistant to the effects of melanin (**Appendix 3.4**). No allele drop out was observed at the five concentrations of melanin.

With the ForenSeq™ DNA Signature Prep Kit, melanin was a strong inhibitor of the analytical process. The STR loci average read depth of the DNA samples amplified with five concentrations of melanin ranged from 29X to 3X, respectively (**Appendix 3.2**). The SNP loci average read depth showed the same trend as that of the STR loci. Melanin decreased the SNP loci average read depth substantially to 2X compared with 280X for the reference sample (**Appendix. 3.3**). Consistent with read depth, STR and SNP typing success decreased as the concentration of melanin increased. Melanin was a strong inhibitor in which only 39.4% of STR alleles were typed at the lowest concentration of melanin, and less than 10% of alleles were observed at the highest concentration of melanin

(Fig. 4). Only 1% of SNP alleles could be typed at 12 ng/µL of melanin (Fig. 3.5). With increased concentrations of melanin, one or both alleles of all of the heterozygote STR and SNP loci dropped out, whereby the average heterozygote allele balance decreased to zero (Figs. 3.6-3.7). The STR and SNP loci most refractory to each inhibitor were identified. Only the TH01 locus, rs576261, and rs737681 could be detected at 5 ng/µL of melanin in the three replicates (**Appendix 3.5-3.6**).

The markers in the AmpliSeq™ Library Kit and ID panel and the ForenSeq™ DNA Signature Prep Kit were susceptible to the effects of melanin. Locus read depth, number of alleles typed, and heterozygote allele balance generally were inversely correlated with increasing concentrations of melanin. However, melanin showed stronger inhibitor effects with the ForenSeq™ DNA Signature Prep Kit.

Hematin

For the AmpliSeq™ Library Kit and ID panel, hematin was a very strong inhibitor at all concentrations higher than 1 µM (the lowest concentration). Compared with the average SNP locus read depth of the reference samples (2587X), four concentrations of hematin (Table 3.1) showed rapidly decreasing SNP locus read depth; the only exception was at 1 µM (3099X) (**Appendix 3.1**). SNP typing success also decreased as the concentration of hematin increased. The percentage of SNPs that were typed dropped considerably from 99% (without inhibitor) to 19% at 10 µM, with the exception of two concentrations (5 µM and 7 µM) reporting 7% of alleles (Fig. 3.1). Heterozygote allele balance decreased rapidly followed by a slight increase at the two highest hematin concentrations (7 µM and 10 µM) (Fig. 3.2). No SNPs could be detected at all three replicates with ≥ 5 µM of hematin. There were 18 SNPs (with small amplicon size)

relatively resistant to the effects of hematin compared with other markers (**Appendix. 3.7**). For these 18 markers, full profiles were observed at all three replicates with 1 µM and 3 µM of hematin.

In contrast, using the MiSeq FGx™, the results indicated that STR typing with the ForenSeq™ DNA Signature Prep Kit was not affected by hematin. The STR and SNP loci average read depth of the DNA samples that included hematin were higher than the reference samples (**Appendix 3.2-3.3**). Hematin had no influence on STR and SNP alleles typing; all alleles were observed (Figs. 3.4-3.5). In addition, STR and SNP heterozygote allele balance was not affected by the concentrations of hematin tested herein (Figs. 3.6-3.7). All STRs and SNPs in ForenSeq™ DNA Signature Prep Kit were resistant to hematin (**Appendix 3.8-3.9**). While Jäger et al. [43] also tested the effects of hematin, the results herein cannot be compared because the concentration ranges did not overlap between the two studies.

Overall, hematin is a strong inhibitor to Ion Ampliseq™ Library Kit and ID panel. Locus read depth, number of alleles typed, and heterozygote allele balance generally were inversely correlated with increasing concentrations of hematin. However, the ForenSeq™ DNA Signature Prep Kit was resistant to the effects of hematin at all concentrations tested in this study. Currently, there is no explanation for the enhancing effect of hematin on the read depth of ForenSeq™ DNA Signature Prep Kit.

Collagen

For the AmpliSeq™ Library Kit and ID panel, the presence of collagen had a slight effect on average SNP locus read depth (**Appendix 3.1**). Compared with the average read depth of the reference samples (2587X), with increasing concentrations of collagen read

depth decreased gradually (**Appendix 3.1**). The SNP typing success decreased slightly as the concentration increased, but samples were found to be more tolerable to collagen than the other inhibitors tested in this study (Fig. 3.1). The percentage of SNPs reported dropped from 99% (without inhibitor) to 92% at the highest inhibitor concentration, with the exception of 300 ng/µL and 350 ng/µL (79% and 61%) (Fig. 3.1). Heterozygote allele balance did decrease gradually as the concentration of collagen increased, with the exception of an increase with the final inhibitor concentration (400 ng/µL) (Fig. 3.2). Of all 124 SNPs, there were 15 SNPs that were resistant to the effects of collagen (**Appendix 3.10**). They were detected at all five concentrations of collagen.

Using the ForenSeq™ DNA Signature Prep Kit, collagen was found to negatively impact STR loci average read depth which dropped from 790X (reference samples) to 11X (350 ng/µL of collagen) and 30X (400 ng/µL of collagen) (**Appendix 3.2**). The SNP loci average read depth showed the same trend as that of the STR loci (**Appendix 3.3**). DNA samples with three concentrations of collagen (180 ng/µL, 250 ng/µL, and 300 ng/µL) generated full STR profiles (Fig. 3.4). For the higher concentrations of collagen (350 ng/µL and 400 ng/µL), 82.5% and 91.9% of STR alleles were typed, respectively. The corresponding STR loci average read depth were 11X with 350 ng/µL of collagen and 30X with 400 ng/µL of collagen (**Appendix 3.2**). The percentages of SNP alleles reported had the same trend as observed for the STRs (Fig. 3.5). SNP typing success decreased as the concentration of collagen increased. SNP drop out was observed starting at 300 ng/µL collagen. The 400 ng/µL of collagen sample generated more SNP alleles (79%) than the 350 ng/µL of collagen (61%). The first three concentrations had no apparent influence on STR heterozygote allele balance, but the greater concentrations decreased STR

heterozygote allele balance (Fig. 3.6). SNP heterozygote balance ratios were decreased substantially (Fig. 3.7). There were 30 STRs and 32 SNPs resistant to the five concentrations of collagen (**Appendix 3.11-3.12**).

Overall the markers in both kits showed some susceptibility to the presence of collagen, but the effects were greater in the ForenSeq™ DNA Signature Prep Kit. Locus read depth, number of alleles typed, and heterozygote allele balance generally were inversely correlated with increasing concentrations of collagen.

Calcium

For the AmpliSeq™ Library Kit and ID panel, samples with calcium showed a gradual increase in SNP read depth followed by a severe decrease as the concentration increased to more than 500 µM of inhibitor. Compared with the average read depth of the reference samples (2587X), five concentrations of calcium showed decreased SNP locus read depth after the first two concentrations increased (**Appendix 3.1**). Like collagen, the system was more tolerant to samples inhibited with calcium than with the other inhibitors tested in this study. The reportable number of alleles decreased as the inhibitor concentration increased, with the exception of the final inhibitor, which increased considerably (from 63% to 97%) (Fig.3.1). The number of alleles reported for the first two concentrations remained at the same level as the reference (99%). At 650 µM and 850 µM, 83% and 63% of alleles were reported, respectively. Heterozygote allele balance decreased gradually as the concentration of calcium increased, with the exception of an increase at the final concentration (Fig. 3.2). Of all 124 SNPs, 52 SNPs could be detected at all five concentrations of calcium (**Appendix 3.13**). They were the most refractory markers to calcium.

In contrast, the results of STR and SNP typing in ForenSeq™ DNA Signature Prep Kit were not affected by calcium at the concentrations tested in this study. The STR and SNP loci average read depth of the DNA samples that included calcium were higher than the reference samples except at 350 µM of calcium (**Appendix 3.2-3.3**). Calcium had no influence on STR typing, and all alleles were observed (Fig. 3.4). However, two SNP loci (rs1736442 and rs1031825) were not detected in one of the three replicates at the first concentration (350 µM) (Fig. 3.5). These two SNPs are low performers in ForenSeq™ DNA Signature Prep Kit (36). Calcium did not affect heterozygote peak height ratios (Figs. 3.6-3.7). All STRs were resistant to the effects of calcium (**Appendix 3.14**). All SNPs generally were refractory to the concentrations of calcium used in this study (**Appendix 3.15**).

In summation, the AmpliSeq™ Library Kit and ID panel was susceptible to the presence of calcium. Locus read depth, number of alleles typed, and heterozygote allele balance generally were inversely correlated with increasing concentrations of calcium. However, the ForenSeq™ DNA Signature Prep Kit was resistant to calcium concentrations tested in this study.

Noise Assessment

The influence of five inhibitors on noise generated during SNP sequencing with the AmpliSeq™ Library Kit and ID panel was investigated. Overall noise (i.e., PCR/sequence error) did not appear to increase with exposure to inhibitors. However, percent noise did increase with increasing concentrations of inhibitors as a result of a decrease in read depth of the true allele sequence (note that the true allele sequence is based on the major representative allele(s) of the non-inhibitor reference sample) (**Appendix 3.16-3.19**). For

example, the noise percentage of rs338882 was 22.2% in one of the three replicates that included the highest concentration of humic acid; but the noise reads were only 2X with a locus read depth of 9X (Fig. 3.10).

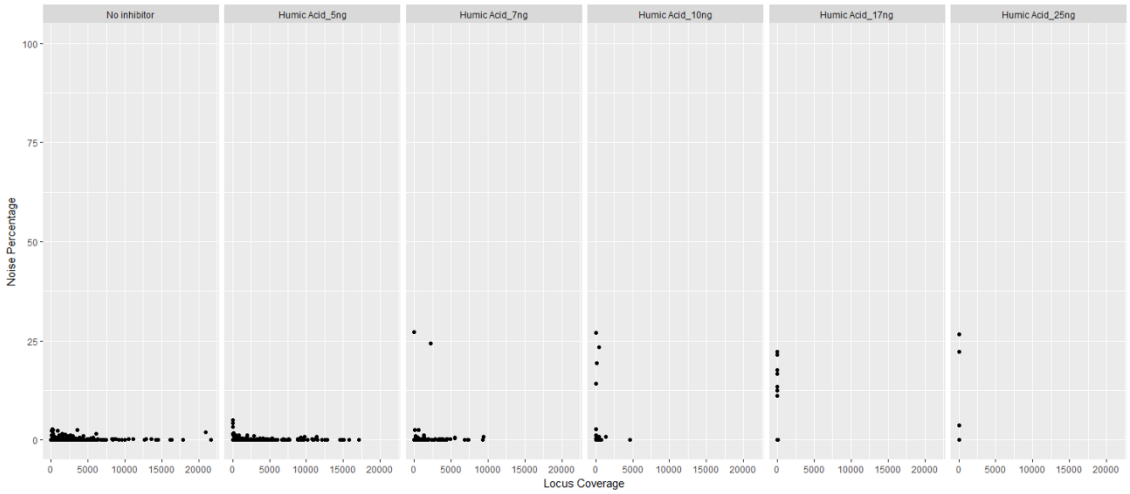


Figure 3.10 The noise percentages of SNPs (Ion Ampliseq^{TM} Library Kit and ID panel) of reference samples and DNA samples spiked with five concentrations of humic acid. X axis is locus coverage, Y axis is noise percentage.

The influence of five inhibitors on the STRs sequence noise (ForenSeq^{TM} DNA Signature Prep Kit) was examined. In this study, for each sample, only the homozygous loci and heterozygous loci that two alleles have at least four repeats difference were used. The sequence noise was divided into three categories for analysis purpose: noise at allele position, noise at -1 repeat position, and artifact [44]. In this study, stutter and sequence noise at -2 repeat and +1 repeat positions were combined into stutter. In this study, all three sequence noise categories were combined for investigating the influence of inhibitors on sequence noise (**Appendix 3.20-3.24**). In addition, the influence of five inhibitors on the SNPs sequence noise (ForenSeq^{TM} DNA Signature Prep Kit) was investigated (**Appendix 3.25-3.29**). The same trends were observed regarding noise as with the AmpliSeq™ kit.

Percent noise does not appear to be affected by the presence of inhibitors in this study; and the percent noise increased as a result of decreasing locus read depth.

Capillary Electrophoresis STR Inhibition Study

The effects of the same five inhibitors (humic acid, melanin, hematin, collagen, and calcium) on the STR testing by CE were conducted by Elwick et al (paper submitted). The performances of two STR kits (GlobalFiler PCR Amplification kit and Investigator® 24plex QS kit) were evaluated. Therefore, a comparison of results from CE data was not reported herein.

Conclusions

The two multiplexes with different chemistries were exposed to DNA samples containing a number of inhibitors over a range of concentrations. As expected, increasing concentrations of inhibitors had an inverse effect on locus read depth and typing success, with a few exceptions. The most noted outcome was that the two kits were not always susceptible to the effects of inhibitors in a similar fashion. For example, the Ion AmpliSeq™ panel was more susceptible to the presence of hematin and calcium with little or no effect observed for the ForenSeq panel. In contrast, the Forenseq panel was more susceptible to melanin and collagen compared with the Ion AmpliSeq™ kit. Possible explanations for different performance between the kits and platforms may be due to PCR conditions or library and/or sequencing chemistry differences. Overall, large amplicon STR loci were less resistant to inhibitors compared with small STR and especially SNP loci. In all cases in which STR or SNP results were obtained, the correct result was obtained. When one allele dropout occurred for heterozygotes, the allele that was observed was consistent with one of the alleles observed in the reference sample. In some of the

increasing inhibitor concentrations, there were differences within a series in which read depth and/or typing success (although slight) were not always consistent with the trend. These differences are likely due to run-to-run variations in MPS and/or stochastic effects. As others attempt to replicate this work, in part or in total, the actual cause of these slight fluctuations may be better elucidated. When there is inhibition, the effective template may not be the same as the estimated input and thus stochastic effects may be more pronounced. The overall outcome was that inhibitors, when they do have a negative effect on typing performance, can reduce typing success but do not contribute to sequencing error.

Acknowledgements

The authors would like to thank the individuals who provided the biological sample used in this study. This project was supported in full by National Institute of Justice (NIJ) Award Number 2015-DN-BX-K066 awarded to Sam Houston State University (SHSU) and the University of North Texas Health Science Center (UNTHSC).

References

1. Ritter N (2007) Missing persons and unidentified remains: The nation's silent mass disaster. NIJ Journal 256:2-7
2. Kovras I, Robins S (2016) Death as the border: Managing missing migrants and unidentified bodies at the EU's Mediterranean frontier. Political Geography 55:40-49
3. Dhanardhono T, Wulandari N, Bhima SKL, Ahmad HJ, Widodo PT (2013) DNA profiling of disaster victim identification in Trenggalek shipwreck case. Forensic Sci Int Genet Suppl Ser 4:e5-e6
4. Barbisin M, Fang R, O'Shea CE, Calandro LM, Furtado MR, Shewale JG (2009) Developmental Validation of the Quantifiler® Duo DNA Quantification Kit for Simultaneous Quantification of Total Human and Human Male DNA and Detection of PCR Inhibitors in Biological Samples. J Forensic Sci 54:305-319
5. Mulero JJ, Chang CW, Lagace RE, Wang DY, Bas JL, McMahon TP, Hennessy LK (2008) Development and validation of the AmpFℓSTR® MiniFilerTM PCR Amplification Kit: a miniSTR multiplex for the analysis of degraded and/or PCR inhibited DNA. J Forensic Sci 53:838-852
6. Thompson RE, Duncan G, McCord BR (2014) An Investigation of PCR Inhibition Using Plexor®-Based Quantitative PCR and Short Tandem Repeat Amplification. J Forensic Sci 59:1517-1529
7. Opel KL, Chung D, McCord BR (2010) A Study of PCR Inhibition Mechanisms Using Real Time PCR. J Forensic Sci 55:25-33
8. Collins PJ, Hennessy LK, Leibelt CS, Roby RK, Reeder DJ, Foxall PA (2004) Developmental validation of a single-tube amplification of the 13 CODIS STR loci,

D2S1338, D19S433, and amelogenin: the AmpFlSTR Identifiler PCR Amplification Kit. J Forensic Sci 49:1265–1277

9. Krenke BE, Tereba A, Anderson SJ, Buel E, Culhane S, Finis CJ, Tomsey CS, Zachetti JM, Masibay A, Rabbach DR, Amiott EA, Sprecher CJ (2002) Validation of a 16-locus fluorescent multiplex system. J Forensic Sci 47:773–785
10. Martín P, de Simón LF, Luque G, Farfán MJ, Alonso A (2014) Improving DNA data exchange: validation studies on a single 6 dye STR kit with 24 loci. Forensic Sci Int Genet 13C:68–78
11. Oostdik K, Lenz K, Nye J, Schelling K, Yet D, Bruski S, Strong J, Buchanan C, Sutton J, Linner J, Frazier N, Young H, Matthies L, Sage A, Hahn J, Wells R, Williams N, Price M, Koehler J, Staples M, Swango KL, Hill C, Oyerly K, Duke W, Katzilierakis L, Ensenberger MG, Bourdeau JM, Sprecher CJ, Krenke B, Storts DR (2014) Developmental validation of the PowerPlex((R)) Fusion System for analysis of casework and reference samples: a 24- locus multiplex for new database standards. Forensic Sci Int Genet 12:69–76
12. Lazaruk K, Walsh PS, Oaks F, Gilbert D, Rosenblum BB, Menchen S, Scheibler D, Wenz HM, Holt C, Wallin J (1998) Genotyping of forensic short tandem repeat (STR) systems based on sizing precision in a capillary electrophoresis instrument. Electrophoresis 19:86–93
13. Butler JM, Buel E, Crivellente F, McCord BR (2004) Forensic DNA typing by capillary electrophoresis using the ABI Prism 310 and 3100 genetic analyzers for STR analysis. Electrophoresis 25:1397–1412

14. Sanchez JJ, Phillips C, Børsting C, Balogh K, Bogus M, Fondevila M, Harrison CD, Musgrave-Brown E, Salas A, Syndercombe-Court D, Schneider PM, Carracedo A, Morling N (2006) A multiplex assay with 52 single nucleotide polymorphisms for human identification. Electrophoresis 27:1713–1724
15. Kidd KK, Pakstis AJ, Speed WC, Grigorenko EL, Kajuna SL, Karoma NJ, Kungulilo S, Kim JJ, Lu RB, Odunsi A, Okonofua F, Parnas J, Schulz LO, Zhukova OV, Kidd JR (2006) Developing a SNP panel for forensic identification of individuals. Forensic Sci Int Genet 164:20–32
16. Yang N, Li H, Criswell LA, Gregersen PK, Alarcon-Riquelme ME, Kittles R, Shigeta R, Silva G, Patel PI, Belmont JW, Seldin MF (2005) Examination of ancestry and ethnic affiliation using highly informative diallelic DNA markers: application to diverse and admixed populations and implications for clinical epidemiology and forensic medicine. Hum Genet 118:382–392
17. Kayser M, Schneider PM (2009) DNA-based prediction of human externally visible characteristics in forensics: motivations, scientific challenges, and ethical considerations. Forensic Sci Int Genet 3:154-161
18. Spichenok O, Budimlija ZM, Mitchell AA, Jenny A, Kovacevic L, Marjanovic D, Caragine T, Prinz M, Wurmbach E (2011) Prediction of eye and skin color in diverse populations using seven SNPs. Forensic Sci Int Genet 5:472-478
19. Walsh S, Liu F, Wollstein A, Kovatsi L, Ralf A, Kosiniak-Kamysz A, Branicki W, Kayser M (2013) The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. Forensic Sci Int Genet 7:98-115

20. Chakraborty R, Stivers DN, Su B, Zhong Y, Budowle B (1999) The utility of short tandem repeat loci beyond human identification: implications for development of new DNA typing systems. Electrophoresis 20:1682-1696
21. Zeng X, King JL, Stoljarova M, Warshauer DH, LaRue BL, Sajantila A, Patel J, Storts DR, Budowle B (2015) High sensitivity multiplex short tandem repeat loci analyses with massively parallel sequencing. Forensic Sci Int Genet 16:38-47
22. Zeng X, King J, Hermanson S, Patel J, Storts DR, Budowle B (2015) An evaluation of the PowerSeq™ Auto System: A multiplex short tandem repeat marker kit compatible with massively parallel sequencing. Forensic Sci Int Genet 19:172-179
23. Fordyce SL, Avila-Arcos MC, Rockenbauer E, Borsting C, Frank-Hansen R, Petersen FT, Willerslev E, Hansen AJ, Morling N, Gilbert MT (2011) High-throughput sequencing of core STR loci for forensic genetic investigations using the Roche Genome Sequencer FLX platform. BioTechniques 51:127–133
24. Fordyce SL, Mogensen HS, Børsting C, Lagacé RE, Chang CW, Rajagopalan N, Morling N (2015) Second-generation sequencing of forensic STRs using the Ion Torrent™ HID STR 10-plex and the Ion PGM™. Forensic Sci Int Genet 14:132-140
25. Friis SL, Buchard A, Rockenbauer E, Borsting C, Morling N (2016) Introduction of the Python script STRinNGS for analysis of STR regions in FASTQ or BAM files and expansion of the Danish STR sequence database to 11 STRs. Forensic Sci Int Genet 21:68–75
26. Hoogenboom J, van der Gaag KJ, de Leeuw RH, Sijen T, de Knijff P, Laros JF (2017) FDSTools: a software package for analysis of massively parallel sequencing data with

the ability to recognise and correct STR stutter and other PCR or sequencing noise. Forensic Sci Int Genet 27:27–40

27. Mikkelsen M, Hansen RF, Hansen AJ, Morling N (2014) Massively parallel pyrosequencing 454 methodology of the mitochondrial genome in forensic genetics. Forensic Sci Int Genet 12:30–37
28. King JL, Wendt FR, Sun J, Budowle B (2017) STRait Razor v2s: Advancing sequence-based STR allele reporting and beyond to other marker systems. Forensic Sci Int Genet 29:21-28
29. Novroski NM, King JL, Churchill JD, Seah LH, Budowle B (2016) Characterization of genetic sequence variation of 58 STR loci in four major population groups. Forensic Sci Int Genet 25:214-226
30. Van Neste C, Vandewoestyne M, Van Criekinge W, Deforce D, Van Nieuwerburgh F (2014) My-Forensic-Loci-queries (MyFLq) framework for analysis of forensic STR data generated by massive parallel sequencing. Forensic Sci Int Genet 9:1–8
31. Parson W, Strobl C, Huber G, Zimmermann B, Gomes SM, Souto L, Fendt L, Delport R, Langit R, Wootton S, Lagacé R, Irwin J (2013) Evaluation of next generation mtGenome sequencing using the ion torrent personal genome machine (PGM). Forensic Sci Int Genet 7:543–549
32. King JL, LaRue BL, Novroski NM, Stoljarova M, Seo SB, Zeng X, Warshauer DH, Davis CP, Parson W, Sajantila A, Budowle B (2014) High quality and high-throughput massively parallel sequencing of the human mitochondrial genome using the Illumina MiSeq. Forensic Sci Int Genet 12:128–135

33. Seo SB, Zeng X, King JL, Larue BL, Assidi M, Al-Qahtani MH, Sajantila A, Budowle B (2015) Underlying data for sequencing the mitochondrial genome with the massively parallel sequencing platform Ion Torrent™ PGM™. BMC Genomics 16 (Suppl 1):S4
34. Eduardoff M, Santos C, de la Puente M, Gross M, Strobl C, Sobrino B, Ballard D, Schneider PM, Carracedo Á, Lareu MV, Parson W, Phillips C (2015) Inter-laboratory evaluation of SNP-based forensic identification by massively parallel sequencing using the Ion PGM. Forensic Sci Int Genet 17:110–121
35. Churchill JD, Chang J, Ge J, Rajagopalan N, Lagacé R, Liao W, King JL, Budowle B (2015) Blind study evaluation illustrates utility of the Ion PGM™ System for use in human identity DNA typing. Croat Med J 56:218-229
36. Churchill JD, Schmedes SE, King JL, Budowle B (2015) Evaluation of the Illumina® beta version ForenSeq™ DNA Signature Prep Kit for use in genetic profiling. Forensic Sci Int Genet 20:20-29
37. Ion AmpliSeq™ Library Preparation for Human Identification Applications (2015) https://tools.thermofisher.com/content/sfs/manuals/MAN0010640_AmpliSeqLibraryPrep_HID_UG.pdf
38. ForenSeq™ DNA Signature Prep Reference Guide (2015) https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/forenseq/forenseq-dna-signature-prep-guide-15049528-01.pdf
39. Pionzio AM, McCord BR (2014) The effect of internal control sequence and length on the response to PCR inhibition in real-time PCR quantitation. Forensic Sci Int Genet 9:55-60

40. Guo F, Shen H, Tian H, Jin P, Jiang X (2014) Development of a 24-locus multiplex system to incorporate the core loci in the Combined DNA Index System (CODIS) and the European Standard Set (ESS). Forensic Sci Int Genet 8:44-54
41. Zhou H, Wu D, Chen R, Xu Y, Xia Z, Guo Y, Zhang F, Zheng W (2014) Developmental validation of a forensic rapid DNA-STR kit: Expressmarker 16. Forensic Sci Int Genet 11:31-38
42. R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
43. Jäger AC, Alvarez ML, Davis CP, Guzmán E, Han Y, Way L, Walichiewicz P, Silva D, Pham N, Caves G, Bruand J, Schlesinger F, Pond SJ, Varlaro J, Stephens KM, Holt CL (2017) Developmental validation of the MiSeq FGx Forensic Genomics System for Targeted Next Generation Sequencing in Forensic DNA Casework and Database Laboratories. Forensic Sci Int Genet 28:52-70
44. Zeng X, King JL, Budowle B (2017) Investigation of the STR loci noise distribution of PowerSeq™ Auto System. Croat Med J 58:214-221

CHAPTER IV

Assessment of impact of DNA extraction methods on analysis of human remain samples on massively parallel sequencing success[1]

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

[1]Zeng X, Elwick K, Mayes C, King JL, Takahashi M, Gangitano D, Budowle B, Hughes-Stamm S (2018) Assessment of impact of DNA extraction methods on analysis of human remain samples on massively parallel sequencing success. Int J Leg Med https://doi.org/10.1007/s00414-018-1955-9

Abstract

Skeletal remains recovered from missing persons' cases are often exposed to harsh environmental conditions resulting in the DNA being damaged, degraded, and/or the samples containing PCR inhibitors. In this study, the efficacy of common extraction methods was evaluated to remove high levels of PCR inhibitors commonly encountered with human remains, and their downstream compatibility with the two leading sequencing chemistries and platforms for human identification purposes.

Blood, hair, and bone samples were spiked with high levels of inhibitors commonly identified in each particular substrate in order to test the efficiency of various DNA extraction methods prior to sequencing. Samples were extracted using three commercial extraction kits (DNA IQ, DNA Investigator, and PrepFiler BTA), organic (blood and hair only), and two total demineralization protocols (bone only). Massively parallel sequencing (MPS) was performed using two different systems: Precision ID chemistry and an early access degradation panel on the Ion S5™ System and the ForenSeq™ DNA Signature Prep Kit on the MiSeq FGx™.

The overall results showed that all DNA extraction methods were efficient and are fully compatible with both MPS systems. Key performance indicators such as STR and SNP reportable alleles, read depth, and heterozygote balance were comparable for each extraction method. In samples where CE-based STRs yielded partial profiles (bone), MPS-based STRs generated more complete or full profiles. Moreover, MPS panels contain more STR loci than current CE-based STR kits and also include SNPs, which can further increase the power of discrimination obtained from these samples, making MPS a desirable choice for the forensic analysis of such challenging samples.

Keywords: Massively Parallel Sequencing, Ion S5™, MiSeq FGx™, PCR inhibitors, DNA extraction, Missing persons

Introduction

In some forensic and many missing persons' cases bone, teeth, hair, and severely decomposed tissues are the only samples remaining for human identification (HID) purposes [1,2]. However, exposure to harsh environmental conditions such as heat, humidity, burial environment, bacteria and mold, or UV light may cause DNA degradation and damage making these samples challenging to process [1,3,4]. In addition, human remains may also contain PCR inhibitory agents such as collagen, calcium, humic acid, melanin, and hematin [5]. Inhibitors may be co-extracted with the DNA [1,6], which can interfere with PCR and reduce downstream DNA typing success [5,7,8]. The current gold standard in HID is the analysis of short tandem repeats (STRs) via capillary electrophoresis (CE) [9-13]. However, massively parallel sequencing (MPS) offers an ability to analyze challenging forensic samples, sequence the entire mitochondrial genome, determine ancestry, provide phenotypic information, and better resolve DNA mixtures [14-16]. MPS has the ability to expand our current capabilities as more genetic information can be retrieved from each sample via the simultaneous analysis of different (and more) markers (e.g. STRs [17], identity SNPs (iiSNPs) [17], and ancestry SNPs (aiSNPs)) [14].

An effective DNA extraction method is critical to obtaining as much high-quality DNA (i.e., sufficiently pure for downstream assay) as possible from difficult samples [18,19]. Following extraction, DNA quantification is used to determine DNA quantity as well as the level of inhibition in a sample using the internal PCR control (IPC), but several studies have shown that the IPC is not always the best or only indicator of PCR inhibition [6,20-22]. In addition, relatively small amounts of DNA extract (2 µL) are used for quantification, while often up to 15 µL of extract are amplified for genotyping. Whilst no

PCR inhibition may be detected during quantification, the input of much larger volumes of neat DNA extract will also increase the amount of any inhibitory agents present and may cause PCR inhibition. Therefore, it is necessary to determine if the MPS systems are able to overcome inhibitors as detected by a current quantitation system. STR-CE amplification kits have been reported as being extremely tolerant to forensically relevant inhibitors [23,24], and commonly employed DNA extraction methods have been shown to be highly compatible with these CE-based STR chemistries. However, this work has not been demonstrated with MPS systems, and little is known regarding the compatibility of these DNA extraction methods with MPS chemistries. The goals of this study were to: 1) evaluate the efficiency of various DNA extraction methods to remove high amounts of PCR inhibitors from challenging samples prior to MPS, and 2) compare the quality of STR/SNP analysis using an early access panel for degraded samples with Precision ID chemistry on the Ion S5™, the ForenSeq™ DNA Signature Prep Kit (using Primer Mix A) on the MiSeq FGx™, and traditional CE-based STR typing for the identification of human remains. Blood, hair, and bone samples were collected and spiked with high amounts of the relevant inhibitor (humic acid, melanin, hematin, or calcium). Blood and hair samples were extracted using the three most commonly used commercial kits in forensic laboratories: PrepFiler® BTA (Life Technologies™, Carlsbad, CA), DNA IQ™ (Promega, Madison, WI), and DNA Investigator (QIAGEN, Valencia, CA), and an organic method. Bone samples were extracted using the same three commercial extraction kits and two different total demineralization protocols.

Materials and Methods

Sample and Inhibitor Preparation

Blood and hair samples were obtained from the same live donor in accordance with Sam Houston State University Institutional Review Board Guidelines #2015-12-26123. Bone samples were harvested from a single body willed to the Applied Anatomical Research Center (AARC) at Sam Houston State University in Huntsville, Texas. Bone samples (approximately 5 cm x 3 cm) were sanded and chipped into small pieces using a Dremel® tool, washed with 10% bleach, diH_2O, and 70% ethanol, and powdered using a SPEX CertiPrep 6750 Freezer/Mill Cryogenic Grinder.

A high concentration of each PCR inhibitor was added to the appropriate biological sample prior to DNA extraction: hematin (10 µL) was added to blood samples (15 µL) for a final inhibitor amount of 17420 ng (in 25 µL volume); melanin was added to the hair samples for an inhibitor amount of 750 ng; calcium was added to bone samples for an inhibitor amount of 22.5 mM; and humic acid was added to bone samples for an inhibitor amount of 3750 ng. The inhibitor amounts chosen are represented in Table 4.1. Hematin, melanin, calcium, and humic acid inhibitors were prepared as stated in Elwick et al. [25] and all subsequent working solutions were prepared with deionized water.

Table 4.1 Final Inhibitor amounts spiked in to their respective substrates.

Sample	Substrate Amount	Inhibitor	Inhibitor Amount
Blood	15 µL	Hematin	17420 ng
Hair	1 hair (with root)	Melanin	750 ng
Bone	50 mg	Calcium	22.5 mM
Bone	50 mg	Humic Acid	3750 ng

Three spiked replicates and one control (no inhibitor) of each of the blood and hair samples were subjected to four different extraction methods: an organic extraction method [26], DNA IQ™ [27-29], PrepFiler® BTA [30], and QIAamp® DNA Investigator [31] all following recommended protocols. In addition to the commercial kits, bone samples also underwent total demineralization (TD) protocols, using TD1 [32] and TD2 [33], as previously described. The extracted DNA was quantified with the Quantifiler® Trio DNA Quantification Kit (Thermo Fisher Scientific, Waltham, MA) on a 7500 Real-Time PCR System (Thermo Fisher Scientific).

Capillary Electrophoresis-based STR genotyping

CE-based STR typing was performed using the GlobalFiler® PCR Amplification Kit (Thermo Fisher Scientific) as per the manufacturer's instructions [34]. All samples were amplified using a target of 0.8 ng of DNA. STR genotyping was performed using the 3500 Genetic Analyzer with a 36 cm capillary array and POP-4 (Thermo Fisher Scientific). Data were analyzed using GeneMapper ID-X v. 1.4 and an in-house workbook. Alleles were assigned using an analytical threshold of 150 RFUs and a stochastic threshold of 600 RFUs.

Ion S5™ Library Preparation and Sequencing

DNA (1 ng) from blood, hair, and bone samples was amplified using the Precision ID DL8 Kit and an early access degradation primer panel (Thermo Fisher Scientific) on the Ion Chef™ System. This panel includes 33 STR markers, 1 Y-STR, 1 Y-indel, amelogenin, 41 iiSNPs, and 34 Y-SNPs. All 16 samples were sequenced in the same run to minimize run-to-run variability. Each DL8 plate amplified 7 samples and one 007 control DNA sample (Thermo Fisher Scientific). Pooled libraries were quantified using the

Ion Library TaqMan® Quantitation Assay (Thermo Fisher Scientific) and diluted to 50 pM for templating. Templating and chip loading were conducted using the Ion Chef™ System with Ion 530™ semiconductor chips. Four sequencing runs were performed using the Ion S5™ System and the Ion S5™ Precision ID Chef and Sequencing Kit (Thermo Fisher Scientific). Positive and negative control samples were also sequenced.

Data analysis was performed using Converge™ 2.0 (Thermo Fisher Scientific) and in-house workbooks. STR and SNP allele typing success was calculated as the percentage of concordant alleles reported. An arbitrary detection threshold was set at 5X. STR read depth was calculated by summing the coverage of all STR loci. STR heterozygote balance was calculated by dividing the read depth of the higher coverage STR allele divided by the read depth of the lower coverage STR allele so that the value would also indicate which allele had a greater read depth. SNP read depth was calculated by summing the coverage of all SNP loci. SNP heterozygote balance was calculated by the read depth of the higher coverage SNP allele divided by the read depth of the lower coverage SNP allele. Because of the number of markers between kits, and cluster density vs loading we cannot make a direct comparison between the two platforms.

MiSeq FGx™ Library Preparation and Sequencing

The same extracted DNA (1 ng) from blood, hair, and bone samples was amplified using the ForenSeq™ DNA Signature Prep Kit (using Primer Mix A) according to the manufacturer's specifications [35]. Sequencing was performed using the MiSeq FGx™ system (10 µL pooled libraries were analyzed). Three sequencing runs were performed. Positive and negative controls were included in each sequencing run. Data analysis was conducted using STRait Razor v2s [36]. The same approach as above was used to calculate

STR/SNP allele typing success, STR/SNP read depth, and STR/SNP heterozygote balance as described in the previous section.

Results and Discussion

To test the efficacy of each extraction method to remove high amounts of common inhibitors from samples typically recovered from decomposed human remains (blood, hair, and bone) additional amounts of each inhibitor specific to each tissue were added to further challenge the performance of each extraction method. For this comparative study, blood samples spiked with hematin and hairs (with roots) spiked with melanin were extracted using multiple methods (DNA IQ™, DNA Investigator, PrepFiler® BTA, and an organic method). In addition, powdered bone samples spiked with either humic acid or calcium were extracted using five methods (DNA IQ™, DNA Investigator, PrepFiler® BTA, or two different total demineralization protocols).

Hematin

Blood samples were spiked with hematin and extracted with four common DNA extraction methods. CE-generated STRs resulted in complete profiles for all samples with each of the four extraction methods. The APHs ranged from 167108 ± 2120 RFUs (organic method) to 247216 ± 31410 RFUs (PrepFiler® BTA) (Table 4.2). The APHRs ranged from 86% ± 10% (DNA IQ™) to 90% ± 6% (PrepFiler® BTA) (Table 4.2). For the blood samples spiked with hematin, no notable difference in the STR metrics was observed between the four extraction methods tested. All extraction methods yielded amplifiable DNA in blood samples and resultant profiles did not indicate the presence of inhibitors. DNA quantitation results showed the internal PCR control (IPC) with a $\Delta C_T < 1$ for all extraction methods indicating no detectable PCR inhibition and thus efficient DNA

purification. These results support data by Hu et al. [37] for the DNA IQ™ system and organic extraction methods, who also obtained full DNA profiles for high amounts of spiked hematin (532 ng/µL) prior to extraction.

Table 4.2 Average percent of alleles called, average peak height, and average peak height ratios for each DNA extraction method and inhibitor.

Inhibitor	Extraction Method	Alleles Called	Average Peak Height	Average Peak Height Ratio
Hematin	DNA IQ	100	181108 ± 42930	86 ± 10
	DNA Investigator	100	169894 ± 33157	87 ± 8
	PrepFiler	100	247216 ± 31410	90 ± 6
	Organic	100	167108 ± 2120	88 ± 9
Melanin	DNA IQ	100	145038 ± 128489	87 ± 8
	DNA Investigator	100	227592 ± 28296	89 ± 8
	PrepFiler	100	242936 ± 24482	91 ± 6
	Organic	100	210590 ± 31075	86 ± 10
Calcium	DNA IQ	91 ± 5	58477 ± 10176	69 ± 29
	DNA Investigator	95 ± 3	98008 ± 28532	65 ± 32
	PrepFiler	99 ± 1	107053 ± 33499	69 ± 23
	Total Demin 1	96 ± 6	62758 ± 27409	73 ± 24
	Total Demin 2	96 ± 3	71928 ± 15363	78 ± 20
Humic Acid	DNA IQ	91 ± 8	61258 ± 6744	64 ± 6
	DNA Investigator	91 ± 5	85029 ± 8063	62 ± 2
	PrepFiler	97 ± 3	92430 ± 8553	73 ± 5
	Total Demin 1	96 ± 3	56129 ± 15003	72 ± 8
	Total Demin 2	96 ± 3	79586 ± 6046	70 ± 7

The same DNA extracts used for CE-based STR genotyping were also sequenced on both the Ion S5™ and MiSeq FGx™ systems. For the Ion S5™ System, all but one blood sample produced complete STR profiles. When amplifying samples using DL8 library preparation kits, on occasion individual samples would unexpectedly fail to amplify, possibly due to a liquid handling issue on the Ion Chef™. In this particular case, the blood organic extraction control completely failed to produce data and was therefore excluded from data analysis. The read depth ranged from 74891X ± 4548X (DNA Investigator) to 82818X ± 4684X (organic) (Fig. 4.1A). The heterozygote balance ranged from 80% ± 12-16% (DNA IQ™ and organic) to 84% ± 12% (PrepFiler® BTA) (Fig. 4.2A). SNP profiles produced 100% of alleles for all extraction methods. Read depth ranged from 133251X ± 2615X (DNA IQ™) to 147079X ± 3404X (organic) for test samples (Fig. 4.3A). SNP heterozygote balance ranged from 80% ± 6% (organic) to 85% ± 3-5% (DNA IQ™ and PrepFiler® BTA) (Fig. 4.4A). Read depth, allele typing success, and heterozygote balance were not notably different between extraction methods. Therefore, results indicate that the choice of extraction method does not have any negative effect on sequencing efficiency due to the quality of the DNA extracts.

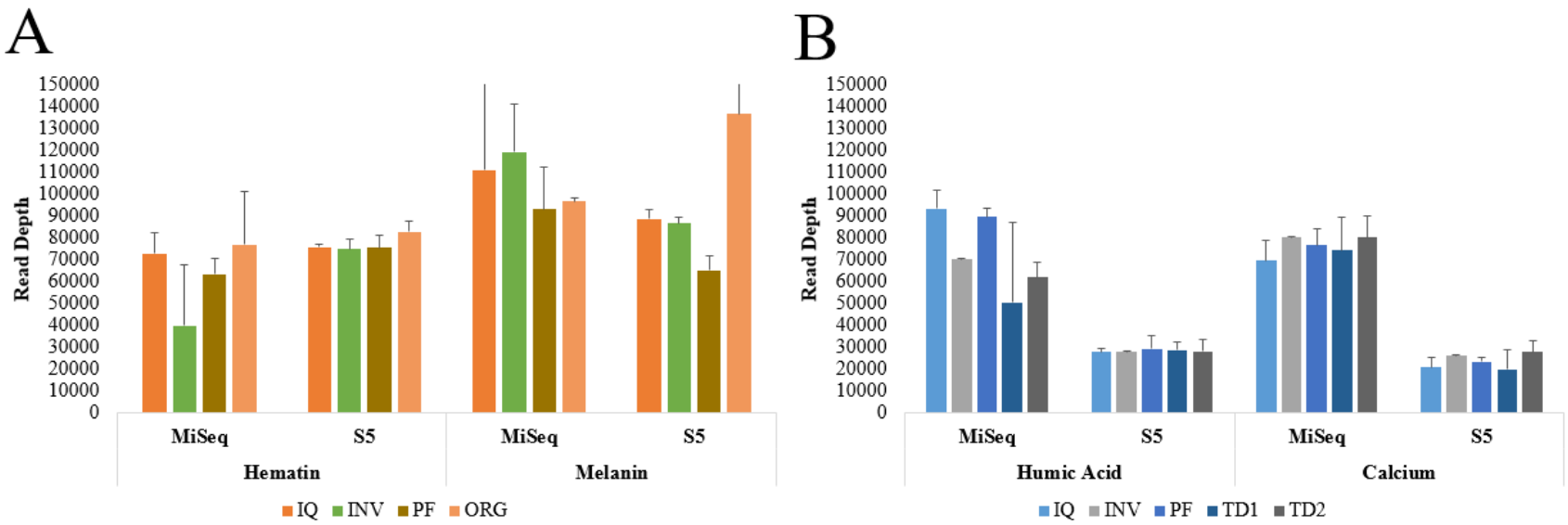


Figure 4.1 STR read depth of **A.)** blood (spiked with hematin) and hair (spiked with melanin) extracted with three commercial kits and an organic method and **B.)** bone (spiked with humic acid and calcium) extracted with three commercial kits and two total demineralization methods, while comparing two sequencing platforms (MiSeq vs. S5). Data presented as average ± SD (N = 3).

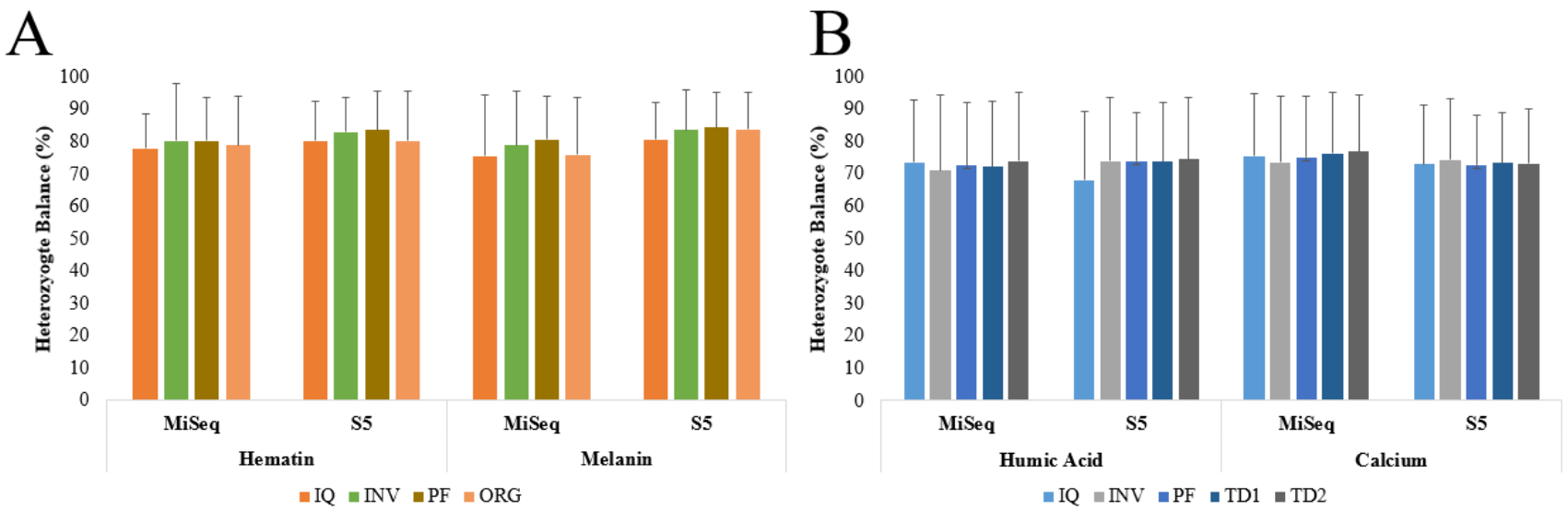


Figure 4.2 STR heterozygote balance of **A.)** blood (spiked with hematin) and hair (spiked with melanin) extracted with three commercial kits and an organic method and **B.)** bone (spiked with humic acid and calcium) extracted with three commercial kits and two total demineralization methods, while comparing two sequencing platforms (MiSeq vs. S5). Data presented as average ± SD (N = 3).

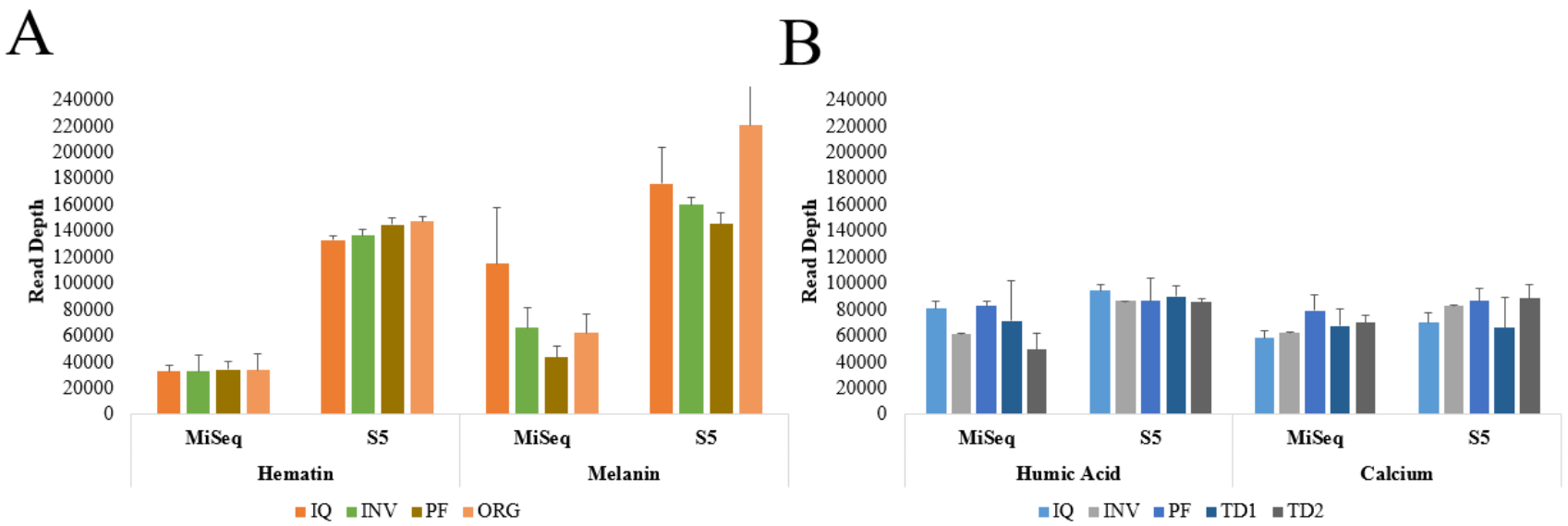


Figure 4.3 SNP read depth of **A.)** blood (spiked with hematin) and hair (spiked with melanin) extracted with three commercial kits and an organic method and **B.)** bone (spiked with humic acid and calcium) extracted with three commercial kits and two total demineralization methods, while comparing two sequencing platforms (MiSeq vs. S5). Data presented as average ± SD (N = 3).

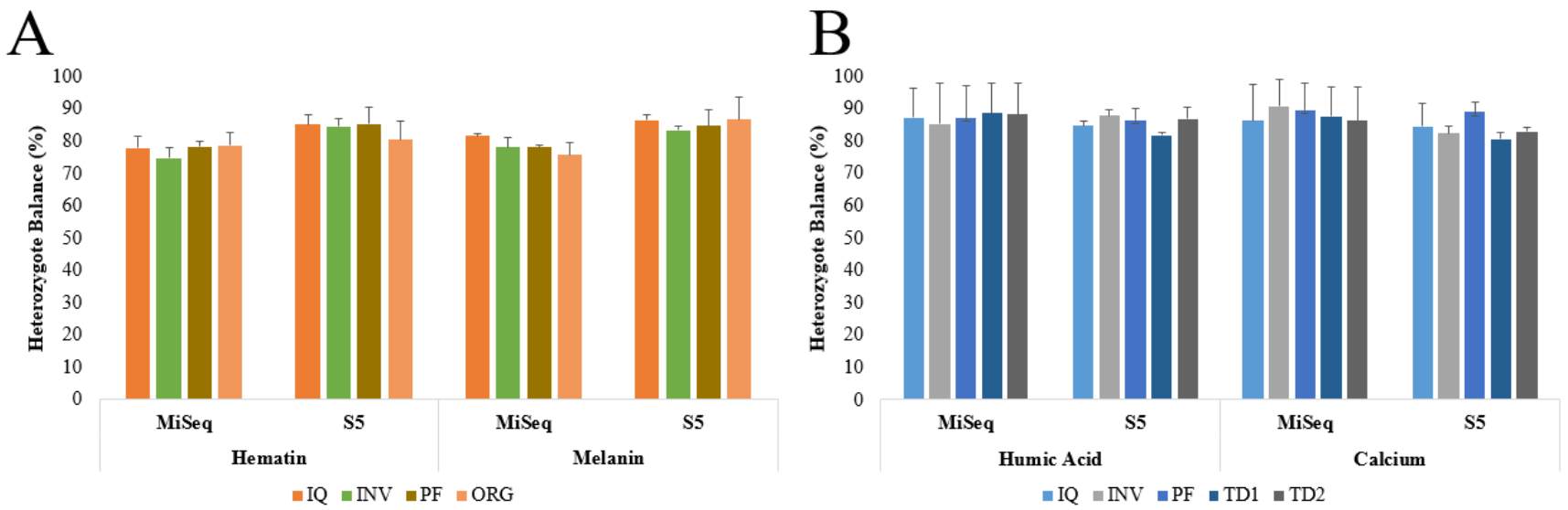


Figure 4.4 SNP heterozygote balance of **A.)** blood (spiked with hematin) and hair (spiked with melanin) extracted with three commercial kits and an organic method and **B.)** bone (spiked with humic acid and calcium) extracted with three commercial kits and two total demineralization methods, while comparing two sequencing platforms (MiSeq vs. S5). Data presented as average ± SD (N = 3).

For the MiSeq FGx™ platform, all hematin-spiked blood samples produced near complete STR profiles with one sample producing 95% of alleles. Read depth ranged from 39858X ± 27639X (DNA Investigator) to 76945X ± 24290X (organic) (Fig. 4.1A). Heterozygote balance ranged from 78% ± 10% (DNA IQ™) to 80% ± 13-18% (PrepFiler® BTA and DNA Investigator) (Fig. 4.2A). SNP profiles were also near complete with only one replicate producing 99% reportable alleles. Read depth ranged from 32643X ± 4596X (DNA IQ™) to 34342X ± 6089X (PrepFiler® BTA) (Fig. 4.3A). SNP heterozygote balance ranged from 75% ± 3% (DNA Investigator) to 79% ± 4% (organic) (Fig. 4.4A). There was no notable difference between extraction methods using the Illumina sequencing platform.

For hematin spiked blood samples, all CE and sequencing based STRs were concordant across all methods. All extraction methods resulted in complete or near complete profiles for sequence based STRs on both S5 and MiSeq platforms. Overall, no notable differences were observed between extraction methods for hematin-spiked blood samples for reportable alleles, read depth, and heterozygote balance. Additionally, both the S5 and MiSeq platforms produced comparable results for these samples.

Melanin

Hair roots were spiked with melanin and extracted with four common DNA extraction methods. Quantitation results showed all IPC ΔC_T values < 1.5 indicating little or no effect of PCR inhibition. DNA IQ™ samples showed a ΔC_T value of 1.40, whereas the other methods demonstrated values ≤ 1. CE-generated STRs resulted in complete profiles for all samples except one hair sample (possible partial, damaged, or missing root) extracted with DNA IQ™, which failed to yield amplifiable amounts of DNA. This sample

was excluded from all further analyses. APHs ranged from 210590 ± 31075 RFUs (organic) to 242936 ± 24482 RFUs (PrepFiler® BTA) (Table 4.2). APHRs ranged from 86% ± 10% (organic) to 91% ± 6% (PrepFiler® BTA) (Table 4.2). However, there was no notable difference between extraction methods, suggesting that all extraction methods produced comparable quality DNA extracts for CE-based STR profiling. These results are similar to those of Faber et al. [7] for samples inhibited with high amounts of melanin [7]. In contrast, Hu et al. [37] did not observe successful data using an organic extraction method with lower amounts of melanin than tested in our study (55 ng compared to 750 ng). These inconsistent results may be due to differences in organic protocols used in each study.

All hair DNA extracts used for CE-based STR analyses were also sequenced using the Ion S5™ and MiSeq FGx™ systems. Hair samples analyzed using the Ion S5™ produced complete STR profiles for each extraction method. Read depth ranged from 65119X ± 2805X (PrepFiler® BTA) to 136395X ± 95709X (organic) (Fig. 4.1A). Heterozygote balance ranged from 81% ± 12% (DNA IQ™) to 84% ± 11% (PrepFiler® BTA) (Fig. 4.2A). SNPs produced complete profiles for each extraction method. SNP read depth ranged from 145220X ± 8671X (PrepFiler® BTA) to 220615X ± 108109X (organic) (Fig. 4.3A). Heterozygote balance ranged from 83% ± 1% (DNA Investigator) to 87% ± 2-7% (DNA IQ™ and organic) (Fig. 4.4A). No notable difference was found between extraction methods and all methods produced high quality data.

When analyzing hair samples with the MiSeq FGx™ sequencing platform, complete profiles were generated for each extraction method in this study. STR read depth ranged from 93118X ± 19325X (PrepFiler® BTA) to 119233X ± 21615X (DNA

Investigator) for test samples (Fig. 4.1A). Heterozygote balance ranged from 75% ± 19% (DNA IQ™) to 81% ± 14% (PrepFiler® BTA) (Fig. 4.2A). All samples produced complete SNP profiles with melanin-spiked test samples producing a read depth ranging from 43553X ± 8959X (PrepFiler® BTA) to 114709X ± 43373X (DNA IQ™) (Fig. 4.3A). Heterozygote balance ranged from 76% ± 4% (organic) to 82% ± 1% (DNA IQ™) (Fig. 4.4A). However, no notable differences were observed between extraction methods.

For melanin-spiked hair samples, all extraction methods resulted in complete profiles and similar heterozygote balance for sequence-based STRs and SNPs on both the S5 and MiSeq platforms. While Elwick et al. [25] previously reported melanin as a strong inhibitor for the MiSeq chemistry but not for the S5 chemistry, extraction methods were able to remove enough inhibitor so that neither chemistry was affected in this study. While the variation among replicates on the MiSeq was greater than the S5, no notable differences were observed between extraction methods for melanin-spiked hair samples for reportable alleles, read depth, and heterozygote balance. Moreover, both the S5 and MiSeq platforms produced comparable high quality data for these samples.

Calcium

Bone samples were spiked with calcium and extracted using five common DNA extraction methods. Quantitation results showed IPC $\Delta C_T < 1$ for all extraction methods confirming no effect due to PCR inhibition. The average percent of STR alleles (by CE) reported ranged from 91% ± 5% (DNA IQ™) to 99% ± 1% (PrepFiler® BTA) (Table 4.2). APHs ranged from 58477 ± 10176 RFUs (DNA IQ™) to 107053 ± 33499 RFUs (PrepFiler® BTA) (Table 4.2). APHRs ranged from 65% ± 32% (DNA Investigator) to 78% ± 19% (TD2) (Table 4.2). According to a previous study [37], when samples were

spiked with calcium chloride (~2-8 μg/μL), the DNA IQ™ system recovered full DNA profiles, however, organic extraction recovered <10% of alleles [37]. In this study, DNA IQ™ recovered >90% of alleles and organic extraction (TD2) recovered >96% of alleles when spiked with high amounts of calcium. This difference may be due to the organic protocols used in each study, and that our the TD2 protocol included an extra purification step prior to downstream processing.

Using the S5 platform, bone samples spiked with calcium produced near complete sequencing profiles barring one DNA IQ™ and one PrepFiler® BTA replicate, generating 96% and 97% of alleles, respectively. Read depth ranged from 20017X ± 9094X (TD1) to 27955X ± 5066X (TD2) (Fig. 4.1B). All control samples produced a read depth higher than that of the test samples, suggesting that some residual amount of inhibitor may still be present in the test samples. However, because the sample size was small and only one control was evaluated for comparison, no definitive conclusions can be made. Heterozygote balance ranged from 73% ± 16-18% (PrepFiler®, TD1, TD2, and DNA IQ™) to 74% ± 3% (DNA Investigator) (Fig. 4.2B). Calcium-spiked bone samples produced near complete SNP profiles for all extraction methods except for one DNA IQ™ replicate with a single SNP dropping out (rs2032599). SNP read depth ranged from 65984X ± 23096X (TD1) to 88501X ± 10861X (TD2) (Fig. 4.3B). Heterozygote balance ranged from 81% ± 2% (TD1) to 89% ± 3% (PrepFiler® BTA) (Fig. 4.4B). Overall, no notable differences were reported between extraction methods when comparing these STR and SNP data.

Using the MiSeq platform, all calcium-spiked bone samples produced complete STR and SNP profiles. Read depth for the STRs ranged from 69542X ± 8049X (DNA

IQ™) to 80504X ± 9430X (DNA Investigator) (Fig. 4.1B), and heterozygote balance ranged from 73% ± 21% (DNA Investigator) to 77% ± 18% (TD2) (Fig. 4.2B). For SNPs, read depth ranged from 58149X ± 3842X (DNA IQ™) to 79252X ± 11510X (PrepFiler® BTA) (Fig. 4.3B), and heterozygote balance ranged from 86% ± 11% (DNA IQ™ and TD2) to 91% ± 8% (DNA Investigator) (Fig. 4.4B). All extraction methods were effective in removing calcium from bone samples.

For calcium spiked bone samples, all extraction methods resulted in complete profiles for the MiSeq platform and near complete profiles using the S5 system for both STRs and SNPs. STR and SNP heterozygote balance was similar for both chemistries. Elwick et al. [25] previously reported that the Ion Torrent chemistry was most tolerant to calcium and the MiSeq chemistry was highly tolerant to calcium. STRs and most SNPs were not affected using the MiSeq chemistry, however, only 52 out of 124 SNPs were refractory to calcium using the S5 chemistry. In this study, we observed that STRs and SNPs were unaffected by calcium using the MiSeq chemistry and S5 chemistries.

Humic Acid

Bone samples were spiked with humic acid and extracted using five common DNA extraction methods. DNA quantitation results produced IPC $\Delta C_T < 1$ for all extraction methods indicating there was no detectable PCR inhibition. For samples spiked with humic acid, the average percent of alleles reported using the GlobalFiler® PCR Amplification kit ranged from 91% ± 5% (DNA Investigator) to 97% ± 3% (PrepFiler® BTA) (Table 4.2). APHs ranged from 56129 ± 15003 RFUs (TD1) to 92430 ± 8553 RFUs (PrepFiler® BTA) (Table 4.2). APHRs ranged from 66% ± 27% (DNA Investigator) to 73% ± 25% (PrepFiler® BTA) (Table 4.2). Hu et al. [37] reported full DNA profiles with high amounts

of humic acid (340 ng/µL) when extracted with the DNA IQ™ kit, and no reportable alleles were detected when spiked samples were extracted using an organic method [37]. However, in our study, near complete profiles (91-96% alleles) were obtained with high amounts of humic acid that were extracted using both the DNA IQ™ and organic methods. Again, the inconsistency may be due to the difference in organic extraction protocols.

Bone samples spiked with humic acid produced complete profiles using the Ion S5™ platform. Read depth ranged from 27925X ± 3602X (DNA IQ™) to 29266X ± 5995X (PrepFiler® BTA) (Fig. 4.1B). Heterozygote balance ranged from 68% ± 21% (DNA IQ™) to 74% ± 15-20% (PrepFiler®, TD1, TD2, and DNA Investigator) (Fig. 4.2B). SNP profiles were complete with 100% of alleles reported for each extraction method. Read depth ranged from 85507X ± 3104X (TD2) to 94239X ± 2211X (DNA IQ™) (Fig. 4.3B). SNP heterozygote balance ranged from 82% ± 1% (TD1) to 88% ± 2% (DNA Investigator) (Fig. 4.4B). There was no notable difference between extraction methods for the data reported for samples spiked with humic acid.

All bone samples produced near complete STR profiles using the MiSeq platform with one replicate producing 62% of alleles. Read depth ranged from 50511X ± 36377X (TD1) to 93309X ± 7326X (DNA IQ™) (Fig. 4.1B). STR heterozygote balance ranged from 71% ± 23% (DNA Investigator) to 74% ± 19-21% (DNA IQ™ and TD2) (Fig. 4.2B). SNP profiles produced 100% of alleles for all samples. Read depth ranged from 49322X ± 12273X (TD2) to 82497X ± 3775X (PrepFiler® BTA) (Fig. 4.3B). Heterozygote balance ranged from 85% ± 13% (DNA Investigator) to 89% ± 9% (TD1) (Fig. 4.4B). There was no notable difference between extraction methods with samples that were spiked with humic acid and sequenced using the Illumina platform.

For humic acid-spiked bone samples, all extraction methods resulted in near complete to complete genetic profiles using both the Illumina and S5 platforms. Overall, both the S5 and MiSeq systems produced comparable data for these bone samples.

Conclusions

The overall results of this study demonstrate that all of the common DNA extraction methods tested were effective in removing high amounts of inhibitors from spiked blood, hair, and bone tissues. These extraction methods all produced sufficiently pure DNA extracts that were equally compatible with both the Precision ID chemistry on the Ion S5™ System and the ForenSeq™ DNA Signature Prep Kit on the MiSeq FGx™ system. All extraction methods produced quantifiable DNA with little or no PCR inhibition detected, indicating that all extraction methods were effective and suitable for preparing samples for MPS. Very little dropout was observed for either platform for both STRs and SNPs. The results of this study demonstrate that the extraction methods commonly used in most crime laboratories are compatible with MPS sequencing chemistries and platforms.

Acknowledgements

This project was supported in part by Award No. NIJ 2015-DN-BX-K066, awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect those of the Department of Justice. We would also like to thank Thermo Fisher Scientific and Joe Chang for all their help and support with troubleshooting.

References

1. Marshall PL, Stoljarova M, Schmedes SE, King JL, Budowle B (2014) A high volume extraction and purification method for recovering DNA from human bone. Forensic Sci Int Genet 12:155-160. doi:10.1016/j.fsigen.2014.06.011

2. Karija Vlahovic M, Kubat M (2012) DNA extraction method from bones using Maxwell(R) 16. Leg Med (Tokyo) 14 (5):272-275. doi:10.1016/j.legalmed.2012.04.004

3. Kus M, Ossowski A, Zielinska G (2016) Comparison of three different DNA extraction methods from a highly degraded biological material. J Forensic Leg Med 40:47-53. doi:10.1016/j.jflm.2016.03.002

4. Barbisin M, Fang R, O'Shea CE, Calandro LM, Furtado MR, Shewale JG (2009) Developmental validation of the Quantifiler® Duo DNA quantification kit for simultaneous quantification of total human and human male DNA and detection of PCR inhibitors in biological samples. J Forensic Sci 54 (2):305-319

5. Schmedes S, Marshall P, King JL, Budowle B (2013) Effective removal of co-purified inhibitors from extracted DNA samples using synchronous coefficient of drag alteration (SCODA) technology. Int J Leg Med 127 (4):749-755. doi:10.1007/s00414-012-0810-7

6. Pionzio AM, McCord BR (2014) The effect of internal control sequence and length on the response to PCR inhibition in real-time PCR quantitation. Forensic Sci Int Genet 9:55-60. doi:10.1016/j.fsigen.2013.10.010

7. Faber KL, Person EC, Hudlow WR (2013) PCR inhibitor removal using the NucleoSpin(R) DNA Clean-Up XS kit. Forensic Sci Int Genet 7 (1):209-213. doi:10.1016/j.fsigen.2012.06.013

8. Thompson RE, Duncan G, McCord BR (2014) An Investigation of PCR Inhibition Using Plexor®-Based Quantitative PCR and Short Tandem Repeat Amplification. J Forensic Sci 59 (6):1517-1529

9. Wang DY, Chang CW, Lagacé RE, Calandro LM, Hennessy LK (2012) Developmental validation of the AmpFℓSTR® Identifiler® Plus PCR Amplification Kit: an established multiplex assay with improved performance. J Forensic Sci 57 (2):453-465

10. Parson W, Ballard D, Budowle B, Butler JM, Gettings KB, Gill P, Gusmao L, Hares DR, Irwin JA, King JL, Knijff P, Morling N, Prinz M, Schneider PM, Neste CV, Willuweit S, Phillips C (2016) Massively parallel sequencing of forensic STRs: Considerations of the DNA commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements. Forensic Sci Int Genet 22:54-63. doi:10.1016/j.fsigen.2016.01.009

11. Flores S, Sun J, King J, Budowle B (2014) Internal validation of the GlobalFiler Express PCR Amplification Kit for the direct amplification of reference DNA samples on a high-throughput automated workflow. Forensic Sci Int Genet 10:33-39. doi:10.1016/j.fsigen.2014.01.005

12. Van Neste C, Van Nieuwerburgh F, Van Hoofstat D, Deforce D (2012) Forensic STR analysis using massive parallel sequencing. Forensic Sci Int Genet 6 (6):810-818

13. Zubakov D, Kokmeijer I, Ralf A, Rajagopalan N, Calandro L, Wootton S, Langit R, Chang C, Lagace R, Kayser M (2015) Towards simultaneous individual and tissue

identification: A proof-of-principle study on parallel sequencing of STRs, amelogenin, and mRNAs with the Ion Torrent PGM. Forensic Sci Int Genet 17:122-128. doi:10.1016/j.fsigen.2015.04.002

14. Børsting C, Morling N (2015) Next generation sequencing and its applications in forensic genetics. Forensic Sci Int Genet 18:78-89
15. Berglund EC, Kiialainen A, Syvanen AC (2011) Next-generation sequencing technologies and applications for human genetic history and forensics. Investig Genet 2 (1):23. doi:10.1186/2041-2223-2-23
16. Yang Y, Xie B, Yan J (2014) Application of next-generation sequencing technology in forensic science. Genomics Proteomics Bioinformatics 12 (5):190-197. doi:10.1016/j.gpb.2014.09.001
17. Iozzi S, Carboni I, Contini E, Pescucci C, Frusconi S, Nutini AL, Torricelli F, Ricci U (2015) Forensic genetics in NGS era: New frontiers for massively parallel typing. Forensic Sci Int Genet Suppl Ser 5:E418-E419. doi:10.1016/j.fsigss.2015.09.166
18. Huynen L, Lambert DM (2015) A Concentrated Hydrochloric Acid-based Method for Complete Recovery of DNA from Bone. J Forensic Sci 60 (6):1553-1557. doi:10.1111/1556-4029.12846
19. Kitayama T, Ogawa Y, Fujii K, Nakahara H, Mizuno N, Sekiguchi K, Kasai K, Yurino N, Yokoi T, Fukuma Y, Yamamoto K, Oki T, Asamura H, Fukushima H (2010) Evaluation of a new experimental kit for the extraction of DNA from bones and teeth using a non-powder method. Leg Med 12 (2):84-89. doi:10.1016/j.legalmed.2009.12.004

20. Combs LG, Warren JE, Huynh V, Castaneda J, Golden TD, Roby RK (2015) The effects of metal ion PCR inhibitors on results obtained with the Quantifiler((R)) Human DNA Quantification Kit. Forensic Sci Int Genet 19:180-189. doi:10.1016/j.fsigen.2015.06.013
21. Alaeddini R (2012) Forensic implications of PCR inhibition--A review. Forensic Sci Int Genet 6 (3):297-305. doi:10.1016/j.fsigen.2011.08.006
22. Hudlow WR, Chong MD, Swango KL, Timken MD, Buoncristiani MR (2008) A quadruplex real-time qPCR assay for the simultaneous assessment of total human DNA, human male DNA, DNA degradation and the presence of PCR inhibitors in forensic samples: a diagnostic tool for STR typing. Forensic Sci Int Genet 2 (2):108-125. doi:10.1016/j.fsigen.2007.09.001
23. Elwick K, Mayes C, Hughes-Stamm S (2018) Comparative Sensitivity and Inhibitor Tolerance of GlobalFiler® PCR Amplification and Investigator® 24plex QS Kits for Challenging Samples. Leg Med
24. Mulero JJ, Chang CW, Lagace RE, Wang DY, Bas JL, McMahon TP, Hennessy LK (2008) Development and validation of the AmpFℓSTR® MiniFilerTM PCR Amplification Kit: a miniSTR multiplex for the analysis of degraded and/or PCR inhibited DNA. J Forensic Sci 53 (4):838-852
25. Elwick K, Zeng X, King J, Budowle B, Hughes-Stamm S (2018) Comparative tolerance of two massively parallel sequencing systems to common PCR inhibitors. Int J Leg Med 132 (4):983-995. doi:10.1007/s00414-017-1693-4
26. Laboratory F (1994) PCR-Based Typing Protocols. Federal Bureau of Investigation,

27. Corporation P (2013) Technical Bulletin: DNA IQ™ System - Database Protocol. TB297 edn. Promega Corporation, Madison, WI

28. Corporation P (2013) Technical Bulletin: DNA IQ™ System - Small Sample Casework Protocol. TB296 edn. Promega Corporation, Madison, WI

29. Corporation P (2015) Technical Bulletin: Tissue and Hair Extraction Kit (for use with DNA IQ™). TB307 edn. Promega Corporation, Madison, WI

30. Technologies L (2012) PrepFiler ® and PrepFiler ® BTA Forensic DNA Extraction Kits User Guide. vol 4463348, Rev. C edn. Life Technologies, Carlsbad, CA

31. QIAGEN (2012) QIAamp ® DNA Investigator Handbook. QIAGEN,

32. Loreille OM, Parr RL, McGregor KA, Fitzpatrick CM, Lyon C, Yang DY, Speller CF, Grimm MR, Grimm MJ, Irwin JA, Robinson EM (2010) Integrated DNA and fingerprint analyses in the identification of 60-year-old mummified human remains discovered in an Alaskan glacier. J Forensic Sci 55 (3):813-818. doi:10.1111/j.1556-4029.2010.01356.x

33. Identification UCfH (2008) Procedure Manual - Forensic Laboratory. Demineralization Extraction of Skeletal Remains, Rev 2 edn., Fort Worth, TX

34. Scientific TF (2015) GlobalFiler™ PCR Amplification Kit. Thermo Fisher Scientific, Thermo Fisher Scientific

35. Illumina (2015) ForenSeq™ DNA Signature Prep Reference Guide. vol 15049528, v01 edn., San Diego, CA

36. King JL, Wendt FR, Sun J, Budowle B (2017) STRait Razor v2s: Advancing sequence-based STR allele reporting and beyond to other marker systems. Forensic Sci Int Genet 29:21-28. doi:10.1016/j.fsigen.2017.03.013

37. Hu Q, Liu Y, Yi S, Huang D (2015) A comparison of four methods for PCR inhibitor removal. Forensic Sci Int Genet 16:94-97. doi:10.1016/j.fsigen.2014.12.001

CHAPTER V

Utility of the Ion S5™ and MiSeq FGx™ sequencing platforms to characterize challenging human remains[1]

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

[1]Elwick K, Bus M, King JL, Chang J, Hughes-Stamm S, Budowle B (2018). Submitted to *Forensic Science International: Genetics*

Abstract

Often in missing persons' and mass disaster cases the samples remaining for analysis typically are hard tissues such as bones, teeth, nails, and hair. These remains may have been exposed to harsh environmental conditions such as heat, humidity, fire, UV radiation, and microorganisms, which pose challenges for downstream genotyping. The harsh conditions may have damaged, degraded, or introduced PCR inhibitors to the DNA prior to analysis. Short tandem repeat analysis (STR) via capillary electrophoresis (CE) is still the gold standard for DNA typing; however, a newer technology known as massively parallel sequencing (MPS) is an alternate method that contributes to higher success for human identification (HID) and forensic intelligence purposes. This technology could improve upon our current techniques by typing different and more markers in a single analysis, and consequently improving the power of discrimination.

In this study, bone and tooth samples exposed to a variety of DNA insults (cremation, embalming, decomposition, thermal degradation, and fire) were assessed and sequenced using the Precision ID chemistry and a custom AmpliSeq™ STR and iiSNP panel on the Ion S5™ System, and the ForenSeq™ DNA Signature Prep Kit on the MiSeq FGx™ system, as well as the GlobalFiler® PCR Amplification Kit on the 3500 Genetic Analyzer.

The results demonstrated that using traditional CE-based genotyping performed as expected, producing a partial or full DNA profile for all samples, and that both sequencing chemistries and platforms were able to recover sufficient STR and SNP information from a majority of the same challenging and degraded human remains. Run metrics including profile completeness, mean read depth produced good results with each system,

considering the degree of damage of some samples. Most sample insults (except decomposed) produced similar numbers of alleles for both MPS systems. Comparable markers produced full concordance between the two platforms.

Keywords: Massively parallel sequencing, Ion S5™, MiSeq FGx™, Missing Persons, Human remains, Challenged remains

Introduction

Missing persons' cases, unidentified human remains, and mass disasters are problems encountered worldwide [1]. An overwhelming number of migrants and refugees have died or gone missing due to their efforts to cross borders or seas [2] or through human trafficking [3]. Routinely when identifying human remains in missing persons' cases, skeletal remains (bone, teeth) are the only samples available for DNA analysis [4-9]. However, some samples are more challenging to process than others due to their biological composition, environmental exposure (humidity, temperature, UV light, and microorganisms), DNA damage and/or degradation, the presence of inhibitors, and the possibility of contamination or comingled remains [4,7,10,11].

Currently, amplification of short tandem repeat (STR) loci in tandem with capillary electrophoresis (CE) is most commonly used to analyze such remains [10]. STRs are most frequently used because of their high discriminatory power. However, these severely compromised samples may not have suitable fragment lengths to generate a full CE-based STR profiles, decreasing the power of discrimination [12]. Therefore, other methods and genetic markers are being explored that may be more amenable to typing challenged samples. Single nucleotide polymorphisms (SNPs) using massively parallel sequencing (MPS) may be applicable for some degraded samples [4,12]. MPS demonstrates promising capabilities such as large sample multiplexing, improved mixture deconvolution, and the simultaneous analysis of different types of markers (e.g., identity informative SNPs (iiSNPs), ancestry informative SNPs (aiSNPs), STRs, and phenotypic informative SNPs (piSNPs) [13-15]. The use of multiple marker systems simultaneously (STRs and SNPs) can provide higher powers of discrimination and greater typing success with challenged

samples than analyzing solely STRs. Furthermore, MPS can also detect sequence variation within the amplicons of these markers, many revealing SNPs within STR repeat regions [16] and unreported microvariants [17-19], which were previously undetected using CE technology.

The goal of this study was to evaluate two MPS chemistries and platforms and compare their performance with traditional CE-based genotyping using challenged human remains that may be encountered in missing persons' cases. Bone and tooth samples were extracted using a total demineralization (TD) protocol [20]. The extracted DNA was quantified, STR-typed via CE, and then sequenced using both a custom AmpliSeq™ STR and iiSNP panel for degraded remains with Precision ID chemistry on the Ion S5™ system and the ForenSeq™ DNA Signature Prep Kit (using Primer Mix A) on the MiSeq FGx™. Performance between the two systems was determined by comparing read depth, heterozygote balance, and the total number of alleles or percentage of alleles. Percentage/number of alleles and the performance of the CODIS loci were compared between the three systems (two MPS systems and CE).

Materials and Methods

Sample Preparation

Bone (N = 19) and teeth (N = 5) samples from 14 cadavers were collected from the Applied Anatomical Research Center (AARC) at Sam Houston State University in Huntsville, Texas. These samples were subjected to a range of insults including cremation, embalming, decomposition, thermal degradation, and fire (Table 5.1). The remains were cremated in an oven at 900°C for 2.5 hours; embalmed remains were preserved with 30% glutaraldehyde for 880 days; teeth were thermally degraded in an oven at 232°C for 45

minutes; decomposed remains were surface exposed for 12-18 months; and burned remains were ignited with gasoline in a house (mock arson scene) and burned until they self-extinguished.

Table 5.1 Sample information including bone type, environmental insult, and donor.

Cadaver	Bone	Insult
1	Top Vertebral Arch	Cremated
2	Femur	Embalmed
3	Femur	Burned
4	Premolar	Thermally Degraded
5	Molar	Thermally Degraded
6	Premolar	Thermally Degraded
7	Premolar	Thermally Degraded
8	Molar	Thermally Degraded
9	Femur	Decomposed
	Humerus	Decomposed
	Tibia	Decomposed
10	Femur	Decomposed
	Humerus	Decomposed
	Tibia	Decomposed
11	Femur	Burned
	Humerus	Burned
	Tibia	Burned
12	Femur	Burned
	Humerus	Burned
	Tibia	Burned
13	Femur	Burned
	Humerus	Burned
	Tibia	Burned
14	Fibula	Burned

Bone sections were cleaned, chipped, and powdered as described in Zeng et al. [21]. Teeth were cleaned with a sterile toothbrush using 10% bleach, rinsing with DI H_2O, brushing with 70% ethanol, and rinsing again with DI H_2O. Teeth were individually wrapped in large task wipes, lightly crushed with a hammer, and powdered using a SPEX CertiPrep 6750 Freezer/Mill Cryogenic Grinder.

Three samples of each bone and tooth powders (300 mg) were extracted using a TD protocol [20]. Reference buccal swabs were collected before the cadavers were exposed to any insults (burning, decomposition, etc.). Reference swabs were then extracted using the AutoMate *Express*™ Forensic DNA Extraction System and PrepFiler *Express*™ (Thermo Fisher Scientific) according to the manufacturer's protocol [22]. Extracted DNA was quantified with Quantifiler® Trio DNA Quantification Kit (Thermo Fisher Scientific, Waltham, MA, USA) using a 7500 Real-Time PCR System (Thermo Fisher Scientific) according to manufacturer's instructions [23].

CE-based STR Analysis

PCR amplification of STRs was performed using the GlobalFiler® PCR Amplification Kit (Thermo Fisher Scientific) on a ProFlex™ 96-well PCR System in accordance with the manufacturer's protocol [24]. DNA target input was 0.8 ng, whereas for low template samples (<0.05 ng/µL) the full 15 µL of extract were amplified. Separation and detection were performed using a 3500 Genetic Analyzer with POP-4™ polymer and a 36 cm capillary array (Thermo Fisher Scientific). Data were analyzed with GeneMapper™ ID-X v. 1.4 and an in-house excel workbook. An analytical threshold of 150 RFUs and a stochastic threshold of 600 RFUs were used to assign allele peaks. Average peak height (APH) was calculated by summing the peak heights at each locus of the sample replicates and dividing by the number of replicates. Average peak height ratios (APHR) were calculated by summing the peak height ratios at each locus for the sample replicates and dividing by the number of replicates. If allele or locus dropout occurred, the peak height ratio of that locus was given a value of zero. The standard deviation (SD) was calculated using the three replicates per sample.

Ion S5™ Sequencing

An automated library preparation method was chosen based on sample volume as up to 15 µL of DNA extract can be used with the Precision ID DL8 Kit (Thermo Fisher Scientific) on the Ion Chef™ System (Thermo Fisher Scientific), whereas manual library preparation is limited to 6 µL of extract. All low template samples (i.e., <0.16 ng/µL) were amplified and prepared using the DL8 kit and DNA samples greater than or equal to 0.16 ng were prepared manually using the Precision ID Library Kit (Thermo Fisher Scientific). A custom AmpliSeq™ STR and iiSNP primer panel (Thermo Fisher Scientific), including 32 STR markers, 1 Y-indel, 2 amelogenin sex markers, 41 iiSNPs, and 34 Y-SNPs, was used to amplify the extracted DNA. This panel consists of all STRs from the Precision ID GlobalFiler NGS STR Panel v2 and 75 SNPs from the Precision ID Identity Panel (Thermo Fisher Scientific). All samples (N = 81) were sequenced in four runs. Two control samples (007 control DNA from Thermo Fisher Scientific) and two negative control samples (nuclease-free H_2O) were amplified with the manual library preparation. One control sample (007 control DNA) was amplified with each DL8 IonCode PCR plate. Libraries were quantified using the Ion Library TaqMan® Quantitation Assay (Thermo Fisher Scientific). Two pools of "high" quantity libraries were diluted to 50 pM, one pool of "mid-range" quantity libraries was combined neat at ~26 pM, and one pool of "low" quantity libraries was combined neat at ~12 pM (Supplemental Table 2). Templating and chip loading were performed using the Ion Chef™ System on a 530™ semiconductor chip, and sequencing was performed using the Ion S5™ Precision ID Chef and Sequencing Kit with the Ion S5™ System (Thermo Fisher Scientific). Data analyses were performed using Converge™ 2.0 (Thermo Fisher Scientific) and in-house excel workbooks. For STRs and

SNPs, mean read depth was calculated by summing the total usable reads for the sample replicates and dividing by the number of replicates. STR and SNP heterozygote balance was calculated by averaging the heterozygote balance across sample replicates. Allele and locus dropout were treated as was done with CE-based dropout described previously. A minimum arbitrary detection threshold of 5X was used for both systems.

Table 5.2 Ion S5™ sequencing run metrics. Bolded numbers are not within a recommended range according to Torrent Suite Software (TSS).

Chip	No. Samples	Pooling Concentration	% Chip Loading (40% - 70%)	% Usable Reads (>30%)	% Poly-clonal (20% - 40%)	Total Reads	Mean Read Length
1	24	50 pM	42%	35%	34%	5279709X	78 bp
2	33	50 pM	56%	32%	34%	6576081X	112 bp
3	14	~26 pM	**36%**	**29%**	32%	3748684X	114 bp
4	28	~12 pM	**37%**	**28%**	30%	3732793X	102 bp

MiSeq FGx™ Sequencing

Libraries were prepared using the ForenSeq™ DNA Signature Prep Kit (Verogen, Inc., San Diego, CA, USA) with Primer Mix A following the manufacturer's protocol [25]. Primer Mix A targets 27 autosomal STRs, 24 Y-STRs, and 7 X-STRs) and 94 iiSNPs. Samples with more than 0.2 ng of DNA were normalized to 0.2 ng and samples below 0.2 ng were used neat (0.01 ng – 0.1 ng) (5 µL maximum input). Normalized sample libraries including positive 2800M template control from the ForenSeq™ DNA Signature Prep Kit and a negative control (nuclease-free H_2O) were pooled in equal volumes according to the manufacturer's protocol [25] (Table 5.3). Sequencing was performed on a MiSeq FGx™ (Illumina, San Diego, CA, USA) instrument using the MiSeq FGx™ Reagent Kit

(Verogen) and the manufacturer's protocol [25]. Data analyses were performed using the ForenSeq Universal Analysis Software (Verogen), STRait Razor v2s [26], and in-house excel workbooks. The same data metrics calculated using the Ion S5 were also calculated in the same manner for the MiSeq.

Table 5.3 MiSeq FGx™ sequencing run metrics. Bolded numbers indicate metrics that are not within a recommended range according to ForenSeq™ Universal Analysis Software (UAS).

Run	No. Samples	Cluster Density (400-1650 K/mm^2)	Cluster Passing Filter (≥80%)	Phasing (≤0.25%)	Pre-phasing (≤0.15%)	Total Reads
1	32	642	93.81%	**0.285%**	-	7740000X
1 re-run	24	**294**	97.62%	0.211%	-	3620000X
2	31	1060	90.08%	0.160%	0.032%	12610000X
3	32	1143	88.08%	0.153%	0.098%	13260000X

Results and Discussion

Capillary Electrophoresis (CE)

Reportable Alleles

The number of reportable alleles was determined by the number of alleles present out of the total number of alleles expected. The expected number of alleles was determined by the total number of alleles in each panel. Full female profiles produced 44 alleles and full male profiles produced 46 alleles. All bone and teeth samples amplified with the GlobalFiler™ PCR Amplification Kit produced a STR profile to varying degrees of profile completeness. Reportable alleles ranged from 10 ± 3 to complete profiles across the samples (Fig. 5.1). The thermally degraded teeth samples produced full profiles. The embalmed and cremated samples produced complete and near complete profiles, respectively. Mock arson burned samples produced profiles ranging from 27 ± 6 reportable

alleles to full profiles. Decomposed skeletal remains produced the most degraded DNA profiles, ranging from 10 ± 3 to 32 ± 5.

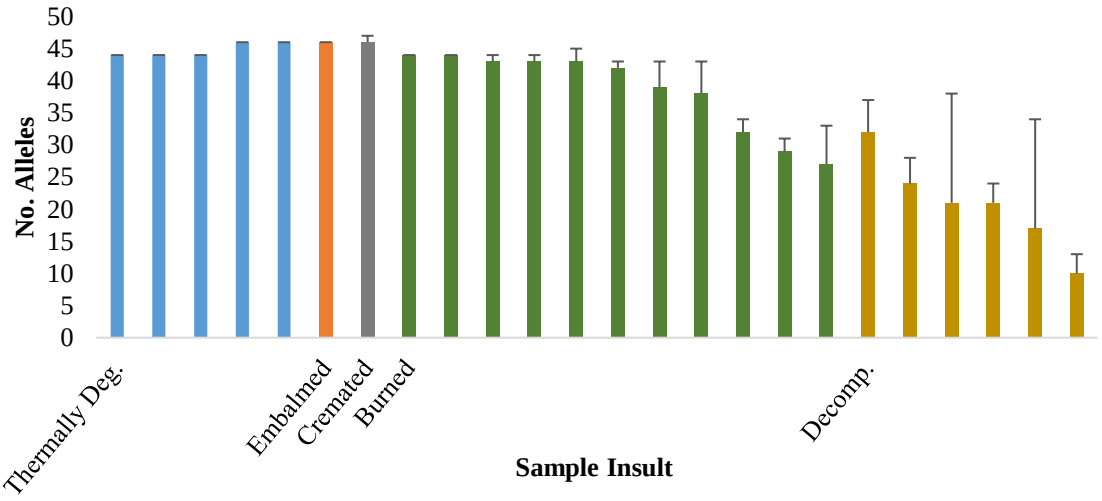


Figure 5.1 Profile completeness of 24 challenging human remains samples genotyped using the GlobalFiler® PCR Amplification Kit. Data presented as average + SD (N = 3).

Peak Height and Peak Height Ratios

APH across all samples ranged from 5154 ± 1952 relative fluorescence units (RFUs) to 210578 ± 8846 RFUs (Fig. 5.2). Overall, the pattern observed across the sample types when considering the APH was consistent with the trend seen with STR profile completeness. The thermally degraded samples produced the highest APHs ranging from 65684 ± 18897 RFUs to 210578 ± 8846 RFUs, while the decomposed remains produced the lowest APHs ranging from 5154 ± 1952 RFUs to 50751 ± 21380 RFUs.

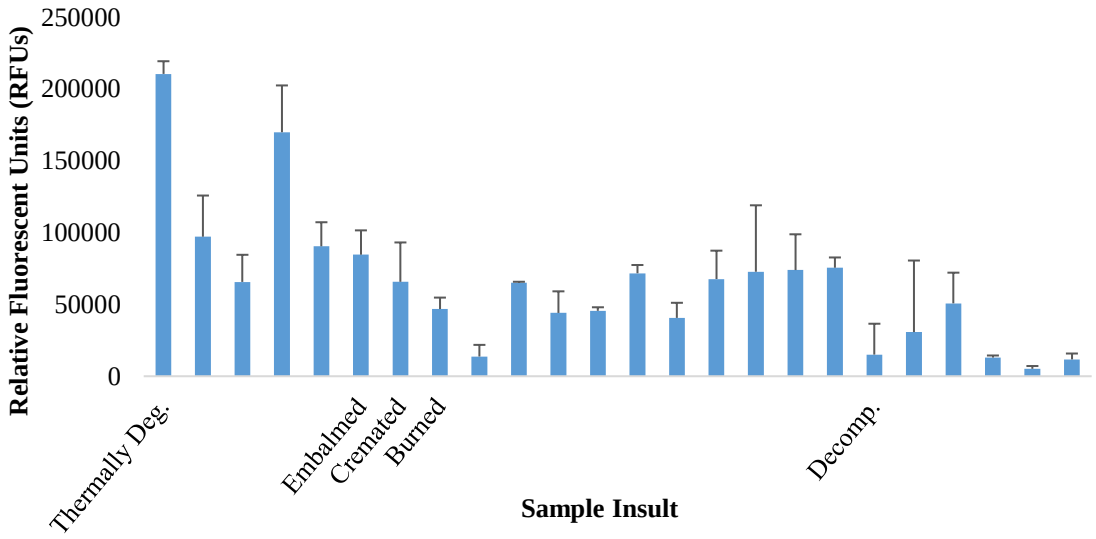


Figure 5.2 Average peak height of 24 challenging human remains samples genotyped using the GlobalFiler® PCR Amplification Kit. Data presented as average + SD (N = 3).

APHRs showed a similar trend to both profile completeness and APH, decreasing from thermally degraded to decomposed samples. APHRs ranged from 8% ± 25% to 87% ± 11% across all samples, with just below half (46%) of the samples showing APHRs below 70% (Fig. 5.3). The thermally degraded teeth ranged from 71% ± 19% to 87% ± 11% while the decomposed remains produced the least balanced profiles ranging from 8% ± 25% to 44% ± 42% APHRs.

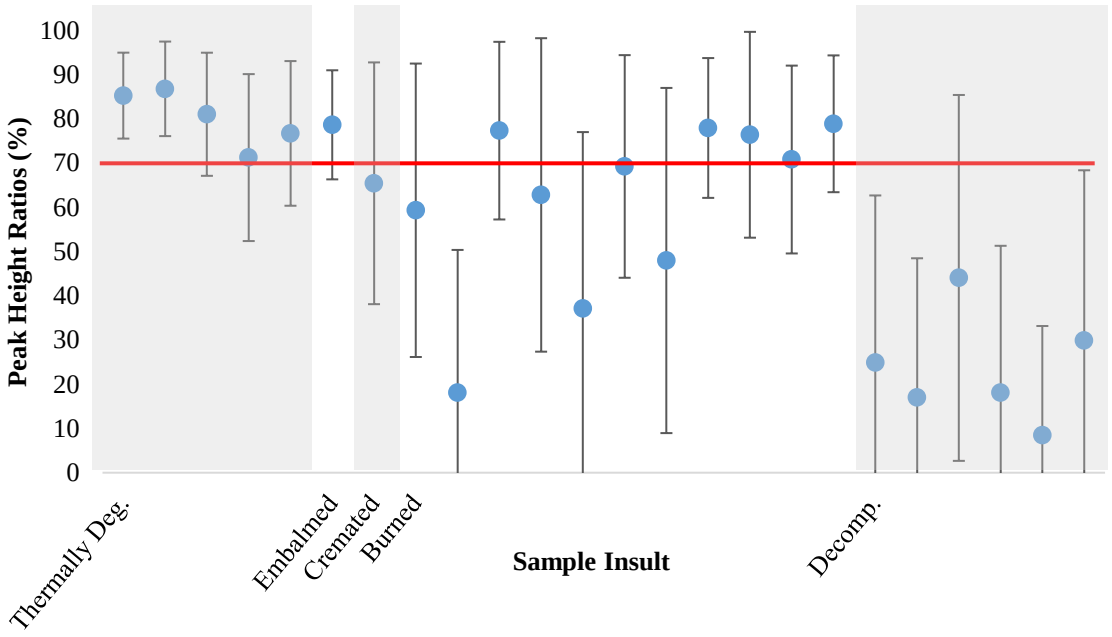


Figure 5.3 Average peak height ratios of 24 challenging human remains samples genotyped using the GlobalFiler® PCR Amplification Kit. Data presented as average ± SD (N = 3). The red line denotes 70% APHR threshold. Shading was included to more easily differentiate between sample insults.

Allelic Dropout

Allele dropout was determined by summing the number of alleles that dropped out at each locus across all samples. Allelic dropout was determined by comparison to a reference sample. No allelic dropout was observed with the thermally degraded (five samples) and embalmed samples (one sample), the cremated sample (one sample) produced one dropout event at the DYS391 locus. The burned samples (eleven samples) produced 153 instances of allele dropout, and the decomposed samples (six samples) resulted in the highest amount of allele dropout with 396 occurrences. As expected, the number of allelic dropout events increased as the size of the locus increased. Alleles at the loci D16S539, D7S820, FGA, CSF1PO, D18S51, TPOX, and SE33 experienced the most allelic dropout (Fig. 5.4), which has been observed in several other studies [27-29]. Alleles at the SE33 locus dropped out

most often with 62 occurrences overall. In contrast, only one instance of allelic dropout occurred at the Y INDEL and D22S1045 loci.

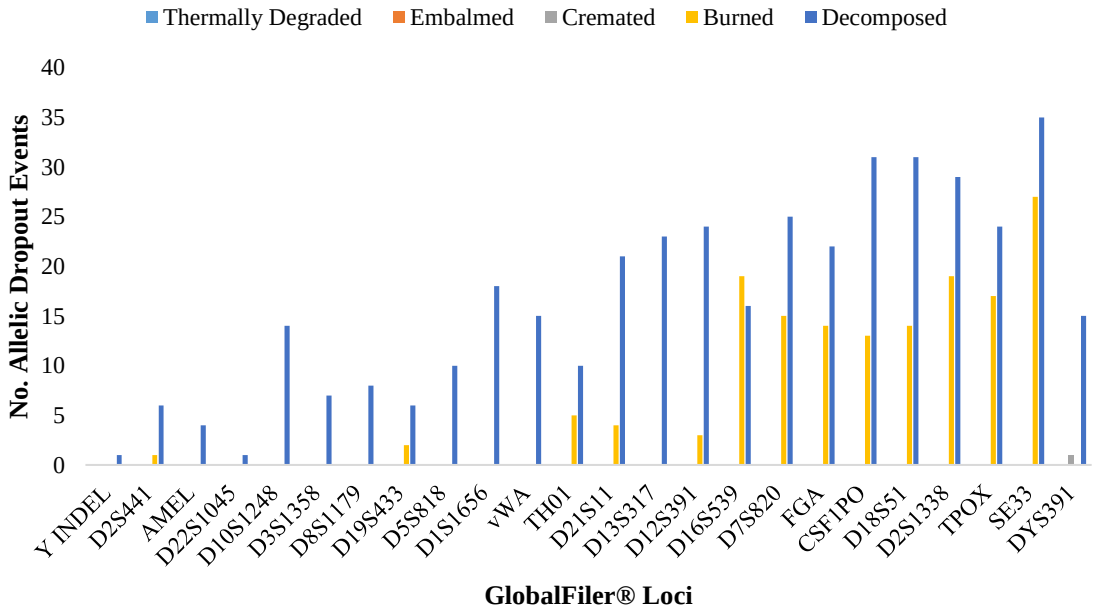


Figure 5.4 Number of allelic dropout events for each sample insult using the GlobalFiler® PCR Amplification Kit. Loci are arranged from smallest (~90 bp) to largest (~380 bp) fragment size.

Massively Parallel Sequencing (MPS)

In this study, STR and SNP typing success was assessed via the number of reportable alleles, read depth, and heterozygote balance. Between the two platforms, all comparable results were concordant.

Reportable Alleles

The number of STR and SNP reportable alleles was calculated in the same manner as CE-based STRs. For the Ion S5™ system, full female STR profiles resulted in 64 alleles and SNP profiles produced 82 alleles. Full male STR profiles resulted in 67 alleles, and SNP profiles produced 75 alleles. For the MiSeq, full female STR profiles resulted in 70

alleles, full male STR profiles resulted in 88 alleles, and SNP profiles produced 188 alleles for both sexes. All Ion S5™ samples sequenced produced reportable alleles ranging from one allele to full profiles. However, using the MiSeq, two samples (one decomposed and one thermally degraded) produced no DNA profile. The decomposed sample that produced no profile with the MiSeq produced one allele with the Ion S5™. In contrast, the thermally degraded sample that produced no profile with the MiSeq resulted in 98% of alleles using the Ion S5™. Metrics are described for each sequencing run using the Ion S5™ (Table 5.2) and the MiSeq FGx™ (Table 5.3).

Using the Ion S5™, STR profiles ranged from 1 ± 1 allele to full profiles (Fig. 5.5A). All samples except decomposed remains produced >90% of alleles. Only three profiles produced below 50% of reportable alleles, all of which were decomposed skeletal remains. Reportable alleles for decomposed remains ranged from 1 ± 1 to 62 ± 3. For SNPs, the Ion S5™ produced profiles ranging from 61 ± 55 alleles to full profiles (Fig. 5.5A). Similar to STRs, all samples except decomposed skeletal samples produced >90% of alleles. Decomposed remains resulted in profiles ranging from 47 ± 42 to 116 ± 0 alleles (Fig. 5.5A).

Using the MiSeq, STR profiles ranged from 0 alleles to complete profiles (Fig. 5.5B). Most thermally degraded, embalmed, and burned samples produced near complete or complete profiles. However, four samples (two burned and two thermally degraded) produced profiles <75%. Decomposed remains demonstrated the highest level of degradation (degradation index (DI) values from 1.6 to 18.5 and IPC ΔC_T values less than 1) with the number of reportable alleles ranging from 0 to 15 ± 14 alleles. SNPs showed a similar pattern to STRs with reportable alleles ranging from 0 alleles to full profiles (Fig.

5.5B). Like STRs, embalmed, and burned samples showed near complete or complete profiles. In contrast to STRs, all but one thermally degraded sample produced profiles ≤75%. The cremated sample produced similar results for SNPs and STRs showing ~70% of reportable alleles. The number of reportable alleles for decomposed skeletal samples ranged from 0 to 20 ± 16 alleles, demonstrating the most degradation as described above (Fig. 5.5B).

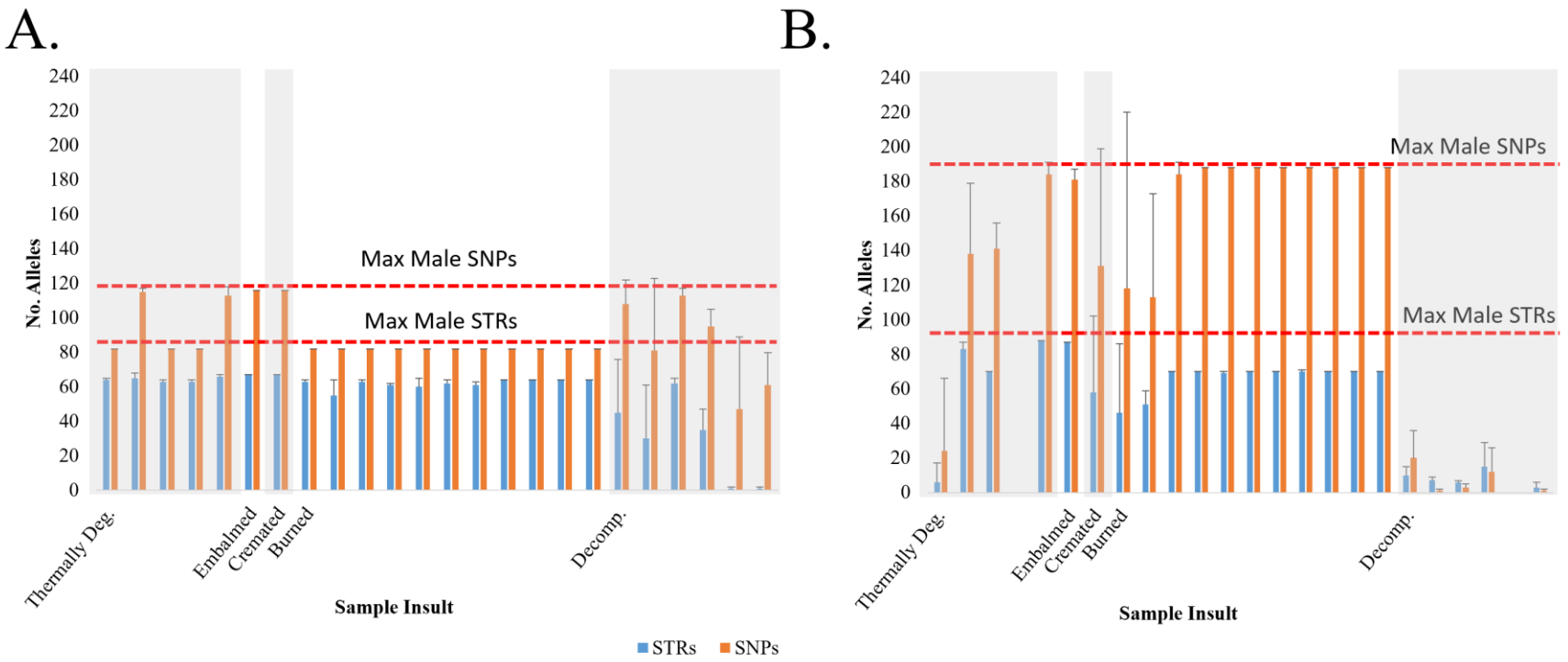


Figure 5.5 Profile completeness (number of alleles) of STRs and SNPs sequenced using the **A.)** Ion S5™ and **B.)** MiSeq FGx™ systems for 24 challenged human remains samples. Data presented as average + SD (N = 3). Dotted red line denotes the maximum number of male STRs and SNPs for each sequencing platform. Minimum detection threshold of 5X coverage.

Overall, both sequencing platforms produced quality data for the types of challenged remains analyzed. On the Ion S5™ the SNPs demonstrated higher profile completeness than that of STRs, producing ~10% more alleles than STRs overall (~93% ± 29% vs ~84% ± 16%). The severely compromised decomposed remains were especially difficult to analyze. However, with Precision ID DL8 library preparation 15 µL of low quantity sample were used increasing the DNA input amount compared to manual library preparation (6 µL). Therefore, the Precision ID DL8 library preparation on the Ion Chef was more flexible than manual library preparation when amplifying low template samples. In this study, there was a correlation observed between DNA input and percentage of reportable alleles ($p < 0.05$) (Fig. 5.6A).

For most MiSeq samples, profile completeness between STRs and SNPs was comparable. In general, slightly more STR alleles were produced than SNPs (~66% ± 44% vs ~63% ± 44%). However, there were a few samples that demonstrated a >20% increase in STR profile completeness compared to SNPs. With MiSeq chemistry, only 5 µL of low template sample could be amplified resulting in a lower number of alleles being genotyped for compromised samples. Sample concentration would likely improve these results; DNA input versus profile completeness demonstrated a correlation in this study ($p < 0.05$) (Fig. 5.6B).

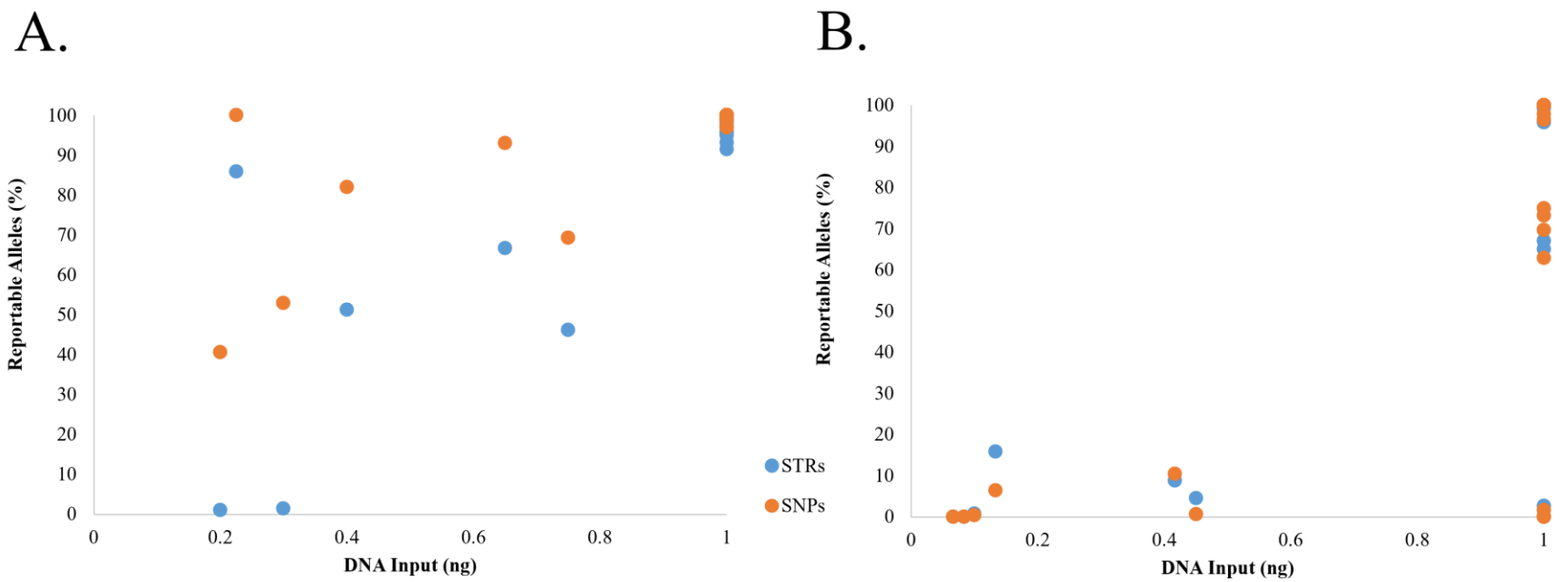


Figure 5.6 DNA input vs. reported alleles (%) of the **A.)** Ion S5™ using the Precision ID chemistry and custom AmpliSeq™ panel and **B.)** MiSeq FGx™ and the ForenSeq™ DNA Signature Prep Kit with Primer Mix A.

CE-based STRs produced more alleles than the Ion S5™ for 2 out of 24 samples and for 8 out of 24 samples when using the MiSeq (Fig. 5.7). For the less compromised remains (embalmed, cremated, and thermally degraded), all methods (CE and MPS) were comparable based on the common loci among the three systems. However, for the severely degraded remains (decomposed), the systems demonstrated variable results. In general, many of the burned remains were comparable for the three systems, but CE results were slightly lower. For the decomposed remains, CE results showed a lower profile completeness for most of the samples than the Ion S5™, but for 2 samples, CE produced results when the other two systems did not (Fig. 5.7). It is possible that these decomposed samples contained PCR inhibitors, affecting the MPS chemistries when maximum volume (15 µL) was amplified. Although the common loci of CE-based STRs are comparable to the MPS results, except for decomposed remains, MPS panels provided more information because they contain more markers (35 STR markers for the Ion S5™ and 58 STRs for the MiSeq vs 24 STRs in GlobalFiler®, and many SNPs). Although, 2 decomposed samples produced almost no results for MPS, CE-based typing was able to recover >30% of alleles (Fig. 5.7).

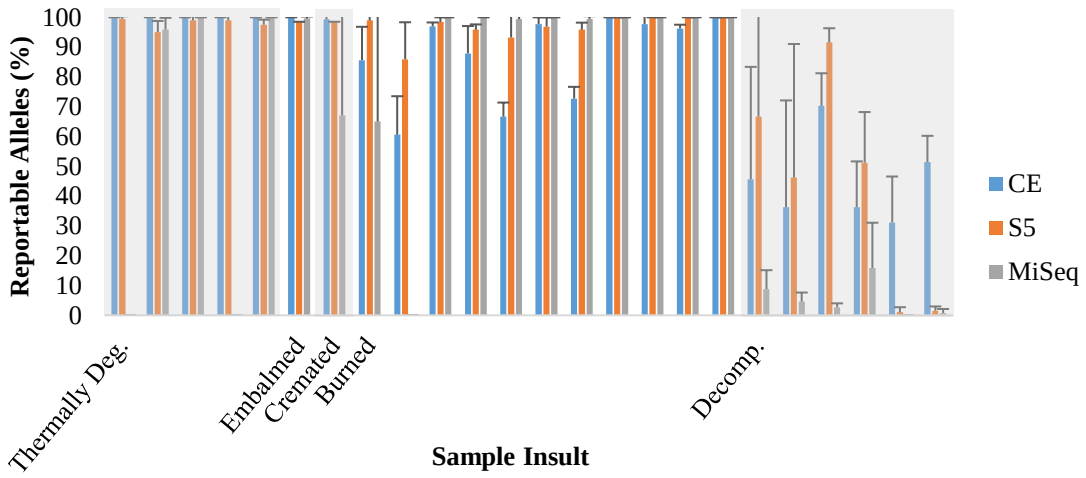


Figure 5.7 Profile completeness across 24 challenging human remains assessed with the GlobalFiler® PCR Amplification Kit using capillary electrophoresis, Precision ID chemistry and a custom AmpliSeq™ panel on the Ion S5™, and the ForenSeq™ DNA Signature Prep Kit on the MiSeq FGx™. Data presented as average + SD (N = 3).

The success of typing 20 core CODIS loci was compared among the three platforms (CE, Ion S5™, and MiSeq). CE-generated STRs produced alleles ranging from 8 ± 3 to 40 (full profiles), with all but 3 samples yielding >50% of alleles. Of the 24 samples processed, half produced a full profile for the 20 CODIS loci (Fig. 5.8). Samples sequenced using the Ion S5™ system generated profiles ranging from 0 to 40 alleles, with all but three samples producing >50% of alleles. Two of the three samples producing <50% alleles generated no profile (both decomposed samples). CE-generated data produced 5 profiles more complete than the Ion S5™ using the CODIS loci (Fig. 5.8). Using the MiSeq system, profiles ranged from 0 to 40 alleles, with 16/24 producing >50% of alleles. Similar to the Ion S5™, only two samples failed to produce any alleles (one decomposed and one thermally degraded). However, 11/24 CE-generated profiles showed more alleles for the CODIS loci. (Fig. 5.8). CE results demonstrated 14/24 samples (Ion S5™) and 19/24 samples (MiSeq) with equal

or greater profile completeness than the MPS systems. Additionally, 12/24 samples produced a full profile for all 3 platforms evaluated (Fig. 5.8).

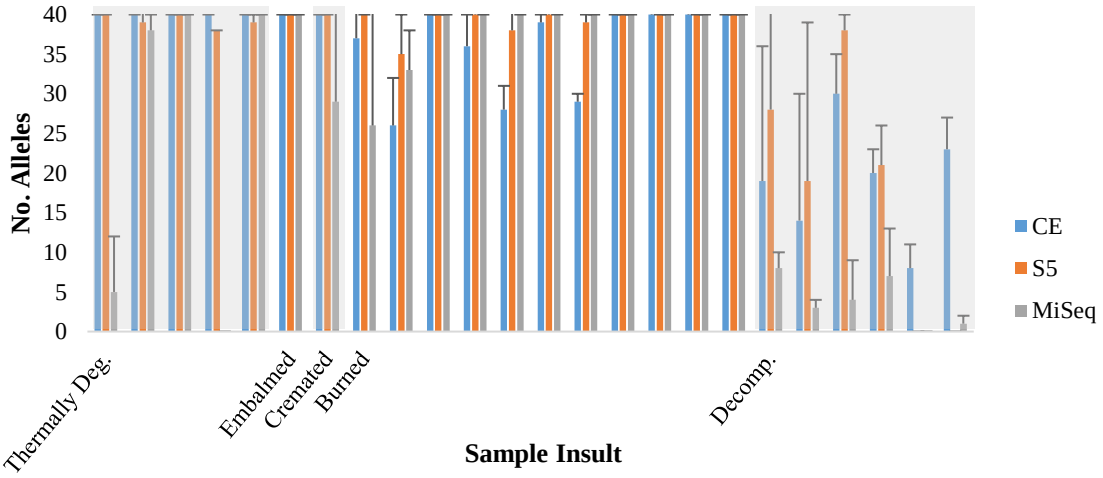


Figure 5.8 Profile completeness (number of alleles) of the 20 core CODIS STRs using CE, the Ion S5™, and the MiSeq FGx™ systems for 24 challenging human remains samples. Data presented as average + SD (N = 3).

Read Depth

Using the Ion S5™, mean read depth of STRs ranged from 19X ± 22X to 53648X ± 7873X, averaging ~17350X across all samples (Fig. 5.9A). The embalmed samples produced the highest mean read depth (53648X ± 7873X), thermally degraded and burned samples produced similar mean read depths, while the decomposed samples produced the lowest values (19X ± 22X to 7410X ± 6563X) (Fig 5.9A). For SNPs, mean read depth ranged from 2848X ± 2378X to 164801X ± 156816X, averaging ~74050X across all samples (Fig. 5.9A). Embalmed and cremated samples produced mean read depths >100000X, 162982X ± 23594X and 100850X ± 24823X, respectively, and decomposed samples produced the lowest mean read depths ranging from 2848X ± 2378X to 84588X ± 27600X (Fig. 5.9A).

For the MiSeq, mean read depth ranged from 0X to 202013X ± 23779X for STRs (Fig. 5.9B). Burned samples produced the highest read depth (3402X ± 2919X to 202013X to 23779X), with the decomposed remains ranging from 0X to 234X ± 298X with the lowest read depth (Fig. 5.9B). For SNPs, mean read depth ranged from 0X to 120061X ± 2690X (Fig. 5.9B). SNP mean read depth demonstrated a similar pattern to that of STRs with burned samples producing the highest read depth (2141X ± 1697X to 120061X ± 2690X), followed by the remaining sample insults and 0X to 109X ± 89X for decomposed samples (Fig. 5.9B).

In general, both platforms performed well and produced high sample read depth. Overall, Ion S5™ SNPs produced higher read depth than STRs (74050X vs 17344X) for every sample. Both STRs and SNPs demonstrated proportional read depth across all samples types. The MiSeq STRs produced higher mean read depths than SNPs (70568X vs 31184X) for most samples. Only two samples produced higher mean read depth for SNPs than STRs, both decomposed remains. The MiSeq showed a large increase in mean read depth for burned samples and one thermally degraded sample for both STRs and SNPs. All other samples produced very low read depth compared to the burned samples, which may be due to run variability or a lower ability to type such samples.

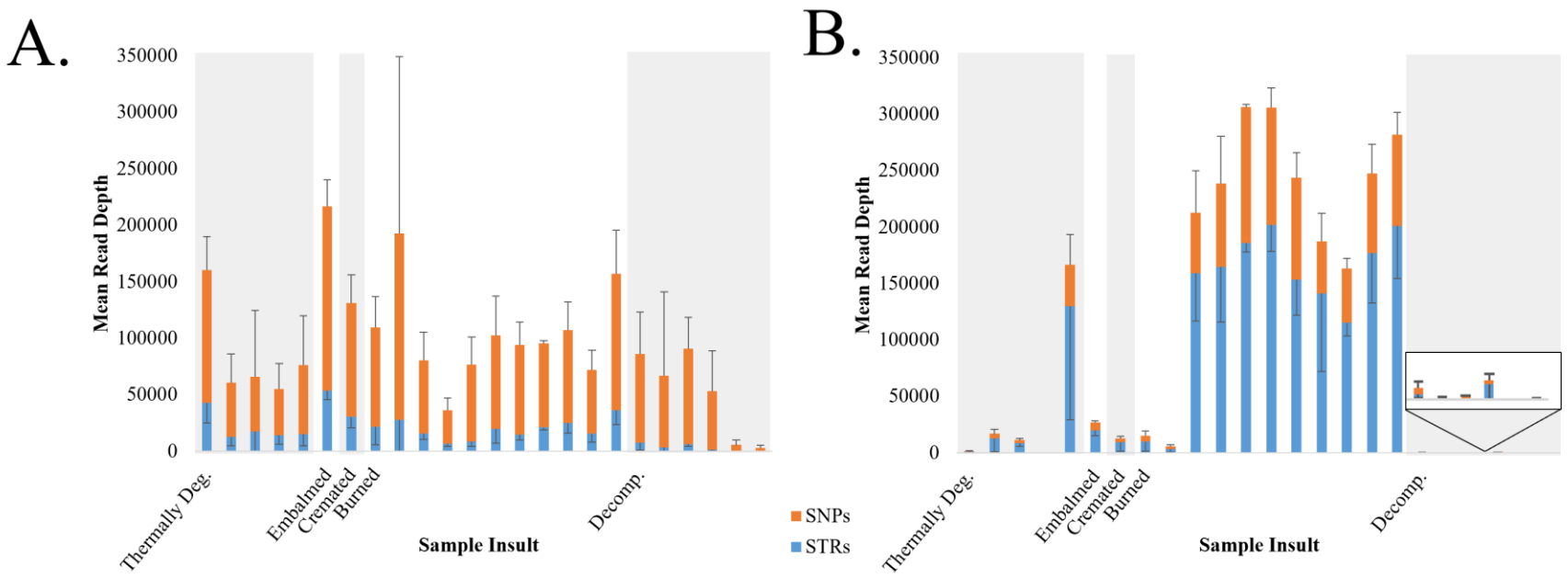


Figure 5.9 Mean read depth of STRs and SNPs sequencing using the **A.)** Ion S5™ and **B.)** MiSeq FGx™ for 24 challenging human remains samples. Data presented as average ± SD (triplicate). Minus error bars represent STR data and plus error bars represent SNP data. SNP mean read depth is determined by subtracting STR mean read depth.

Heterozygote Balance

Due to severe allelic imbalance (dropout) reads with less than 5X coverage were treated as dropout alleles resulting in a heterozygote balance of 0%. STRs produced using both platforms and SNPs produced on the MiSeq system resulted in multiple samples unable to calculate heterozygote balance. Using the Ion S5™ system STR heterozygote balance calculations could not be calculated for two samples (both decomposed) due to severe allelic dropout. However, all SNP samples on the Ion S5™ generated data that could be used for heterozygote balance calculations. For the MiSeq, five samples (one thermally degraded and four decomposed samples) demonstrated substantial allele dropout and heterozygote balance could not be calculated.

For Ion S5™ STRs, heterozygote balance ranged from 0% to 81% ± 14%, averaging ~65% across all samples (Fig. 5.10A). However, most samples generated a heterozygote balance >70%. Out of 24 samples, 8 samples produced <70% heterozygote balance: two burned samples, and all decomposed samples. Thermally degraded teeth samples produced the highest heterozygote balance ranging from 76% ± 21% to 81% ± 14%. Embalmed and cremated samples produced comparable heterozygote balance of 77% ± 14% and 74% ± 19%, respectively. Burned samples resulted in heterozygote balance values ranging from 55% ± 25% to 77% ± 13-17%, and decomposed remains produced the least balanced profiles ranging from 0% to 67% ± 21% (Fig. 5.10A). With Ion S5™ SNPs, heterozygote balance ranged from 32% ± 30% to 86% ± 2-4%, averaging ~79% across all samples (Fig. 5.10A). All but two samples (both decomposed) generated an average heterozygote balance >70%. Similar to STRs, thermally degraded samples produced the highest average heterozygote balance (83% ± 1% to 86% ± 4%), with the lowest

heterozygote balances resulting from decomposed human remains, ranging from 32% ± 30% to 85% ± 1% (Fig. 5.10A).

Heterozygote balance for STRs on the MiSeq ranged from 0% to 82% ± 11% (Fig. 5.10B), averaging ~50% across all samples. Similar to profile completeness and read depth, burned samples generated the highest heterozygote balance ranging from 51% ± 40% to 82% ± 11%, while thermally degraded, embalmed, and cremated remains produced comparable heterozygote balances. Decomposed remains resulted in the lowest heterozygote balances ranging between 0% to 6% ± 22% (Fig. 5.10B). For SNPs, heterozygote balance ranged from 0% to 88% ± 10%, averaging ~51% across all samples (Fig. 5.10B). Analogous to STRs, SNPs demonstrated a similar trend with heterozygote balance (Fig. 5.10B).

In general, for the Ion S5™, the majority of samples showed average heterozygote balances of >70%. Heterozygote balance averaged ~15% higher for SNPs than STRs. SNPs also demonstrated fewer samples with a heterozygote balance <70% compared with STRs (2 vs 8 samples). All samples except decomposed remains resulted in good heterozygote balance for SNPs. Overall, when sequenced on the MiSeq, just under half of the STR and SNP profiles demonstrated heterozygote balances <70%. For both STRs and SNPS, the burned samples demonstrated good heterozygous balance, while decomposed samples consistently demonstrated poor balance (<10%), and variability in all other samples.

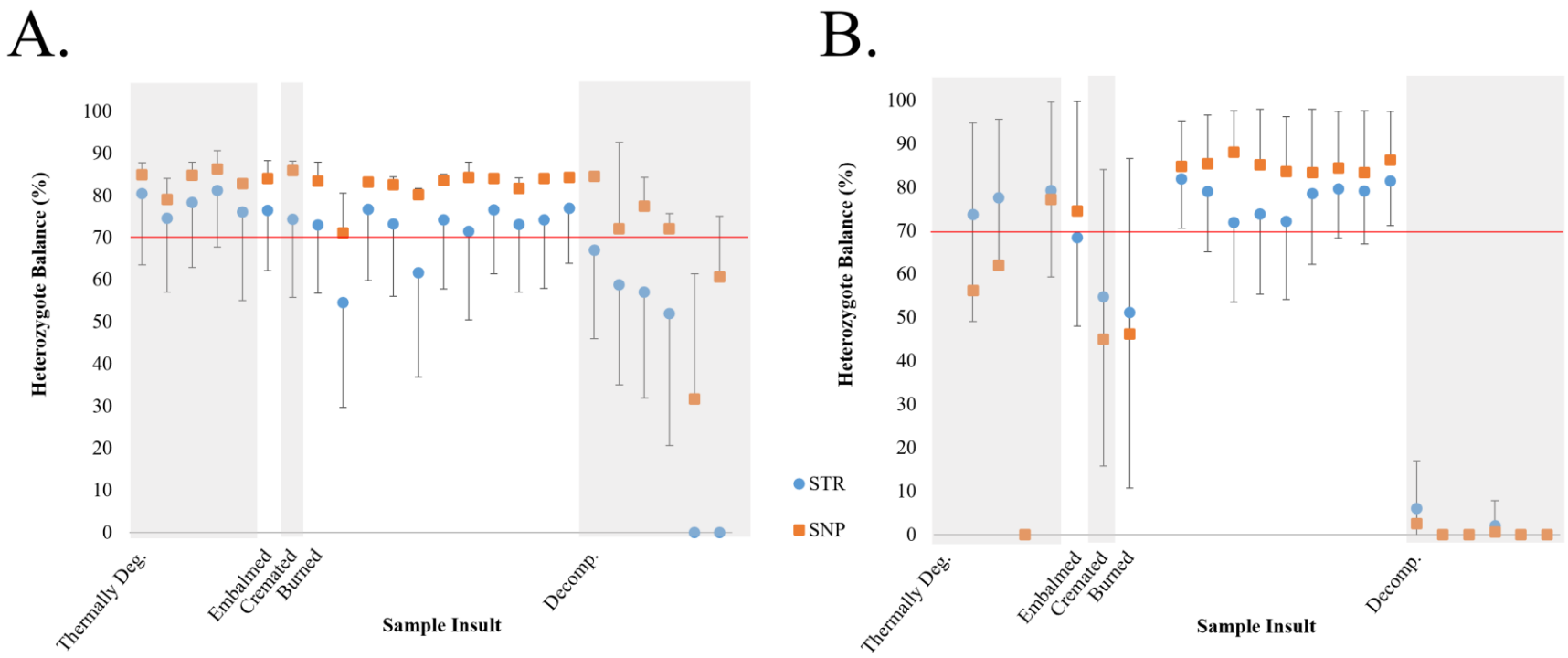


Figure 5.10 Heterozygote balance of STRs and SNPs sequenced using the **A.)** Ion S5™ and **B.)** MiSeq FGx™ for 24 challenging human remains samples. Data presented as average ± SD (N = 3). Minus error bars represent the STR data and plus error bars represent the SNP data.

Conclusions

Overall, MPS generated genetic data from challenged samples and provided more genetic data in 22 samples compared with the CE-based kit. Furthermore, a greater number of alleles will translate in greater power of discrimination. Although CE produced a usable DNA profile for identification purposes, based only on the 20 CODIS core loci for some more difficult samples, the greater number of loci included in MPS multiplexes allowed for more genetic information to be obtained from most samples barring the decomposed remains. Results suggest that MPS may recover more probative information from most samples, but CE-based methods were more robust for identifying skeletal samples. CE chemistry has been substantially developed over the past 25 years, while MPS kits for forensic applications have been around for less than five years. However, improvement in MPS panel design and chemistries could enhance performance.

Acknowledgements

This project was supported by Award No. NIJ 2015-DN-BX-K066, awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect those of the Department of Justice.

References

1. da Silva LAF, Vilaça W, Azevedo D, Majella G, Silva IF, Silva BF (2009) Missing and unidentified persons database. Forensic Sci Int Genet Suppl Ser 2 (1):255-257. doi:10.1016/j.fsigss.2009.08.090

2. Kovras I, Robins S (2016) Death as the border: Managing missing migrants and unidentified bodies at the EU's Mediterranean frontier. Polit Geogr 55:40-49. doi:10.1016/j.polgeo.2016.05.003

3. Brayley H, Cockbain E, Laycock G (2011) The value of crime scripting: Deconstructing internal child sex trafficking. Policing: A Journal of Policy and Practice 5 (2):132-143

4. Alvarez-Cubero MJ, Saiz M, Martinez-Gonzalez LJ, Alvarez JC, Eisenberg AJ, Budowle B, Lorente JA (2012) Genetic identification of missing persons: DNA analysis of human remains and compromised samples. Pathobiology 79 (5):228-238. doi:10.1159/000334982

5. Loreille OM, Diegoli TM, Irwin JA, Coble MD, Parsons TJ (2007) High efficiency DNA extraction from bone by total demineralization. Forensic Sci Int Genet 1 (2):191-195. doi:10.1016/j.fsigen.2007.02.006

6. Ferreira ST, Paula KA, Maia FA, Moraes AV (2013) A comparative study between two protocols for DNA extraction from bones. Forensic Sci Int Genet Suppl Ser 4 (1):e374-e375

7. Rothe J, Nagy M (2016) Comparison of two silica-based extraction methods for DNA isolation from bones. Leg Med (Tokyo) 22:36-41. doi:10.1016/j.legalmed.2016.07.008

8. Karija Vlahovic M, Kubat M (2012) DNA extraction method from bones using Maxwell(R) 16. Leg Med (Tokyo) 14 (5):272-275. doi:10.1016/j.legalmed.2012.04.004

9. Hartman D, Drummer O, Eckhoff C, Scheffer JW, Stringer P (2011) The contribution of DNA to the disaster victim identification (DVI) effort. Forensic Sci Int 205 (1-3):52-58. doi:10.1016/j.forsciint.2010.09.024

10. Alonso A, Andelinovic S, Martin P, Sutlovic D, Erceg I, Huffine E, de Simon LF, Albarran C, Definis-Gojanovic M, Fernandez-Rodriguez A, Garcia P, Drmic I, Rezic B, Kuret S, Sancho M, Primorac D (2001) DNA typing from skeletal remains: evaluation of multiplex and megaplex STR systems on DNA isolated from bone and teeth samples. Croat Med J 42 (3):260-266

11. Mundorff AZ, Bartelink EJ, Mar-Cash E (2009) DNA preservation in skeletal elements from the World Trade Center disaster: recommendations for mass fatality management. J Forensic Sci 54 (4):739-745. doi:10.1111/j.1556-4029.2009.01045.x

12. Freire-Aradas A, Fondevila M, Kriegel AK, Phillips C, Gill P, Prieto L, Schneider PM, Carracedo A, Lareu MV (2012) A new SNP assay for identification of highly degraded human DNA. Forensic Sci Int Genet 6 (3):341-349. doi:10.1016/j.fsigen.2011.07.010

13. Guo F, Yu J, Zhang L, Li J (2017) Massively parallel sequencing of forensic STRs and SNPs using the Illumina® ForenSeq™ DNA Signature Prep Kit on the MiSeq FGx™ Forensic Genomics System. Forensic Sci Int Genet 31:135-148

14. Wang Z, Zhou D, Wang H, Jia Z, Liu J, Qian X, Li C, Hou Y (2017) Massively parallel sequencing of 32 forensic markers using the Precision ID GlobalFiler NGS STR Panel

and the Ion PGM System. Forensic Sci Int Genet 31:126-134. doi:10.1016/j.fsigen.2017.09.004

15. Budowle B, van Daal A (2008) Forensically relevant SNP classes. Biotechniques 44 (5):603-608, 610. doi:10.2144/000112806
16. Novroski NMM, Woerner AE, Budowle B (2018) Potential highly polymorphic short tandem repeat markers for enhanced forensic identity testing. Forensic Sci Int Genet 37:162-171. doi:10.1016/j.fsigen.2018.08.011
17. Gettings KB, Aponte RA, Vallone PM, Butler JM (2015) STR allele sequence variation: Current knowledge and future issues. Forensic Sci Int Genet 18:118-130. doi:10.1016/j.fsigen.2015.06.005
18. Gettings KB, Kiesler KM, Faith SA, Montano E, Baker CH, Young BA, Guerrieri RA, Vallone PM (2016) Sequence variation of 22 autosomal STR loci detected by next generation sequencing. Forensic Sci Int Genet 21:15-21. doi:10.1016/j.fsigen.2015.11.005
19. King JL, Churchill JD, Novroski NMM, Zeng X, Warshauer D, Seah LH, Budowle B (2018) Increasing the discrimination power of ancestry-and identity-informative SNP loci within the ForenSeq™ DNA Signature Prep Kit. Forensic Sci Int Genet
20. Loreille OM, Parr RL, McGregor KA, Fitzpatrick CM, Lyon C, Yang DY, Speller CF, Grimm MR, Grimm MJ, Irwin JA, Robinson EM (2010) Integrated DNA and fingerprint analyses in the identification of 60-year-old mummified human remains discovered in an Alaskan glacier. J Forensic Sci 55 (3):813-818. doi:10.1111/j.1556-4029.2010.01356.x

21. Zeng X, Elwick K, Mayes C, Takahashi M, King JL, Gangitano D, Budowle B, Hughes-Stamm S (2018) Assessment of impact of DNA extraction methods on analysis of human remain samples on massively parallel sequencing success. Int J Leg Med. doi:10.1007/s00414-018-1955-9

22. Thermo Fisher Scientific (2018) AutoMate *Express*™ Instrument User Guide. F edn. Thermo Fisher Scientific,

23. Thermo Fisher Scientific (2017) Quantifiler™ HP and Trio DNA Quantification Kits: User Guide. F edn. Thermo Fisher Scientific,

24. Thermo Fisher Scientific (2015) GlobalFiler™ PCR Amplification Kit. Thermo Fisher Scientific, Thermo Fisher Scientific

25. Illumina (2015) ForenSeq DNA Signature Prep Reference Guide. vol 15049528, v01 edn., San Diego, CA

26. King JL, Wendt FR, Sun J, Budowle B (2017) STRait Razor v2s: Advancing sequence-based STR allele reporting and beyond to other marker systems. Forensic Sci Int Genet 29:21-28. doi:10.1016/j.fsigen.2017.03.013

27. Elwick K, Mayes C, Hughes-Stamm S (2018) Comparative Sensitivity and Inhibitor Tolerance of GlobalFiler® PCR Amplification and Investigator® 24plex QS Kits for Challenging Samples. Leg Med

28. Parys-Proszek A, Wrobel M, Marcinska M, Kupiec T (2018) Dual amplification strategy for improved efficiency of forensic DNA analysis using NGM Detect, NGM or Globalfiler kits. Forensic Sci Int Genet 35:46-49. doi:10.1016/j.fsigen.2018.03.018

29. Tucker VC, Hopwood AJ, Sprecher CJ, McLaren RS, Rabbach DR, Ensenberger MG, Thompson JM, Storts DR (2011) Developmental validation of the PowerPlex((R)) ESI

16 and PowerPlex((R)) ESI 17 Systems: STR multiplexes for the new European standard. Forensic Sci Int Genet 5 (5):436-448. doi:10.1016/j.fsigen.2010.09.004

CHAPTER VI

Conclusions

Missing persons' cases, migrant deaths, mass disasters, and mass graves due to military conflicts or natural disasters are problems faced worldwide. Recovery of these remains requires a sizable effort from law enforcement, forensic scientists, and specialty recovery teams. When remains are recovered or a MFI occurs it is of the utmost importance to identify and repatriate the remains to the deceased's loved ones. The condition of the human remains may range from pristine to highly decomposed, skeletonized, or burnt causing the DNA in those tissues to be damaged, degraded, and/or inhibited from the environment/climate (temperature, humidity), the physical impact of the incident, or preservatives used. Common PCR inhibitors include those naturally found in tissues such as hematin, melanin, collagen, and calcium, as well as those found in the environment such as humic acid in soil. STRs are currently the gold standard in the recovery of a DNA profile, but there are some pitfalls to using STRs such as limited available space per dye channel and the limited number of samples that can be processed at one time. Because massively parallel sequencing has many advantages and can improve on CE technologies, it may be a suitable replacement in the future for various forensic applications, including human identification. MPS has the ability to improve powers of discrimination because of the simultaneous analysis of marker systems (STRs, iiSNPs, aiSNPs, microhaplotypes, mtDNA, etc.), increase throughput and sample information, deconvolute mixtures, and has many other applications. However, MPS platforms and chemistries have not been thoroughly evaluated and optimized compared to standard CE chemistries over the last 25

years, and little data has been published to demonstrate the compatibilities of MPS with commonly employed sample preparation chemistries.

Prior to the rapid onset of MPS, CE-based megaplex STR kits were designed in an effort to increase the number of markers in a kit and power of discrimination for identification. In addition, STR kits have been continually improved over the last 30 years to be more sensitive, tolerant to inhibitors, and more robust in general. CODIS also increased the number of core loci from 13 to 20 making it pertinent that STR kits be upgraded. Two kits (GlobalFiler® PCR Amplification and Investigator® 24plex QS kits) were evaluated in this study for sensitivity and tolerance to PCR inhibitors (five concentrations of five inhibitors). The GlobalFiler® kit appeared to be more sensitive than the Investigator® kit generating more complete STR profiles down to 7.8 pg and demonstrating higher peak heights and peak height ratios across the entire sample range. However, the Investigator® kit appeared to be more tolerant to PCR inhibitors than GlobalFiler®, indicating a higher number of reportable alleles, more balanced profiles, and less dropout at two DNA input concentrations (1 ng and 0.1 ng). In general, the number of reportable alleles decreased as the inhibitor concentrations increased. However, both CE chemistries were determined to be comparable in performance for traditional analysis of DNA.

Because MPS is a newer technology than CE, the chemistries have not benefited from the same level of refinement as CE-based products, and therefore optimization of assays for sequencing difficult samples such as low level, inhibited, and/or degraded samples for forensic purposes is warranted. The two most common forensic sequencing platforms available at the time of this study (Ion™ PGM and MiSeq FGx™) and their

complementary identification assay (Ion AmpliSeq™ Identity Panel and Primer Mix A from the ForenSeq™ DNA Signature Prep Kit) were evaluated for tolerance to PCR inhibitors prior to extraction. To evaluate these two chemistries and platforms, inhibitor concentrations from the STR megaplex evaluation were used as a starting point and then modified with each run to determine the limits of tolerance for the MPS systems. The Ion AmpliSeq™ chemistry and ID panel demonstrated fair tolerance to collagen and calcium inhibitors, but showed high susceptibility to humic acid and hematin. In general, mean read depth and heterozygote balance for SNPs decreased as the inhibitor concentration increased. The ForenSeq™ kit using Primer Mix A demonstrated that STR/SNP success and heterozygote balance was extremely tolerant to hematin and calcium but was greatly affected by melanin. Sample success was not affected by humic acid and collagen until higher inhibitor concentrations were introduced. Between the two platforms, the AmpliSeq™ chemistry was more tolerant to melanin than the ForenSeq™ chemistry, and the ForenSeq™ chemistry was highly tolerant to hematin, whereas the AmpliSeq™ chemistry was severely affected. The ForenSeq™ chemistry was determined to be more tolerant overall to inhibitors than the AmpliSeq™ chemistry. PCR or sequencing error (noise) was also assessed, and it appeared that noise was not affected by the presence of inhibitors; however, the percentage of noise increased as a result of decreasing locus coverage. Noise was also more prominent when sequencing STRs than SNPs. For humic acid, melanin, and hematin, CE-based inhibitor concentrations were too concentrated to provide any sequencing data. Therefore, for humic acid, concentrations were decreased for testing the MPS systems from 50 ng/µL – 250 ng/µL to 5 ng/µL – 25 ng/µL; for melanin, concentrations were decreased from 25 ng/µL – 50 ng/µL to 4 ng/µL – 12 ng/µL; and for

hematin, concentrations were decreased from 300 µM – 1100 µM to 1 µM – 10 µM. However, for collagen and calcium, inhibitor concentrations were increased for defining the limits of tolerance for MPS because even the highest CE inhibitor concentrations were producing full profiles when sequenced with MPS. Consequently, collagen concentrations were increased from 50 ng/µL – 160 ng/µL to 180 ng/µL – 400 ng/µL and calcium concentrations were increased from 250 µM – 850 µM to 350 µM – 1100 µM. Overall, the comparative inhibitor tolerance for the two MPS sequencing chemistries and platforms used for forensics applications was reported.

Further investigation was required to determine if three commercial DNA extraction kits and two total demineralization extraction and purification protocols were equally compatible with CE and two MPS sequencing chemistries (Precision ID custom AmpliSeq™ STR and iiSNP panel and ForenSeq™ DNA Signature Prep Kit Primer Mix A) and platforms (Ion S5™ System and MiSeq FGx™). It was also necessary to establish whether these commonly used extraction and purification chemistries could efficiently remove high amounts of PCR inhibitors from human remains samples prior to MPS. Any negative effects on sequencing could suggest that extraction kits widely used within forensic laboratories may be unsuitable for use prior to downstream sequencing. However, as expected, all extraction methods were deemed suitable for removing high concentrations of PCR inhibitors and producing clean DNA extracts when coupled with CE-based STR typing. Results demonstrated profiles with >90% of alleles and very few samples with peak height ratios <70%. In addition, MPS results demonstrated good mean read depth and heterozygote balance for all samples when extracted with each extraction protocol. Overall, all extraction protocols were effective at removing PCR inhibitors from human remains,

producing high quality DNA profiles via MPS. Because these extraction protocols showed no negative effects on downstream sequencing, most commercial and total demineralization extraction methods should be considered equally compatible with both MPS systems.

After defining the inhibitor tolerance and establishing the compatibility of common DNA extraction methods with both MPS systems, we then evaluated the comparative performance of both MPS platforms to identify challenging human remains. Human remains (bones and teeth) that have been burned, embalmed, cremated, decomposed, and thermally degraded were assessed with traditional CE methods and both MPS technologies. CE results showed that thermally degraded, embalmed, and cremated samples produced the best results (reportable alleles, peak height, and peak height ratios), while burned and decomposed remains generated less complete profiles. The MPS results demonstrated that the Ion S5™ system produced more complete profiles, higher heterozygote balance, and similar mean read depth when compared with the same samples sequenced using the MiSeq FGx™ system. Similar to the CE results, decomposed remains proved to be the most difficult samples to sequence, particularly for the MiSeq FGx™ system. However, the ForenSeq™ primer panel has more STRs and SNPs than the custom AmpliSeq ™ primer panel, so the powers of discrimination may be similar for some samples. Overall, challenging samples still pose a problem for DNA analysis, particularly highly decomposed human remains. Every sample processed using CE produced at least a partial DNA profile, while not all samples sequenced via MPS produced a genetic profile. However, MPS panels include more markers than with CE STR kits. Utilizing the two MPS chemistries, between ~65 and ~190 STR and SNP markers were evaluated compared to 24

possible STR loci used in CE. Because more markers can be interrogated with MPS, there is a high likelihood that even when the proportion of loci/alleles successfully amplified is less than CE, the powers of discrimination may still be higher with the MPS systems.

In summary, the two most common MPS platforms marketed for forensic applications have been evaluated in their abilities to analyze challenging and inhibited samples such as may be encountered in missing persons' cases. Because these MPS chemistries and platforms are still relatively new compared to CE workflows, they are not yet fully optimized for the analysis of challenging human remains, and few studies have been published that explore the limits of their performance. MPS systems will continue to be modified, improved, and expanded to provide the scientific community with better performance. The fields of human identification and forensic intelligence will greatly benefit from these future developments.

REFERENCES

1. (FBI) FBI CODIS and NDIS Fact Sheet - FBI. FBI, U.S. government, U.S. Department of Justice https://www.fbi.gov/services/laboratory/biometric-analysis/codis/codis-and-ndis-fact-sheet.

2. (FBI) FBI Combined DNA Index System (CODIS). FBI, U.S. government, U.S. Department of Justice https://www.fbi.gov/services/laboratory/biometric-analysis/codis.

3. (UNTCHI) UNTCHI UNT Center for Human Identification UNT Center for Human Identification http://untchi.org/.

4. Alaeddini R (2012) Forensic implications of PCR inhibition--A review. Forensic Sci Int Genet 6 (3):297-305. doi:10.1016/j.fsigen.2011.08.006

5. Alaeddini R et al. (2010) Forensic implications of genetic analyses from degraded DNA—a review. Forensic Sci Int Genet 4 (3):148-157

6. Allouche M et al. (2008) Genetic identification of decomposed cadavers using nails as DNA source. Forensic Sci Int Genet 3 (1):46-49. doi:10.1016/j.fsigen.2008.07.008

7. Alonso A et al. (2001) DNA typing from skeletal remains: evaluation of multiplex and megaplex STR systems on DNA isolated from bone and teeth samples. Croat Med J 42 (3):260-266

8. Alonso A et al. (2005) Challenges of DNA profiling in mass disaster investigations. Croat Med J 46 (4):540-548

9. Alvarez-Cubero MJ et al. (2012) Genetic identification of missing persons: DNA analysis of human remains and compromised samples. Pathobiology 79 (5):228-238. doi:10.1159/000334982

10. Amory S et al. (2012) Automatable full demineralization DNA extraction procedure from degraded skeletal remains. Forensic Sci Int Genet 6 (3):398-406. doi:10.1016/j.fsigen.2011.08.004

11. Anđelinović Š et al. (2005) Twelve-year experience in identification of skeletal remains from mass graves. Croat Med J 46 (4):530-539

12. Barbisin M et al. (2009) Developmental validation of the Quantifiler® Duo DNA quantification kit for simultaneous quantification of total human and human male DNA and detection of PCR inhibitors in biological samples. J Forensic Sci 54 (2):305-319

13. Bentley DR et al. (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456 (7218):53-59. doi:10.1038/nature07517

14. Berglund EC et al. (2011) Next-generation sequencing technologies and applications for human genetic history and forensics. Investig Genet 2 (1):23. doi:10.1186/2041-2223-2-23

15. Bertinchamps AJ et al. (1978) Effects of ionizing radiation on DNA.

16. Besecker J et al. (2013) Dynamic passivation with BSA overcomes LTCC mediated inhibition of PCR. Sensors and Actuators B-Chemical 176:118-123. doi:10.1016/j.snb.2012.09.110

17. Bessetti J (2007) An introduction to PCR inhibitors. J Microbiol Meth 28:159-167

18. BiteSize Bio.

19. Bogas V et al. (2011) Comparison of four DNA extraction methods for forensic application. Forensic Sci Int Genet Suppl Ser 3 (1):e194-e195. doi:10.1016/j.fsigss.2011.08.097

20. Bonnet J et al. (2010) Chain and conformation stability of solid-state DNA: implications for room temperature storage. Nucleic Acids Res 38 (5):1531-1546. doi:10.1093/nar/gkp1060

21. Bornman DM et al. (2012) Short-read, high-throughput sequencing technology for STR genotyping. Biotech Rapid Dispatches 2012:1-6

22. Borsting C et al. (2014) Evaluation of the Ion Torrent HID SNP 169-plex: A SNP typing assay developed for human identification by second generation sequencing. Forensic Sci Int Genet 12:144-154. doi:10.1016/j.fsigen.2014.06.004

23. Børsting C, Morling N (2015) Next generation sequencing and its applications in forensic genetics. Forensic Sci Int Genet 18:78-89

24. Bottino CG et al. (2015) STR genotyping using ion torrent PGM and STR 24-plex system: Performance and data interpretation. Forensic Sci Int Genet Suppl Ser 5:E325-E326. doi:10.1016/j.fsigss.2015.09.129

25. Brayley H et al. (2011) The value of crime scripting: Deconstructing internal child sex trafficking. Policing: A Journal of Policy and Practice 5 (2):132-143

26. Brevnov MG et al. (2009) Developmental validation of the PrepFiler™ forensic DNA extraction kit for extraction of genomic DNA from biological samples. J Forensic Sci 54 (3):599-607

27. Brough AL et al. (2015) The basics of disaster victim identification. Journal of Forensic Radiology and Imaging 3 (1):29-37. doi:10.1016/j.jofri.2015.01.002

28. Budowle B et al. (2005) Forensic aspects of mass disasters: strategic considerations for DNA-based human identification. Leg Med (Tokyo) 7 (4):230-243. doi:10.1016/j.legalmed.2005.01.001

29. Budowle B, van Daal A (2008) Forensically relevant SNP classes. Biotechniques 44 (5):603-608, 610. doi:10.2144/000112806

30. Butler JM (2005) Constructing STR multiplex assays. In: Forensic DNA Typing Protocols. Springer, pp 53-65

31. Calacal GC et al. (2015) Comparing different post-mortem human samples as DNA sources for downstream genotyping and identification. Forensic Sci Int Genet 19:212-220. doi:10.1016/j.fsigen.2015.07.017

32. Chaitanya L et al. (2017) Bringing colour back after 70 years: Predicting eye and hair colour from skeletal remains of World War II victims using the HIrisPlex system. Forensic Sci Int Genet 26:48-57. doi:10.1016/j.fsigen.2016.10.004

33. Chen CM et al. (2014) Identification of conserved and polymorphic STRs for personal genomes. BMC Genomics 15 Suppl 10 (10):S3. doi:10.1186/1471-2164-15-S10-S3

34. Churchill JD et al. (2015) Blind study evaluation illustrates utility of the Ion PGM™ system for use in human identity DNA typing. Croat Med J 56 (3):218-229

35. Churchill JD et al. (2017) Working towards implementation of whole genome mitochondrial DNA sequencing into routine casework. Forensic Sci Int Genet Suppl Ser 6:E388-E389. doi:10.1016/j.fsigss.2017.09.167

36. Combs LG et al. (2015) The effects of metal ion PCR inhibitors on results obtained with the Quantifiler((R)) Human DNA Quantification Kit. Forensic Sci Int Genet 19:180-189. doi:10.1016/j.fsigen.2015.06.013

37. Life Technologies Corporation (2010) PrepFiler Express™ and PrepFiler Express BTA™ Forensic DNA Extraction Kits. Life Technologies Corporation,

38. Promega Corporation (2013) Technical Bulletin: DNA IQ™ System - Database Protocol. TB297 edn. Promega Corporation, Madison, WI

39. Promega Corporation (2013) Technical Bulletin: DNA IQ™ System - Small Sample Casework Protocol. TB296 edn. Promega Corporation, Madison, WI

40. Promega Corporation (2015) Technical Bulletin: Tissue and Hair Extraction Kit (for use with DNA IQ™). TB307 edn. Promega Corporation, Madison, WI

41. Cortellini V et al. (2015) Evaluation of mtDNA stability across the maternal line: A study on three generations in two families. Forensic Sci Int Genet Suppl Ser 5:E19-E20. doi:10.1016/j.fsigss.2015.09.008

42. da Silva LAF et al. (2009) Missing and unidentified persons database. Forensic Sci Int Genet Suppl Ser 2 (1):255-257. doi:10.1016/j.fsigss.2009.08.090

43. Daniel R et al. (2015) A SNaPshot of next generation sequencing for forensic SNP analysis. Forensic Sci Int Genet 14:50-60. doi:10.1016/j.fsigen.2014.08.013

44. Davis C et al. (2015) Sequencing the hypervariable regions of human mitochondrial DNA using massively parallel sequencing: Enhanced data acquisition for DNA samples encountered in forensic testing. Leg Med (Tokyo) 17 (2):123-127. doi:10.1016/j.legalmed.2014.10.004

45. Davoren J et al. (2007) Highly effective DNA extraction method for nuclear short tandem repeat testing of skeletal remains from mass graves. Croat Med J 48 (4):478-485

46. de la Puente M et al. (2017) Evaluation of the Qiagen 140-SNP forensic identification multiplex for massively parallel sequencing. Forensic Sci Int Genet 28:35-43. doi:10.1016/j.fsigen.2017.01.012

47. DeNiro MJ, Weiner S (1988) Chemical, enzymatic and spectroscopic characterization of "collagen" and other organic fractions from prehistoric bones. Geochi Cosmo Acta 52 (9):2197-2206

48. Dhanardhono T et al. (2013) DNA profiling of disaster victim identification in Trenggalek shipwreck case. FForensic Sci Int Genet Suppl Ser 4 (1):e5-e6. doi:10.1016/j.fsigss.2013.10.002

49. Diegoli TM et al. (2012) An optimized protocol for forensic application of the PreCR Repair Mix to multiplex STR amplification of UV-damaged DNA. Forensic Sci Int Genet 6 (4):498-503. doi:10.1016/j.fsigen.2011.09.003

50. dos Santos Pereira J et al. (2016) Association of the XPD and XRCC3 gene polymorphisms with oral squamous cell carcinoma in a Northeastern Brazilian population: A pilot study. Arch Oral Biol 64:19-23. doi:10.1016/j.archoralbio.2015.12.004

51. Edson SM et al. (2004) Naming the dead-confronting the realities of the rapid identification of degraded skeletal remains. 16 (1):63-88

52. Eduardoff M et al. (2015) Inter-laboratory evaluation of SNP-based forensic identification by massively parallel sequencing using the Ion PGM. Forensic Sci Int Genet 17:110-121. doi:10.1016/j.fsigen.2015.04.007

53. Edwards A et al. (1991) DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet 49 (4):746-756

54. Elena S et al. (2016) Revealing the challenges of low template DNA analysis with the prototype Ion AmpliSeq Identity panel v2.3 on the PGM Sequencer. Forensic Sci Int Genet 22:25-36. doi:10.1016/j.fsigen.2015.07.011

55. Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nat Rev Genet 5 (6):435-445. doi:10.1038/nrg1348

56. Elwick K et al. (2018) Comparative Sensitivity and Inhibitor Tolerance of GlobalFiler® PCR Amplification and Investigator® 24plex QS Kits for Challenging Samples. Leg Med

57. Elwick K et al. (2018) Comparative tolerance of two massively parallel sequencing systems to common PCR inhibitors. Int J Legal Med 132 (4):983-995. doi:10.1007/s00414-017-1693-4

58. Ensenberger MG et al. (2010) Developmental validation of the PowerPlex 16 HS System: an improved 16-locus fluorescent STR multiplex. Forensic Sci Int Genet 4 (4):257-264. doi:10.1016/j.fsigen.2009.10.007

59. Faber KL et al. (2013) PCR inhibitor removal using the NucleoSpin(R) DNA Clean-Up XS kit. Forensic Sci Int Genet 7 (1):209-213. doi:10.1016/j.fsigen.2012.06.013

60. Ferreira ST et al. (2013) A comparative study between two protocols for DNA extraction from bones. Forensic Sci Int Genet Suppl Ser 4 (1):e374-e375

61. Ferreira STG et al. (2013) Cartilage and phalanges from hallux: alternative sources of samples for DNA typing in disaster victim identification (DVI). A comparative study. Forensic Sci Int Genet Suppl Ser 4 (1):e366-e367

62. Flores S et al. (2014) Internal validation of the GlobalFiler Express PCR Amplification Kit for the direct amplification of reference DNA samples on a high-throughput

automated workflow. Forensic Sci Int Genet 10:33-39. doi:10.1016/j.fsigen.2014.01.005

63. Fondevila M et al. (2008) Challenging DNA: Assessment of a range of genotyping approaches for highly degraded forensic samples. Forensic Sci Int Genet Suppl Ser 1 (1):26-28. doi:10.1016/j.fsigss.2007.10.057
64. Fondevila M et al. (2013) Revision of the SNPforID 34-plex forensic ancestry test: Assay enhancements, standard reference sample genotypes and extended population studies. Forensic Sci Int Genet 7 (1):63-74. doi:10.1016/j.fsigen.2012.06.007
65. Freese E, Cashel M (1964) Crosslinking of Deoxyribonucleic Acid by Exposure to Low Ph. Biochim Biophys Acta 91 (1):67-77
66. Freire-Aradas A et al. (2012) A new SNP assay for identification of highly degraded human DNA. Forensic Sci Int Genet 6 (3):341-349. doi:10.1016/j.fsigen.2011.07.010
67. Ge J et al. (2010) Haplotype block: a new type of forensic DNA markers. Int J Leg Med 124 (5):353-361. doi:10.1007/s00414-009-0400-5
68. Gettings KB et al. (2015) STR allele sequence variation: Current knowledge and future issues. Forensic Sci Int Genet 18:118-130. doi:10.1016/j.fsigen.2015.06.005
69. Gettings KB et al. (2016) Sequence variation of 22 autosomal STR loci detected by next generation sequencing. Forensic Sci Int Genet 21:15-21. doi:10.1016/j.fsigen.2015.11.005
70. Gettings KB et al. (2014) A 50-SNP assay for biogeographic ancestry and phenotype prediction in the U.S. population. Forensic Sci Int Genet 8 (1):101-108. doi:10.1016/j.fsigen.2013.07.010

71. Gill P (2001) An assessment of the utility of single nucleotide polymorphisms (SNPs) for forensic purposes. Int J Leg Med 114 (4-5):204-210

72. Goffin C, Verly WG (1983) Interstrand DNA crosslinks due to AP (apurinic/apyrimidinic) sites. FEBS letters 161 (1):140-144

73. Goodwin S et al. (2016) Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet 17 (6):333-351. doi:10.1038/nrg.2016.49

74. Goodwin WH, Simmons T (2013) Disaster Victim Identification. In: Enc Forensic Sci. Academic Press, Waltham, pp 332-338. doi:10.1016/b978-0-12-382165-2.00061-1

75. Gouveia N et al. (2017) Massively parallel sequencing of forensic samples using precision ID mtDNA whole genome panel on the ion S5™ system. Forensic Sci Int Genet Suppl Ser 6:e167-e168. doi:10.1016/j.fsigss.2017.09.057

76. Green RL et al. (2013) Developmental validation of the AmpFlSTR(R) NGM SElect PCR Amplification Kit: A next-generation STR multiplex with the SE33 locus. Forensic Sci Int Genet 7 (1):41-51. doi:10.1016/j.fsigen.2012.05.012

77. Greenspoon SA et al. (1998) QIAamp spin columns as a method of DNA isolation for forensic casework. J Forensic Sci 43 (5):1024-1030

78. Guo F et al. (2017) Massively parallel sequencing of forensic STRs and SNPs using the Illumina® ForenSeq™ DNA Signature Prep Kit on the MiSeq FGx™ Forensic Genomics System. Forensic Sci Int Genet 31:135-148

79. Guo F et al. (2016) Next generation sequencing of SNPs using the HID-Ion AmpliSeq Identity Panel on the Ion Torrent PGM platform. Forensic Sci Int Genet 25:73-84. doi:10.1016/j.fsigen.2016.07.021

80. Habib M et al. (2017) Direct amplification of biological evidence and DVI samples using the Qiagen Investigator 24plex GO! Kit. Forensic Sci Int Genet Suppl Ser 6:E208-E210. doi:10.1016/j.fsigss.2017.09.079

81. Hall A, Ballantyne J (2004) Characterization of UVC-induced DNA damage in bloodstains: forensic implications. Anal Bioanal Chem 380 (1):72-83. doi:10.1007/s00216-004-2681-3

82. Hall A et al. (2014) Assessment of DNA damage induced by terrestrial UV irradiation of dried bloodstains: forensic implications. Forensic Sci Int Genet 8 (1):24-32. doi:10.1016/j.fsigen.2013.06.010

83. Hartman D et al. (2011) The contribution of DNA to the disaster victim identification (DVI) effort. Forensic Sci Int 205 (1-3):52-58. doi:10.1016/j.forsciint.2010.09.024

84. Heather JM, Chain B (2016) The sequence of sequencers: The history of sequencing DNA. Genomics 107 (1):1-8. doi:10.1016/j.ygeno.2015.11.003

85. Hebda LM, Foran DR (2015) Assessing the Utility of Soil DNA Extraction Kits for Increasing DNA Yields and Eliminating PCR Inhibitors from Buried Skeletal Remains. J Forensic Sci 60 (5):1322-1330. doi:10.1111/1556-4029.12878

86. Hofreiter M et al. (2001) ancient DNA. Nat Rev Genet 2 (5):353

87. Holland MM et al. (2011) Second generation sequencing allows for mtDNA mixture deconvolution and high resolution detection of heteroplasmy. Croat Med J 52 (3):299-313

88. Houck MM (2013) The National Missing and Unidentified Persons System (NamUs). In: Enc Forensic Sci. Academic Press, Waltham, p 375. doi:10.1016/b978-0-12-382165-2.00057-x

89. Hu CT et al. (2016) Genetic analysis of 15 mtDNA SNP loci in Chinese Yi ethnic group using SNaPshot minisequencing. Gene 576 (1 Pt 1):105-108. doi:10.1016/j.gene.2015.09.071

90. Hu Q et al. (2015) A comparison of four methods for PCR inhibitor removal. Forensic Sci Int Genet 16:94-97. doi:10.1016/j.fsigen.2014.12.001

91. Huang QY et al. (2002) Mutation patterns at dinucleotide microsatellite loci in humans. Am J Hum Genet 70 (3):625-634. doi:Doi 10.1086/338997

92. Hudlow WR et al. (2008) A quadruplex real-time qPCR assay for the simultaneous assessment of total human DNA, human male DNA, DNA degradation and the presence of PCR inhibitors in forensic samples: a diagnostic tool for STR typing. Forensic Sci Int Genet 2 (2):108-125. doi:10.1016/j.fsigen.2007.09.001

93. Huynen L, Lambert DM (2015) A Concentrated Hydrochloric Acid-based Method for Complete Recovery of DNA from Bone. J Forensic Sci 60 (6):1553-1557. doi:10.1111/1556-4029.12846

94. ICMP (2017) QIAGEN and ICMP launch next-generation DNA testing solutions for identifying missing persons ICMP,

95. ICMP (2018) International Commission on Missing Persons (ICMP). ICMP. https://www.icmp.int/about-us/history/.

96. Identification UCfH (2008) Procedure Manual - Forensic Laboratory. Demineralization Extraction of Skeletal Remains, Rev 2 edn., Fort Worth, TX

97. Illumina (2015) ForenSeq DNA Signature Prep Reference Guide. vol 15049528, v01 edn., San Diego, CA

98. Illumina (2015) ForenSeq™ DNA Signature Prep Reference Guide. vol 15049528, v01 edn., San Diego, CA

99. Illumina (2016) ForenSeq Universal Analysis Software Guide 15053876 edn. Illumina,

100. Illumina (2016) ForenSeq™ DNA Signature Prep Kit Data Sheet. www.illumina.com. Illumina, Illumina

101. INTERPOL (2009) Disaster Victim Identification Guide. INTERPOL,

102. Iozzi S et al. (2015) Forensic genetics in NGS era: New frontiers for massively parallel typing. Forensic Sci Int Genet Suppl Ser 5:E418-E419. doi:10.1016/j.fsigss.2015.09.166

103. Ip SC et al. (2015) An evaluation of the performance of five extraction methods: Chelex(R) 100, QIAamp(R) DNA Blood Mini Kit, QIAamp(R) DNA Investigator Kit, QIAsymphony(R) DNA Investigator(R) Kit and DNA IQ. Sci Justice 55 (3):200-208. doi:10.1016/j.scijus.2015.01.005

104. Iyavoo S et al. (2013) Evaluation of five DNA extraction systems for recovery of DNA from bone. Forensic Sci Int Genet Suppl Ser 4 (1):e174-e175. doi:10.1016/j.fsigss.2013.10.090

105. Jim Robinson HT, Jill Mesirov (2018) Integrative Genomics Viewer. Broad Institute, Regents of the University of California. http://software.broadinstitute.org/software/igv/home.

106. Joël J et al. (2015) DNA extraction of forensic adhesive tapes—A comparison of two different methods. Forensic Sci Int Genet Suppl Ser 5:e579-e581. doi:10.1016/j.fsigss.2015.09.229

107. NIJ (2018) NamUs 2.0. UNTHSC, National Institute of Justice, U.S. Department of Justice https://www.namus.gov/.

108. Kallassy H et al. (2018) Comparison of four DNA extraction methods to extract DNA from cigarette butts collected in Lebanon. Sci Justice. doi:10.1016/j.scijus.2018.09.003

109. Karija Vlahovic M, Kubat M (2012) DNA extraction method from bones using Maxwell(R) 16. Leg Med (Tokyo) 14 (5):272-275. doi:10.1016/j.legalmed.2012.04.004

110. Kasu M, Shires K (2015) The validation of forensic DNA extraction systems to utilize soil contaminated biological evidence. Leg Med (Tokyo) 17 (4):232-238. doi:10.1016/j.legalmed.2015.01.004

111. Kemp BM et al. (2006) Repeat silica extraction: a simple technique for the removal of PCR inhibitors from DNA extracts. J Archaeo Sci 33 (12):1680-1689. doi:10.1016/j.jas.2006.02.015

112. Kersbergen P et al. (2009) Developing a set of ancestry-sensitive DNA markers reflecting continental origins of humans. BMC Genet 10 (1):69. doi:10.1186/1471-2156-10-69

113. Kidd KK et al. (2006) Developing a SNP panel for forensic identification of individuals. Forensic Sci Int 164 (1):20-32. doi:10.1016/j.forsciint.2005.11.017

114. Kidd KK et al. (2014) Current sequencing technology makes microhaplotypes a powerful new type of genetic marker for forensics. Forensic Sci Int Genet 12:215-224. doi:10.1016/j.fsigen.2014.06.014

115. Kidd KK et al. (2013) Microhaplotype loci are a powerful new type of forensic marker. Forensic Sci Int Genet Suppl Ser 4 (1):e123-e124. doi:10.1016/j.fsigss.2013.10.063

116. King JL et al. (2018) Increasing the discrimination power of ancestry-and identity-informative SNP loci within the ForenSeq™ DNA Signature Prep Kit. Forensic Sci Int Genet

117. King JL et al. (2014) High-quality and high-throughput massively parallel sequencing of the human mitochondrial genome using the Illumina MiSeq. Forensic Sci Int Genet 12:128-135. doi:10.1016/j.fsigen.2014.06.001

118. King JL et al. (2014) mitoSAVE: mitochondrial sequence analysis of variants in Excel. Forensic Sci Int Genet 12:122-125. doi:10.1016/j.fsigen.2014.05.013

119. King JL et al. (2017) STRait Razor v2s: Advancing sequence-based STR allele reporting and beyond to other marker systems. Forensic Sci Int Genet 29:21-28. doi:10.1016/j.fsigen.2017.03.013

120. Kitayama T et al. (2010) Evaluation of a new experimental kit for the extraction of DNA from bones and teeth using a non-powder method. Leg Med (Tokyo) 12 (2):84-89. doi:10.1016/j.legalmed.2009.12.004

121. Kovras I, Robins S (2016) Death as the border: Managing missing migrants and unidentified bodies at the EU's Mediterranean frontier. Polit Geogr 55:40-49. doi:10.1016/j.polgeo.2016.05.003

122. Kraemer M et al. (2017) Developmental validation of QIAGEN Investigator® 24plex QS Kit and Investigator® 24plex GO! Kit: Two 6-dye multiplex assays for the extended CODIS core loci. Forensic Sci Int Genet 29:9-20

123. Kuś M et al. (2016) Comparison of three different DNA extraction methods from a highly degraded biological material. Journal of Forensic and Legal Medicine 40:47-53. doi:https://doi.org/10.1016/j.jflm.2016.03.002

124. Laboratory F (1994) PCR-Based Typing Protocols. Federal Bureau of Investigation,

125. Lee HY et al. (2010) Simple and highly effective DNA extraction methods from old skeletal remains using silica columns. Forensic Sci Int Genet 4 (5):275-280. doi:10.1016/j.fsigen.2009.10.014

126. Lee JC et al. (2017) SEQ Mapper: A DNA sequence searching tool for massively parallel sequencing data. Forensic Sci Int Genet 26:66-69. doi:10.1016/j.fsigen.2016.10.006

127. Li CX et al. (2016) A panel of 74 AISNPs: Improved ancestry inference within Eastern Asia. Forensic Sci Int Genet 23:101-110. doi:10.1016/j.fsigen.2016.04.002

128. Liu F et al. (2009) Eye color and the prediction of complex phenotypes from genotypes. Curr Biol 19 (5):R192-193. doi:10.1016/j.cub.2009.01.027

129. Loreille OM et al. (2007) High efficiency DNA extraction from bone by total demineralization. Forensic Sci Int Genet 1 (2):191-195. doi:10.1016/j.fsigen.2007.02.006

130. Loreille OM et al. (2010) Integrated DNA and fingerprint analyses in the identification of 60-year-old mummified human remains discovered in an Alaskan glacier. J Forensic Sci 55 (3):813-818. doi:10.1111/j.1556-4029.2010.01356.x

131. Luce C et al. (2009) Validation of the AMPFℓSTR® MiniFilerTM PCR Amplification Kit for use in forensic casework. J Forensic Sci 54 (5):1046-1054

132. Mardis ERJN (2011) A decade's perspective on DNA sequencing technology. 470 (7333):198

133. Marshall PL et al. (2014) A high volume extraction and purification method for recovering DNA from human bone. Forensic Sci Int Genet 12:155-160. doi:10.1016/j.fsigen.2014.06.011

134. Martin P et al. (2014) Improving DNA data exchange: validation studies on a single 6 dye STR kit with 24 loci. Forensic Sci Int Genet 13:68-78. doi:10.1016/j.fsigen.2014.07.002

135. McElhoe JA et al. (2014) Development and assessment of an optimized next-generation DNA sequencing approach for the mtgenome using the Illumina MiSeq. Forensic Sci Int Genet 13:20-29. doi:10.1016/j.fsigen.2014.05.007

136. McKernan KJ et al. (2009) Sequence and structural variation in a human genome uncovered by short-read, massively parallel ligation sequencing using two-base encoding. Genome Res 19 (9):1527-1541. doi:10.1101/gr.091868.109

137. Mehta B et al. (2017) Forensically relevant SNaPshot® assays for human DNA SNP analysis: a review. Int J Leg Med 131 (1):21-37

138. Monroe C et al. (2013) Evaluating the efficacy of various thermo-stable polymerases against co-extracted PCR inhibitors in ancient DNA samples. Forensic Sci Int 228 (1-3):142-153. doi:10.1016/j.forsciint.2013.02.029

139. Montelius K, Lindblom B (2012) DNA analysis in Disaster Victim Identification. Forensic Sci Med Pathol 8 (2):140-147. doi:10.1007/s12024-011-9276-z

140. Moretti TR et al. (2001) Validation of short tandem repeats (STRs) for forensic usage: performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples. J Forensic Sci 46 (3):647-660

141. Morey M et al. (2013) A glimpse into past, present, and future DNA sequencing. Mol Genet Metab 110 (1-2):3-24. doi:10.1016/j.ymgme.2013.04.024

142. Mulero JJ et al. (2006) Development and validation of the AmpFℓSTR® Yfiler™ PCR amplification kit: a male specific, single amplification 17 Y-STR multiplex system. J Forensic Sci 51 (1):64-75

143. Mulero JJ et al. (2008) Development and validation of the AmpFℓSTR® MiniFilerTM PCR Amplification Kit: a miniSTR multiplex for the analysis of degraded and/or PCR inhibited DNA. J Forensic Sci 53 (4):838-852

144. Mundorff A, Davoren JM (2014) Examination of DNA yield rates for different skeletal elements at increasing post mortem intervals. Forensic Sci Int Genet 8 (1):55-63. doi:10.1016/j.fsigen.2013.08.001

145. Mundorff AZ et al. (2009) DNA preservation in skeletal elements from the World Trade Center disaster: recommendations for mass fatality management. J Forensic Sci 54 (4):739-745. doi:10.1111/j.1556-4029.2009.01045.x

146. Nachman MW, Crowell SL (2000) Estimate of the mutation rate per nucleotide in humans. Genetics 156 (1):297-304

147. Noor MSM et al. (2017) The clandestine multiple graves in Malaysia: The first mass identification operation of human skeletal remains. Forensic Sci Int 278:410. e411-410. e419. doi:PMID 410.e1

10.1016/j.forsciint.2017.05.014

148. Novroski NMM et al. (2016) Characterization of genetic sequence variation of 58 STR loci in four major population groups. Forensic Sci Int Genet 25:214-226. doi:10.1016/j.fsigen.2016.09.007

149. Novroski NMM et al. (2018) Potential highly polymorphic short tandem repeat markers for enhanced forensic identity testing. Forensic Sci Int Genet 37:162-171. doi:10.1016/j.fsigen.2018.08.011

150. Oostdik K et al. (2014) Developmental validation of the PowerPlex® Fusion System for analysis of casework and reference samples: A 24-locus multiplex for new database standards. Forensic Sci Int Genet 12:69-76

151. Opel KL et al. (2010) A study of PCR inhibition mechanisms using real time PCR. J Forensic Sci 55 (1):25-33. doi:10.1111/j.1556-4029.2009.01245.x

152. Osier MV et al. (2001) ALFRED: an allele frequency database for diverse populations and DNA polymorphisms—an update. Nucleic Acids Res 29 (1):317-319

153. Ozdemir-Kaynak E, Yesil-Celiktas O (2015) Microwave-assisted digestion combined with silica-based spin column for DNA isolation from human bones. Anal Biochem 486:44-50. doi:10.1016/j.ab.2015.06.034

154. Pakstis AJ et al. (2012) Mini-haplotypes as lineage informative SNPs and ancestry inference SNPs. Eur J Hum Genet 20 (11):1148-1154. doi:10.1038/ejhg.2012.69

155. Parson W et al. (2016) Massively parallel sequencing of forensic STRs: Considerations of the DNA commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements. Forensic Sci Int Genet 22:54-63. doi:10.1016/j.fsigen.2016.01.009

156. Parson W et al. (2004) The EDNAP mitochondrial DNA population database (EMPOP) collaborative exercises: organisation, results and perspectives. Forensic Sci Int 139 (2-3):215-226. doi:10.1016/j.forsciint.2003.11.008

157. Parson W et al. (2013) Reprint of: Evaluation of next generation mtGenome sequencing using the Ion Torrent Personal Genome Machine (PGM). Forensic Sci Int Genet 7 (6):632-639. doi:10.1016/j.fsigen.2013.09.007

158. Parsons TJ, Coble MD (2001) Increasing the forensic discrimination of mitochondrial DNA testing through analysis of the entire mitochondrial DNA genome. Croat Med J 42 (3):304-309

159. Parys-Proszek A et al. (2018) Dual amplification strategy for improved efficiency of forensic DNA analysis using NGM Detect, NGM or Globalfiler kits. Forensic Sci Int Genet 35:46-49. doi:10.1016/j.fsigen.2018.03.018

160. Peschel O, Eisenmenger W (2013) Airplane Crashes and Other Mass Disasters. In: Enc Forensic Sci. Academic Press, Waltham, pp 188-192. doi:10.1016/b978-0-12-382165-2.00182-3

161. Pfeiffer H et al. (1999) Mitochondrial DNA typing from human axillary, pubic and head hair shafts–success rates and sequence comparisons. Int J Leg Med 112 (5):287-290

162. Phillips C (2015) Forensic genetic analysis of bio-geographical ancestry. Forensic Sci Int Genet 18:49-65. doi:10.1016/j.fsigen.2015.05.012

163. Phillips C et al. (2013) Eurasiaplex: a forensic SNP assay for differentiating European and South Asian ancestries. Forensic Sci Int Genet 7 (3):359-366. doi:10.1016/j.fsigen.2013.02.010

164. Phillips C et al. (2007) Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs. Forensic Sci Int Genet 1 (3-4):273-280. doi:10.1016/j.fsigen.2007.06.008

165. Phillips K et al. (2012) A comparison of methods for forensic DNA extraction: Chelex-100(R) and the QIAGEN DNA Investigator Kit (manual and automated). Forensic Sci Int Genet 6 (2):282-285. doi:10.1016/j.fsigen.2011.04.018

166. Pionzio AM, McCord BR (2014) The effect of internal control sequence and length on the response to PCR inhibition in real-time PCR quantitation. Forensic Sci Int Genet 9:55-60. doi:10.1016/j.fsigen.2013.10.010

167. Prinz M et al. (2007) DNA Commission of the International Society for Forensic Genetics (ISFG): recommendations regarding the role of forensic genetics for disaster victim identification (DVI). Forensic Sci Int Genet 1 (1):3-12. doi:10.1016/j.fsigen.2006.10.003

168. QIAGEN (2007-2014) AllPrep ® DNA/RNA Micro Handbook. QIAGEN, Hilden, Germany

169. QIAGEN (2012) QIAamp ® DNA Investigator Handbook. QIAGEN,

170. QIAGEN (2012) QIAamp DNA Investigator Handbook. QIAGEN,

171. QIAGEN (2016) Investigator® 24plex QS Handbook. QIAGEN,

172. Rajeevan H et al. (2012) ALFRED: an allele frequency resource for research and teaching. Nucleic Acids Res 40 (Database issue):D1010-1015. doi:10.1093/nar/gkr924

173. Reich DE et al. (2002) Human genome sequence variation and the influence of gene history, mutation and recombination. Nat Genet 32 (1):135-142. doi:10.1038/ng947

174. Ronaghi M et al. (1996) Real-time DNA sequencing using detection of pyrophosphate release. Anal Biochem 242 (1):84-89. doi:10.1006/abio.1996.0432

175. Rothe J, Nagy M (2016) Comparison of two silica-based extraction methods for DNA isolation from bones. Leg Med (Tokyo) 22:36-41. doi:10.1016/j.legalmed.2016.07.008

176. Salado Puerto M, Tuller H (2017) Large-scale forensic investigations into the missing: Challenges and considerations. Forensic Sci Int 279:219-228. doi:10.1016/j.forsciint.2017.08.025

177. Sancar A et al. (2004) Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. Annu Rev Biochem 73 (1):39-85. doi:10.1146/annurev.biochem.73.011303.073723

178. Sanger F, Coulson AR (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Mol Biol 94 (3):441-448

179. Sanger F et al. (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74 (12):5463-5467

180. Santos C et al. (2016) Pacifiplex: an ancestry-informative SNP panel centred on Australia and the Pacific region. Forensic Sci Int Genet 20:71-80. doi:10.1016/j.fsigen.2015.10.003

181. Scherer M et al. (2015) Improving sample analysis and interpretation using QIAGEN's latest Investigator® STR multiplex PCR assays with a novel quality sensor. Forensic Sci Int Genet Suppl Ser 5:e308-e309

182. Schmedes S et al. (2013) Effective removal of co-purified inhibitors from extracted DNA samples using synchronous coefficient of drag alteration (SCODA) technology. Int J Leg Med 127 (4):749-755. doi:10.1007/s00414-012-0810-7

183. Thermo Fisher Scientific (2015) GlobalFiler™ PCR Amplification Kit. Thermo Fisher Scientific, Thermo Fisher Scientific

184. Thermo Fisher Scientific (2015) Ion AmpliSeq™ Library Preparation for Human Identification Applications. Thermo Fisher Scientific, Thermo Fisher Scientific

185. Thermo Fisher Scientific (2017) Precision ID Panels with the Ion S5 System: Application Guide. MAN0015831 Rev. B.0 edn. Thermo Fisher Scientific,

186. Thermo Fisher Scientific (2017) Quantifiler™ HP and Trio DNA Quantification Kits: User Guide. F edn. Thermo Fisher Scientific,

187. Thermo Fisher Scientific (2018) AutoMate *Express*™ Instrument User Guide. F edn. Thermo Fisher Scientific,

188. Seo SB et al. (2013) Single nucleotide polymorphism typing with massively parallel sequencing for human identification. Int J Leg Med 127 (6):1079-1086. doi:10.1007/s00414-013-0879-7

189. Sobrino B et al. (2005) SNPs in forensic genetics: a review on SNP typing methodologies. Forensic Sci Int 154 (2-3):181-194. doi:10.1016/j.forsciint.2004.10.020

190. Strobl C et al. (2018) Evaluation of the precision ID whole MtDNA genome panel for forensic analyses. Forensic Sci Int Genet 35:21-25. doi:10.1016/j.fsigen.2018.03.013

191. Sturk-Andreaggi K et al. (2017) AQME: A forensic mitochondrial DNA analysis tool for next-generation sequencing data. Forensic Sci Int Genet 31:189-197. doi:10.1016/j.fsigen.2017.09.010

192. Life Technologies.

193. Life Technologies (2012) PrepFiler ® and PrepFiler ® BTA Forensic DNA Extraction Kits User Guide. vol 4463348, Rev. C edn. Life Technologies, Carlsbad, CA

194. Templeton JEL et al. (2013) DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification. Investig Genet 4 (1):26

195. Thompson RE et al. (2014) An Investigation of PCR Inhibition Using Plexor®-Based Quantitative PCR and Short Tandem Repeat Amplification. J Forensic Sci 59 (6):1517-1529

196. Thorvaldsdottir H et al. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 14 (2):178-192. doi:10.1093/bib/bbs017

197. Tishkoff SA, Verrelli BC (2003) Patterns of human genetic diversity: implications for human evolutionary history and disease. Annu Rev Genomics Hum Genet 4 (1):293-340. doi:10.1146/annurev.genom.4.070802.110226

198. Tucker VC et al. (2011) Developmental validation of the PowerPlex((R)) ESI 16 and PowerPlex((R)) ESI 17 Systems: STR multiplexes for the new European standard. Forensic Sci Int Genet 5 (5):436-448. doi:10.1016/j.fsigen.2010.09.004

199. Vallee J (2018) Department of Defense DNA Operations. Armed Forces Medical Examiner System, Dover AFB, DE

200. Van Neste C et al. (2012) Forensic STR analysis using massive parallel sequencing. Forensic Sci Int Genet 6 (6):810-818

201. Van Neste C et al. (2014) My-Forensic-Loci-queries (MyFLq) framework for analysis of forensic STR data generated by massive parallel sequencing. Forensic Sci Int Genet 9:1-8. doi:10.1016/j.fsigen.2013.10.012

202. van Oven M, Kayser M (2009) Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum Mutat 30 (2):E386-394. doi:10.1002/humu.20921

203. Vernarecci S et al. (2015) Quantifiler® Trio Kit and forensic samples management: a matter of degradation. Forensic Sci Int Genet 16:77-85

204. Verogen (2018) ForenSeq DNA Signature Prep Reference Guide. VD2018005 Rev. A edn. Verogen,

205. Walsh S et al. (2011) Developmental validation of the IrisPlex system: determination of blue and brown iris colour for forensic intelligence. Forensic Sci Int Genet 5 (5):464-471. doi:10.1016/j.fsigen.2010.09.008

206. Walsh S et al. (2011) IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. Forensic Sci Int Genet 5 (3):170-180. doi:10.1016/j.fsigen.2010.02.004

207. Walsh S et al. (2013) The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. Forensic Sci Int Genet 7 (1):98-115. doi:10.1016/j.fsigen.2012.07.005

208. Walsh S et al. (2012) DNA-based eye colour prediction across Europe with the IrisPlex system. Forensic Sci Int Genet 6 (3):330-340. doi:10.1016/j.fsigen.2011.07.009

209. Wang DY et al. (2012) Developmental validation of the AmpFℓSTR® Identifiler® Plus PCR Amplification Kit: an established multiplex assay with improved performance. J Forensic Sci 57 (2):453-465

210. Wang DY et al. (2015) Developmental validation of the GlobalFiler® Express PCR Amplification Kit: a 6-dye multiplex assay for the direct amplification of reference samples. Forensic Sci Int Genet 19:148-155

211. Wang H et al. (2015) NGS technology makes microhaplotype a potential forensic marker. Forensic Sci Int Genet Suppl Ser 5:E233-E234. doi:10.1016/j.fsigss.2015.09.093

212. Wang Z et al. (2017) Massively parallel sequencing of 32 forensic markers using the Precision ID GlobalFiler NGS STR Panel and the Ion PGM System. Forensic Sci Int Genet 31:126-134. doi:10.1016/j.fsigen.2017.09.004

213. Warshauer DH et al. (2015) STRait razor v2. 0: the improved STR allele identification tool–razor. Forensic Sci Int Genet Suppl Ser 14:182-186

214. Warshauer DH et al. (2013) STRait Razor: a length-based forensic STR allele-calling tool for use with second generation sequencing data. Forensic Sci Int Genet 7 (4):409-417. doi:10.1016/j.fsigen.2013.04.005

215. Watherston J et al. (2018) Current and emerging tools for the recovery of genetic information from post mortem samples: New directions for disaster victim identification. Forensic Sci Int Genet 37:270-282. doi:10.1016/j.fsigen.2018.08.016

216. Weiler NE et al. (2016) Development of a control region-based mtDNA SNaPshot selection tool, integrated into a mini amplicon sequencing method. Sci Justice 56 (2):96-103. doi:10.1016/j.scijus.2015.11.003

217. Weiner S, Price PA (1986) Disaggregation of Bone into Crystals. Calcified Tissue International 39 (6):365-375. doi:Doi 10.1007/Bf02555173

218. Wendt FR et al. (2016) Massively parallel sequencing of 68 insertion/deletion markers identifies novel microhaplotypes for utility in human identity testing. Forensic Sci Int Genet 25:198-209. doi:10.1016/j.fsigen.2016.09.005

219. Woerner AE et al. (2018) Evaluation of the precision ID mtDNA whole genome panel on two massively parallel sequencing systems. Forensic Sci Int Genet 36:213-224. doi:10.1016/j.fsigen.2018.07.015

220. Woerner AE et al. (2017) Fast STR allele identification with STRait Razor 3.0. Forensic Sci Int Genet 30:18-23. doi:10.1016/j.fsigen.2017.05.008

221. Yang DY et al. (1998) Improved DNA extraction from ancient bones using silica-based spin columns. American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists 105 (4):539-543

222. Yang Y et al. (2014) Application of next-generation sequencing technology in forensic science. Genomics Proteomics Bioinformatics 12 (5):190-197. doi:10.1016/j.gpb.2014.09.001

223. Zeng X et al. (2018) Assessment of impact of DNA extraction methods on analysis of human remain samples on massively parallel sequencing success. Int J Leg Med. doi:10.1007/s00414-018-1955-9

224. Zeng X et al. (2015) An evaluation of the PowerSeq Auto System: A multiplex short tandem repeat marker kit compatible with massively parallel sequencing. Forensic Sci Int Genet 19:172-179. doi:10.1016/j.fsigen.2015.07.015

225. Zeng X et al. (2015) High sensitivity multiplex short tandem repeat loci analyses with massively parallel sequencing. Forensic Sci Int Genet 16:38-47. doi:10.1016/j.fsigen.2014.11.022

226. Zgonjanin D et al. (2017) DNA analysis from human skeletal remains in forensic casework. Forensic Sci Int Genet Suppl Ser 6:E342-E345. doi:10.1016/j.fsigss.2017.09.117

227. Zhang S et al. (2017) Developmental validation of a custom panel including 273 SNPs for forensic application using Ion Torrent PGM. Forensic Sci Int Genet 27:50-57. doi:10.1016/j.fsigen.2016.12.003

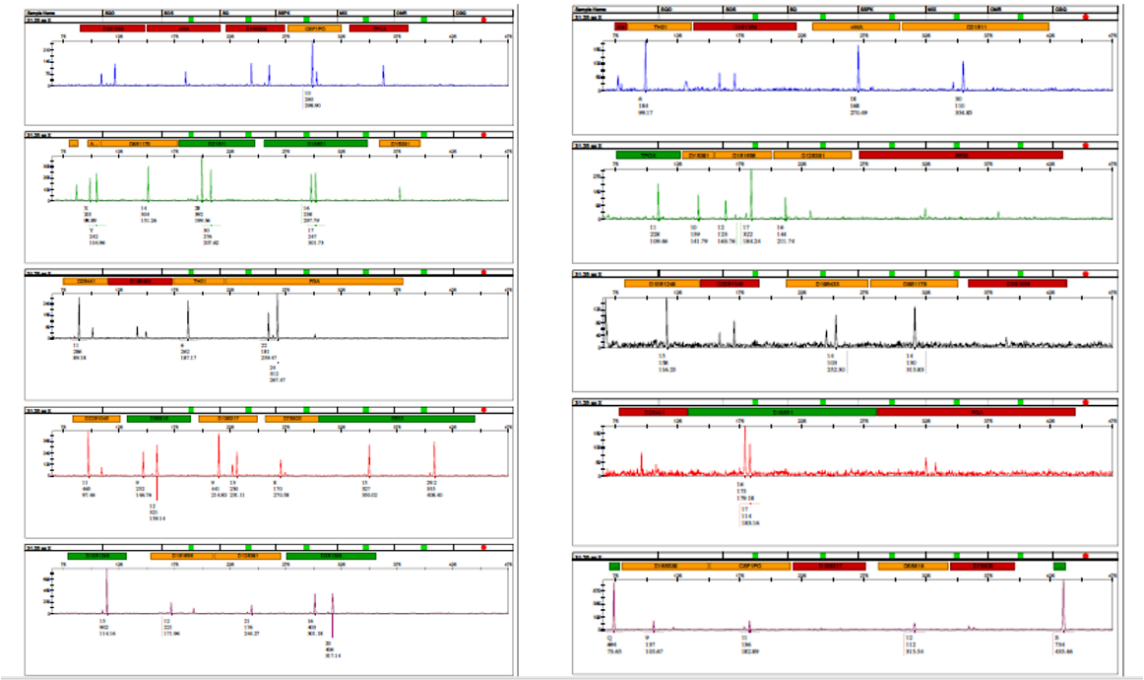
228. Zietkiewicz E et al. (2012) Current genetic methodologies in the identification of disaster victims and in forensic analysis. J Appl Genet 53 (1):41-60. doi:10.1007/s13353-011-0068-7

229. Zubakov D et al. (2015) Towards simultaneous individual and tissue identification: A proof-of-principle study on parallel sequencing of STRs, amelogenin, and mRNAs with the Ion Torrent PGM. Forensic Sci Int Genet 17:122-128. doi:10.1016/j.fsigen.2015.04.002

APPENDIX

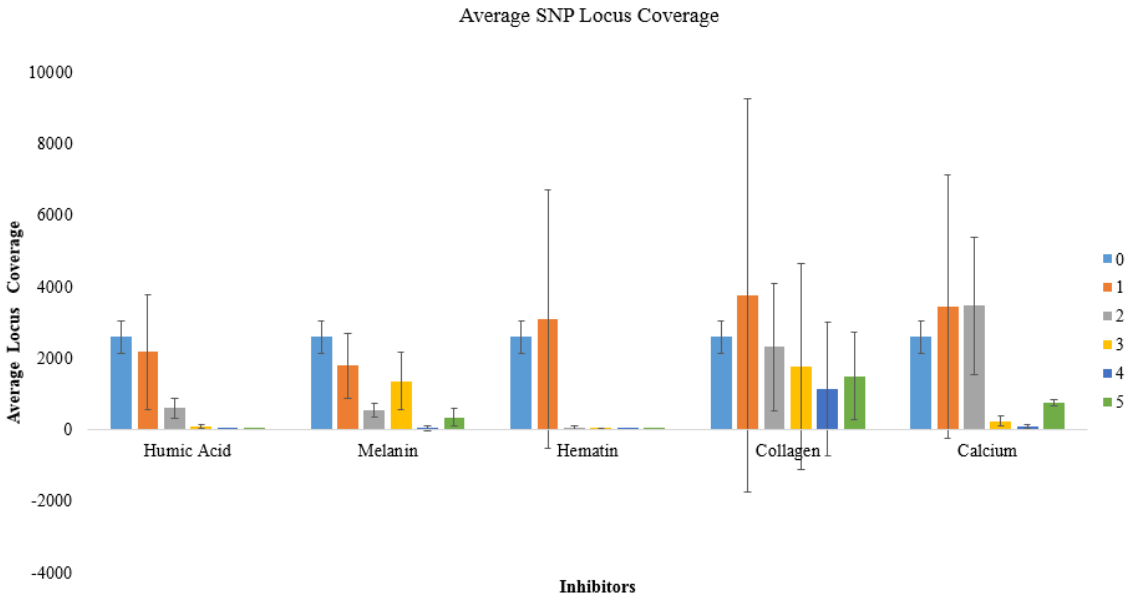
Appendix 2

Appendix 2.1



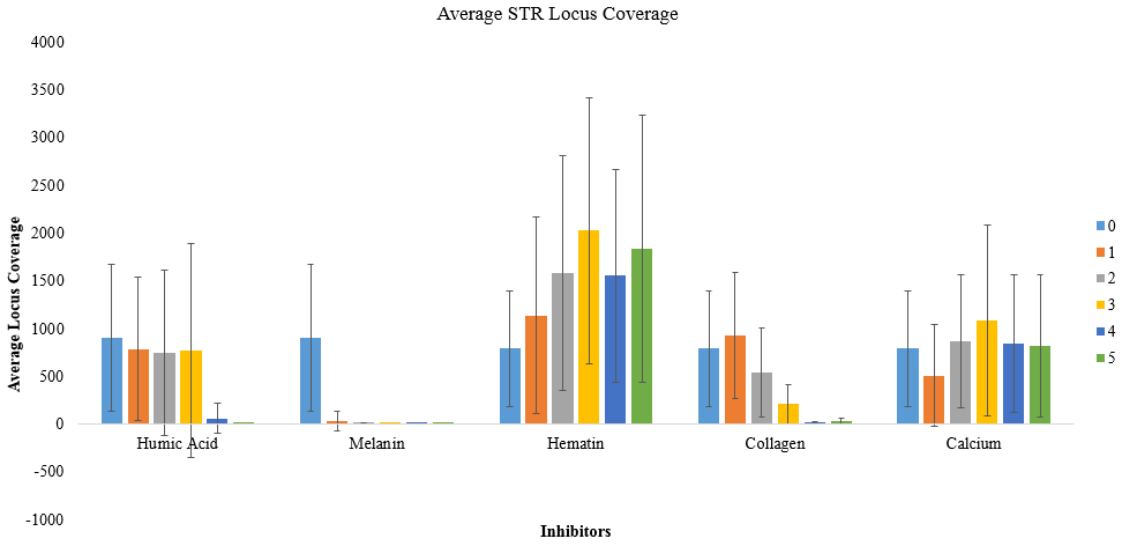
Appendix 3

Appendix 3.1



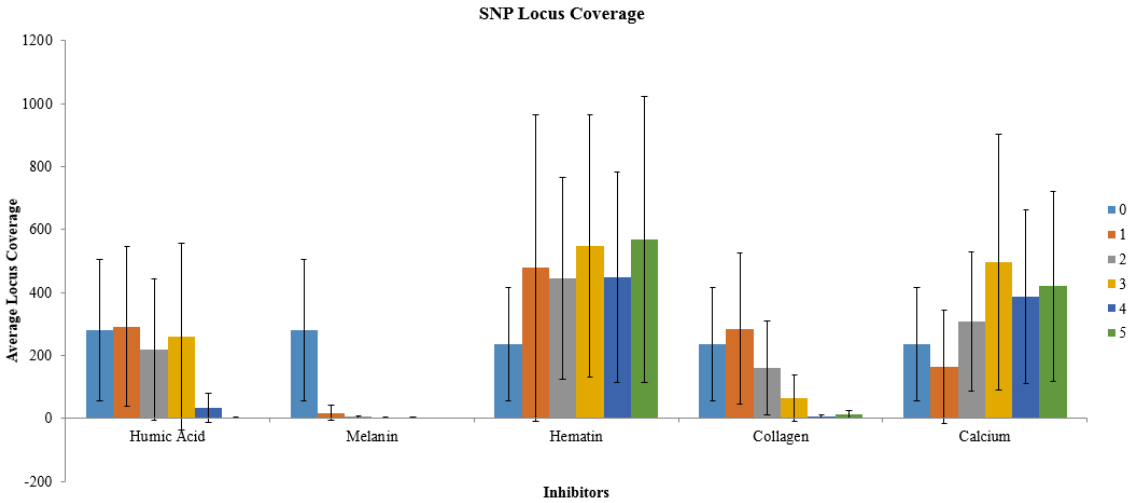
The average SNP loci coverage (Ion Ampliseq™ Library Kit and ID panel) of a 1 ng DNA sample spiked with five concentrations of five PCR inhibitors. Concentration 0 means no inhibitor added. The inhibitor concentrations are listed in Table 3.1. Data presented as average ± standard deviation (N = 3).

Appendix 3.2



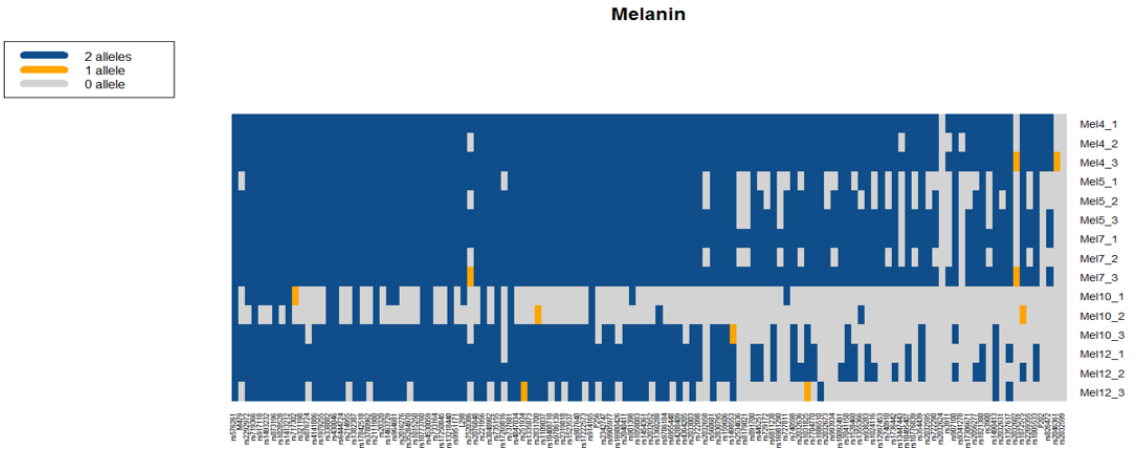
The average STR loci coverage (ForenSeq™ DNA Signature Prep Kit) of a 1 ng DNA sample spiked with five concentrations of five PCR inhibitors. Concentration 0 means no inhibitor added. The inhibitor concentrations were listed in Table 3.1. Data presented as average ± standard deviation (N = 3).

Appendix 3.3



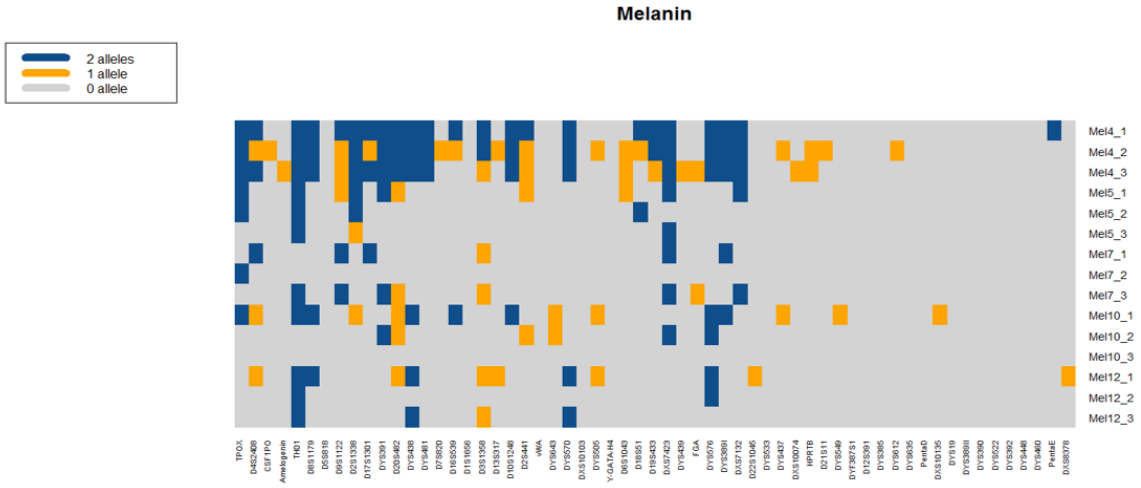
The average SNP loci coverage (ForenSeq™ DNA Signature Prep Kit) of a 1 ng DNA sample spiked with five concentrations of five PCR inhibitors. Concentration 0 means no inhibitor added. The inhibitor concentrations were listed in Table 3.1. Data presented as average ± standard deviation (N = 3).

Appendix 3.4



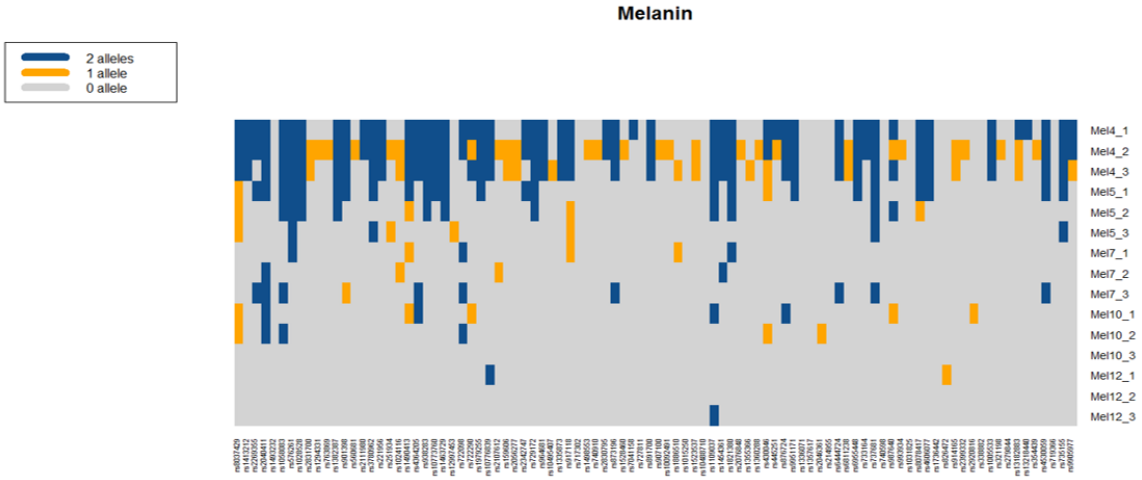
The heatmap of alleles typed at each SNP locus with five concentrations of melanin (Ion Ampliseq™ Library Kit and ID panel).

Appendix 3.5



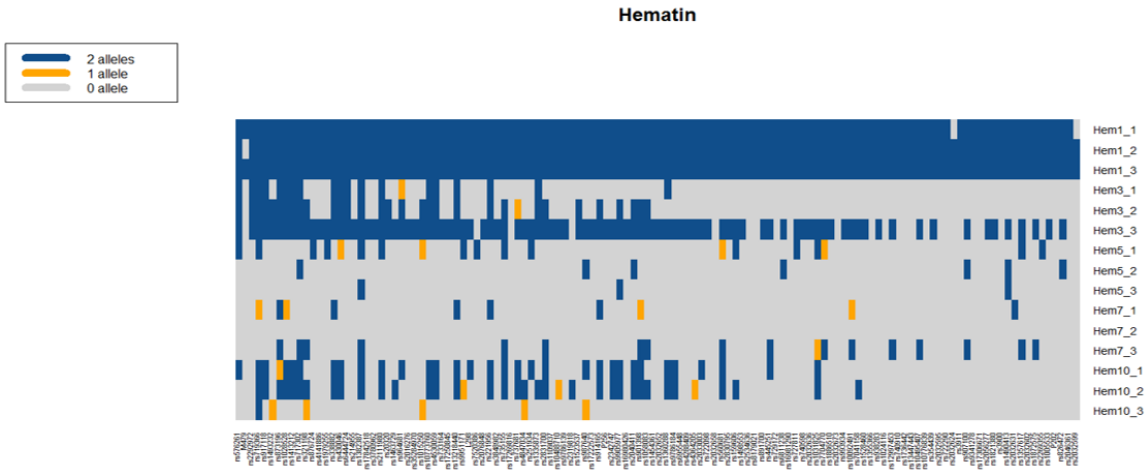
The heatmap of alleles typed at each STR locus with five concentrations of melanin (ForenSeq™ DNA Signature Prep Kit).

Appendix 3.6



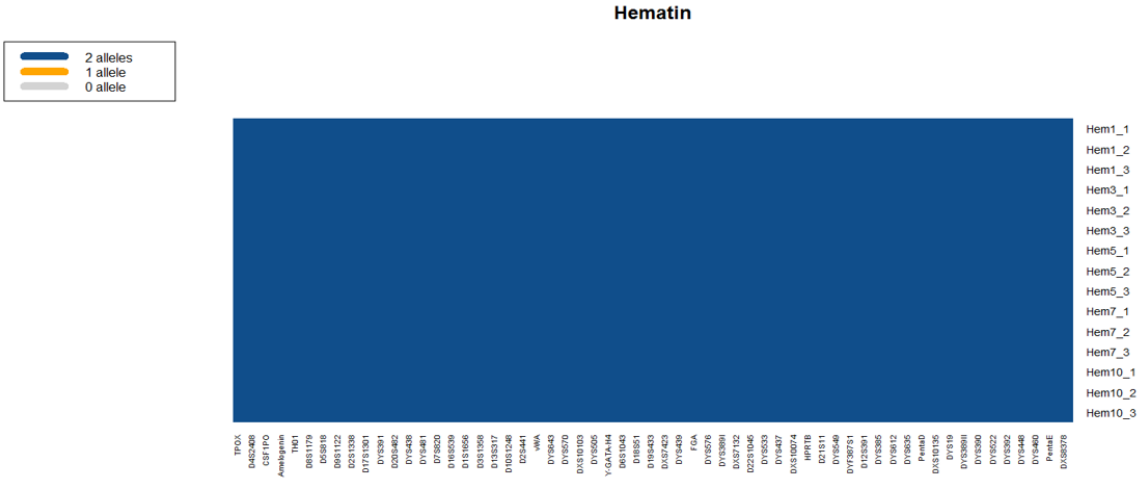
The heatmap of alleles typed at each SNP locus with five concentrations of melanin (ForenSeq™ DNA Signature Prep Kit).

Appendix 3.7



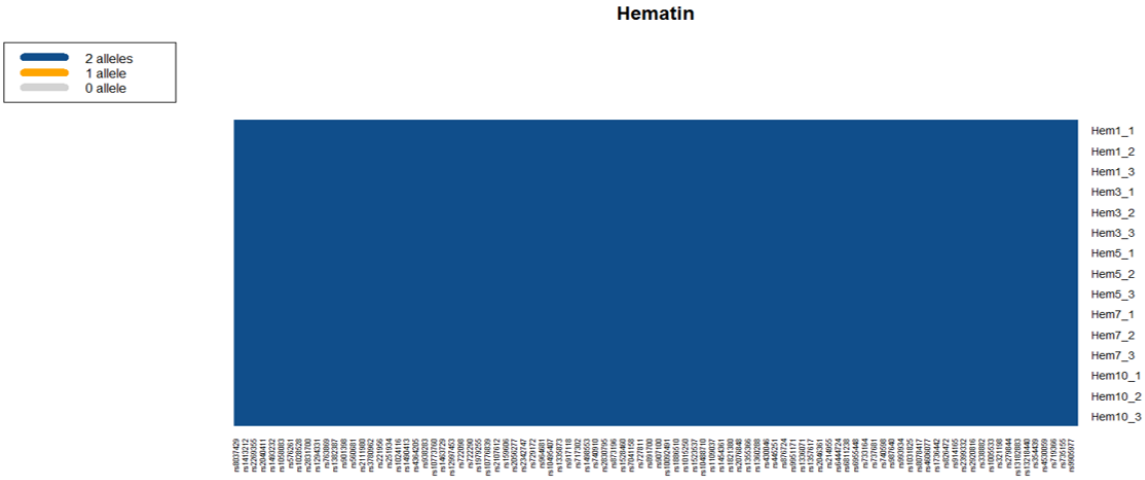
The heatmap of alleles typed at each SNP locus with five concentrations of hematin (Ion Ampliseq™ Library Kit and ID panel).

Appendix 3.8



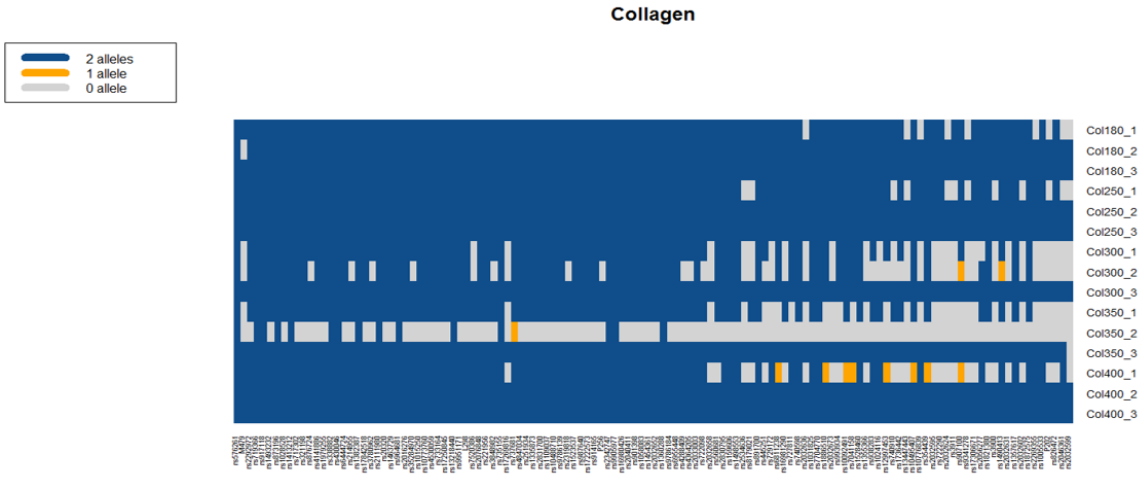
The heatmap of alleles typed at each STR locus with five concentrations of hematin (ForenSeq™ DNA Signature Prep Kit).

Appendix 3.9



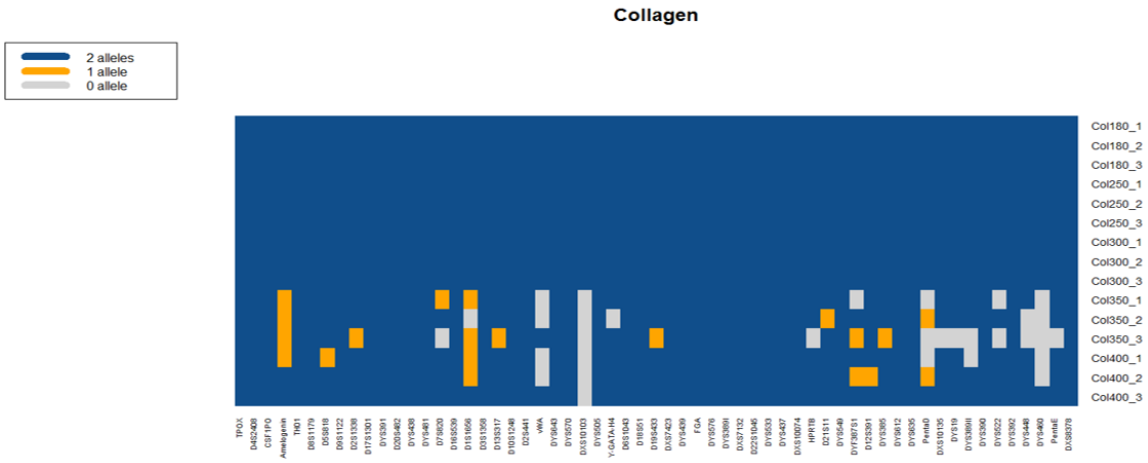
The heatmap of alleles typed at each SNP locus with five concentrations of hematin (ForenSeq™ DNA Signature Prep Kit).

Appendix 3.10



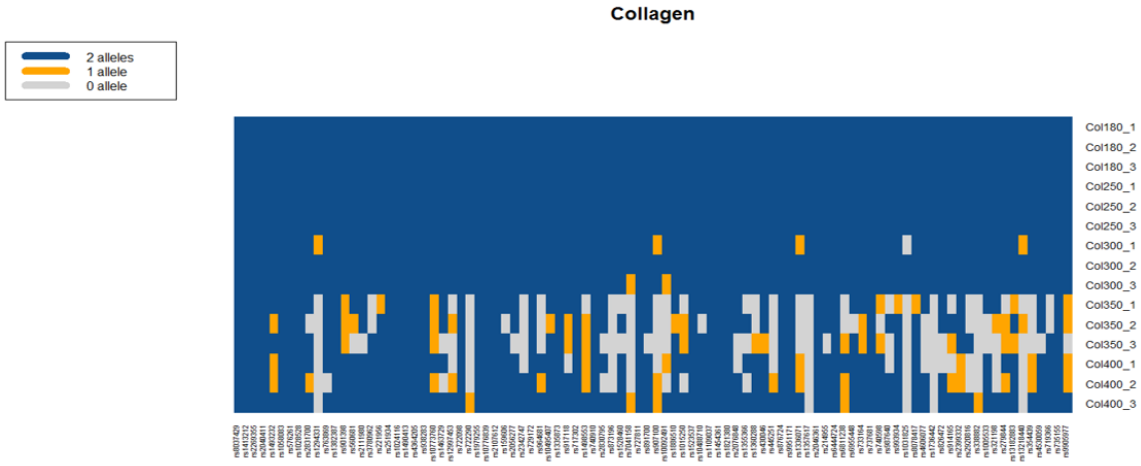
The heatmap of alleles typed at each SNP locus with five concentrations of collagen (Ion Ampliseq^{TM} Library Kit and ID panel).

Appendix 3.11



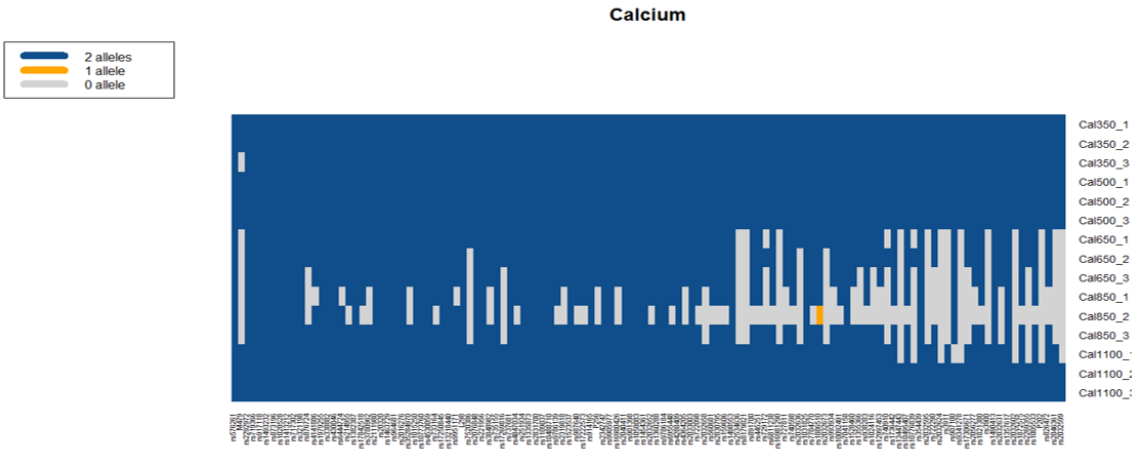
The heatmap of alleles typed at each STR locus with five concentrations of collagen (ForenSeq^{TM} DNA Signature Prep Kit).

Appendix 3.12



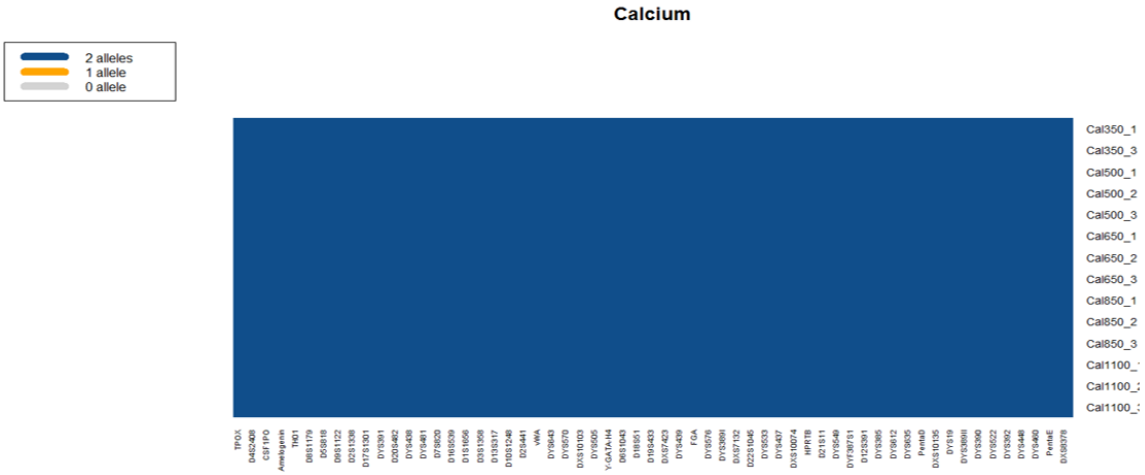
The heatmap of alleles typed at each SNP locus with five concentrations of collagen (ForenSeq^{TM} DNA Signature Prep Kit).

Appendix 3.13



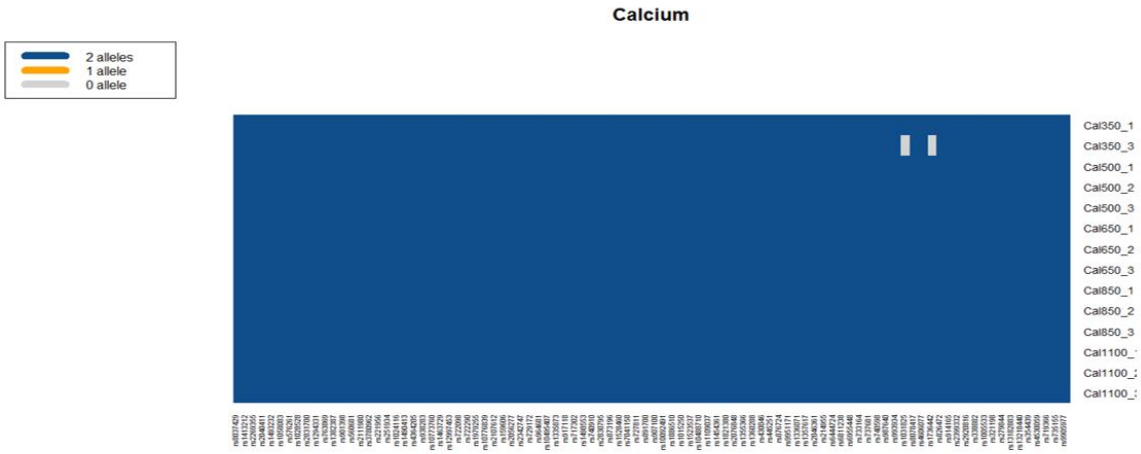
The heatmap of alleles typed at each SNP locus with five concentrations of calcium (Ion Ampliseq^{TM} Library Kit and ID panel).

Appendix 3.14



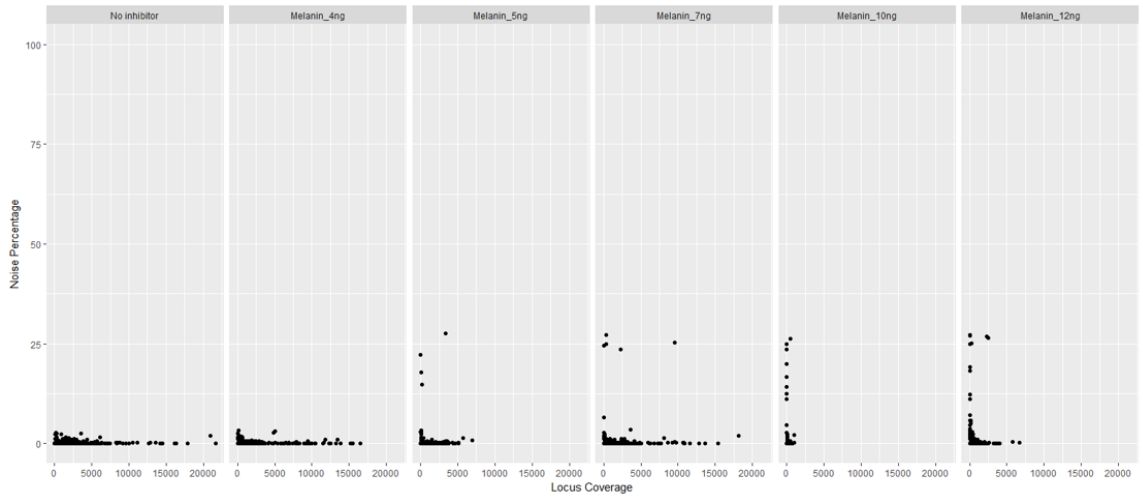
The heatmap of alleles typed at each STR locus with five concentrations of calcium (ForenSeq™ DNA Signature Prep Kit).

Appendix 3.15



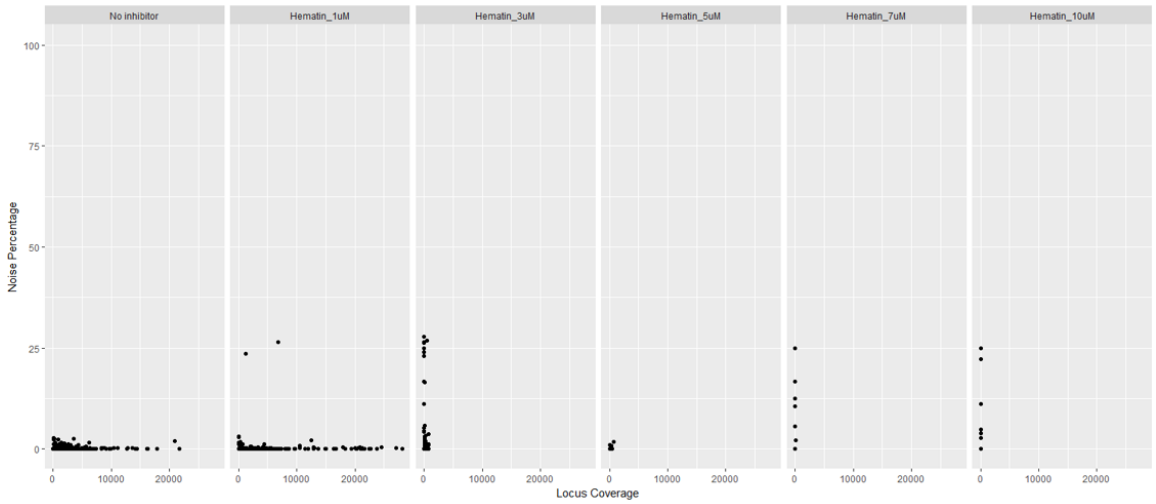
The heatmap of alleles typed at each SNP locus with five concentrations of calcium (ForenSeq™ DNA Signature Prep Kit).

Appendix 3.16



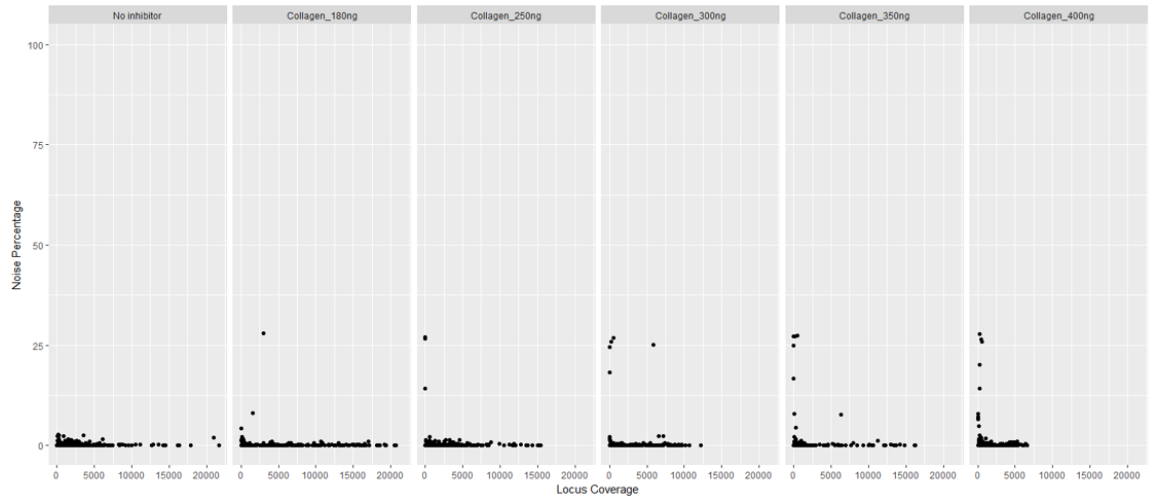
The noise percentages of SNPs (Ion Ampliseq™ Library Kit and ID panel) of reference samples and DNA samples spiked with five concentrations of melanin. X axis is locus coverage, Y axis is noise percentage.

Appendix 3.17



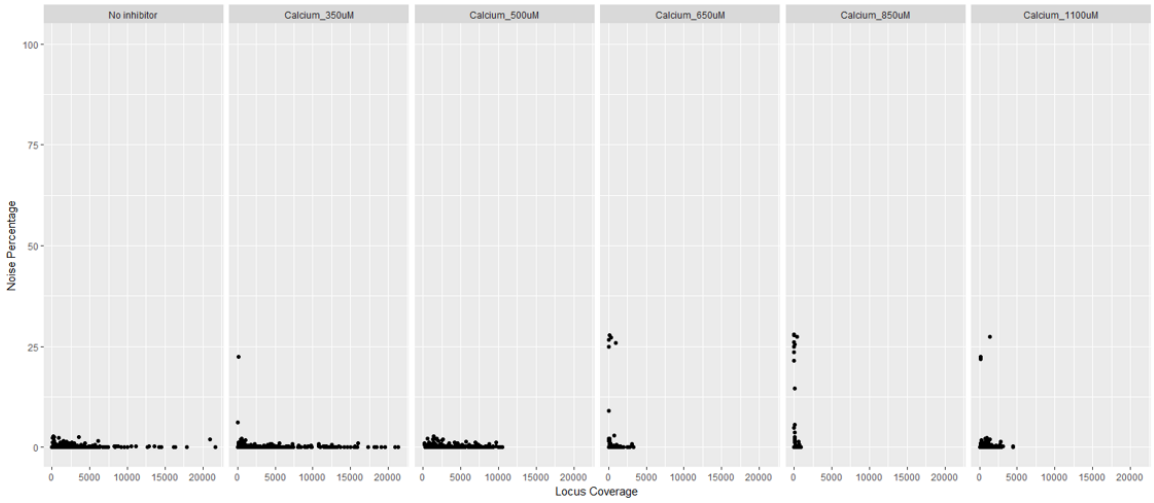
The noise percentages of SNPs (Ion Ampliseq™ Library Kit and ID panel) of reference samples and DNA samples spiked with five concentrations of hematin. X axis is locus coverage, Y axis is noise percentage.

Appendix 3.18



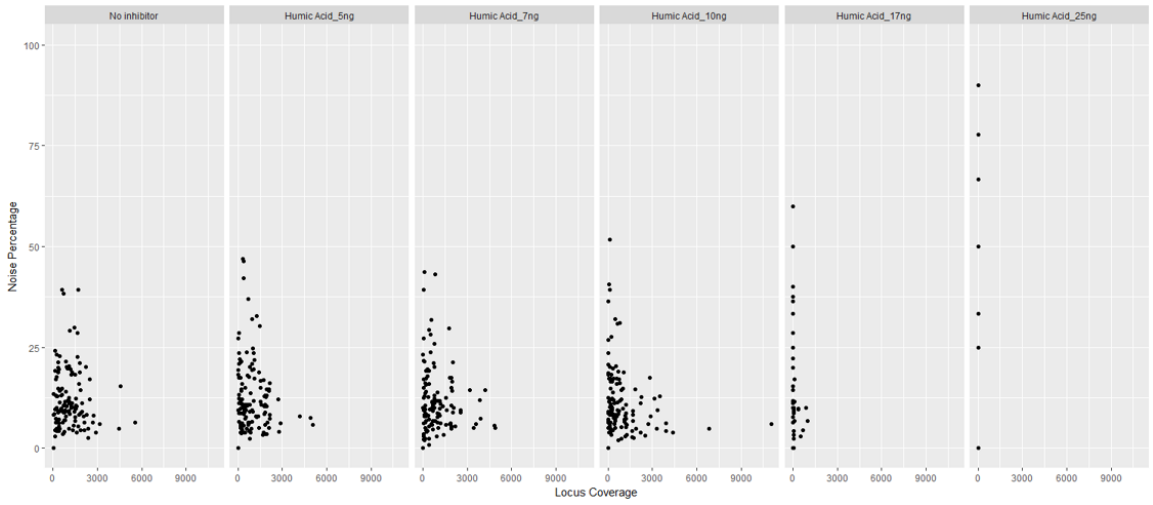
The noise percentages of SNPs (Ion Ampliseq^{TM} Library Kit and ID panel) of reference samples and DNA samples spiked with five concentrations of collagen. X axis is locus coverage, Y axis is noise percentage.

Appendix 3.19



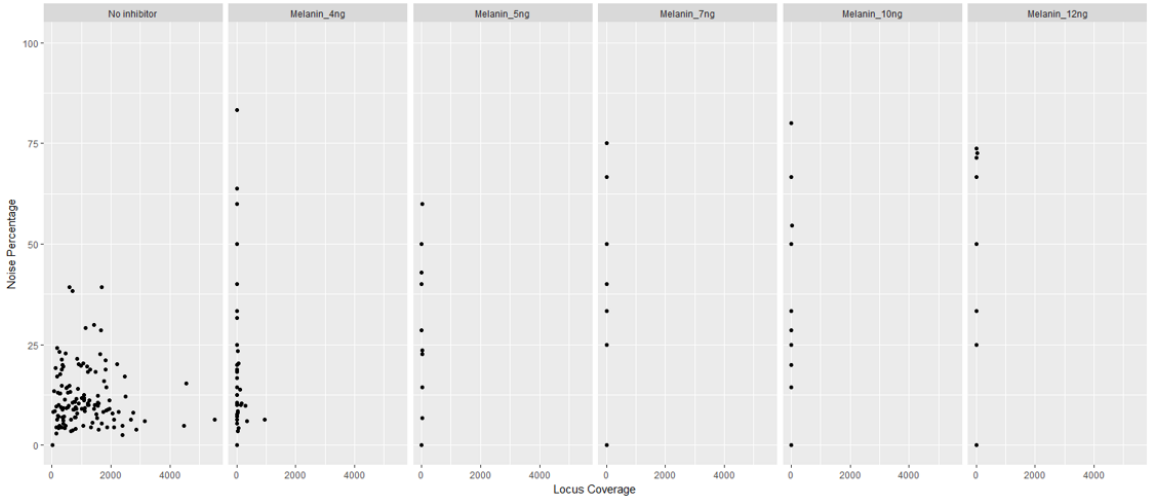
The noise percentages of SNPs (Ion Ampliseq^{TM} Library Kit and ID panel) of reference samples and DNA samples spiked with five concentrations of calcium. X axis is locus coverage, Y axis is noise percentage.

Appendix 3.20



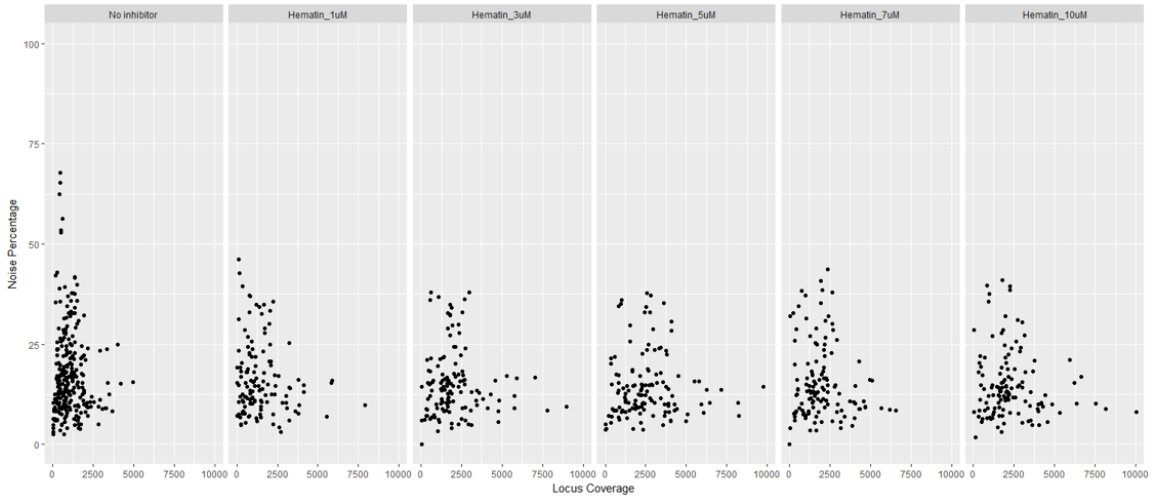
The noise percentages of STRs (ForenSeq™ DNA Signature Prep Kit) of reference samples and DNA samples spiked with five concentrations of humic acid. X axis is locus coverage, Y axis is noise percentage.

Appendix 3.21



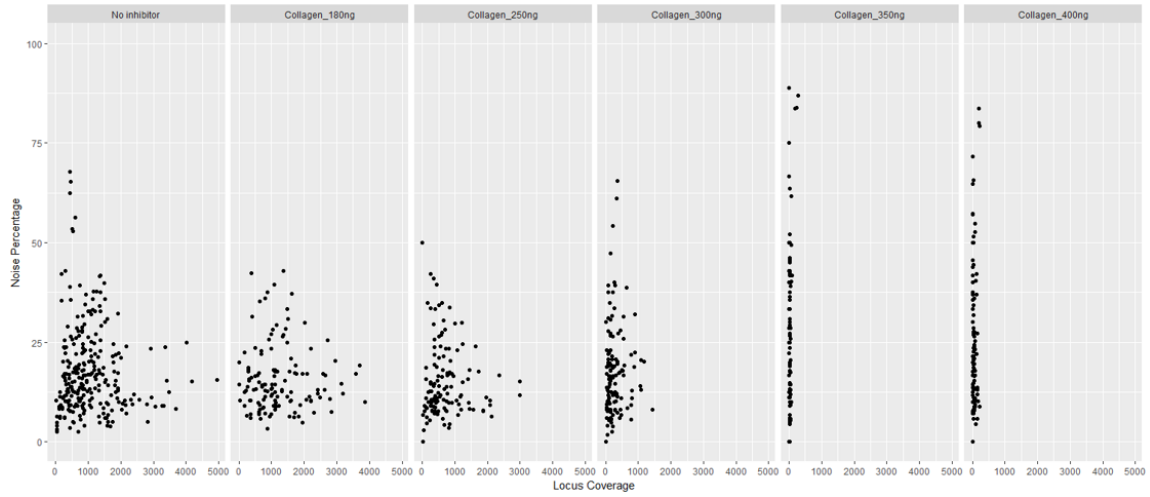
The noise percentages of STRs (ForenSeq™ DNA Signature Prep Kit) of reference samples and DNA samples spiked with five concentrations of melanin. X axis is locus coverage, Y axis is noise percentage.

Appendix 3.22



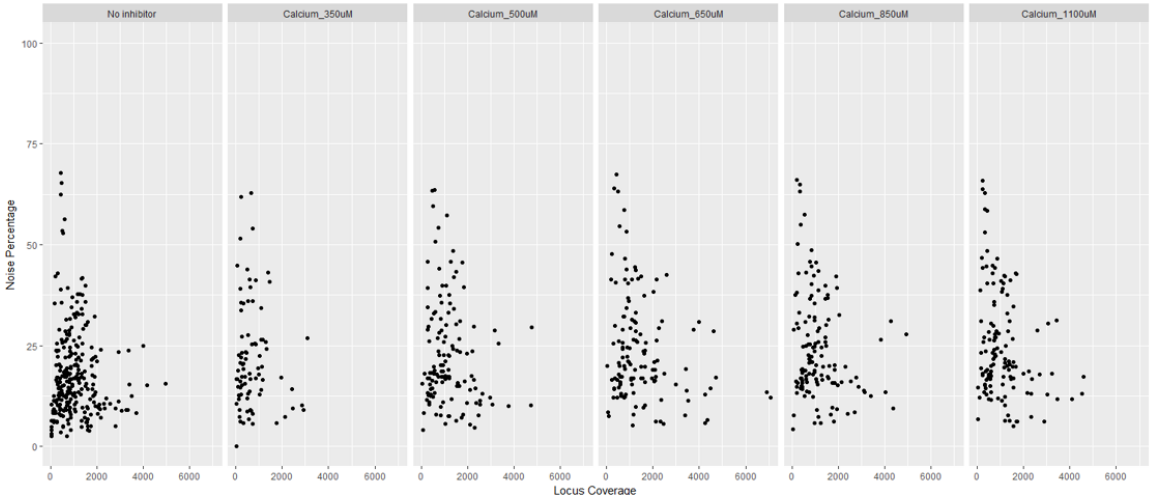
The noise percentages of STRs (ForenSeq™ DNA Signature Prep Kit) of reference samples and DNA samples spiked with five concentrations of hematin. X axis is locus coverage, Y axis is noise percentage.

Appendix 3.23



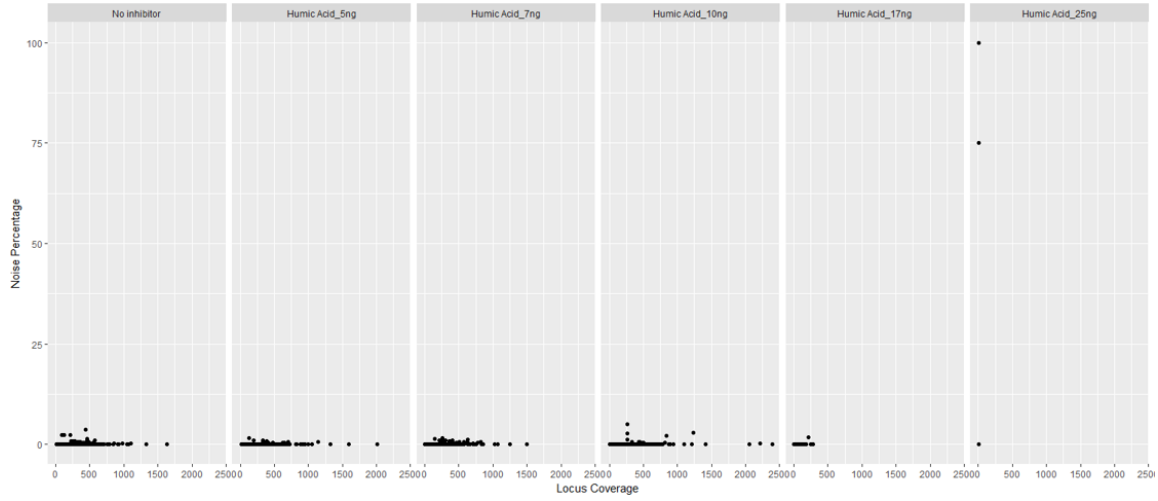
The noise percentages of STRs (ForenSeq™ DNA Signature Prep Kit) of reference samples and DNA samples spiked with five concentrations of collagen. X axis is locus coverage, Y axis is noise percentage.

Appendix 3.24



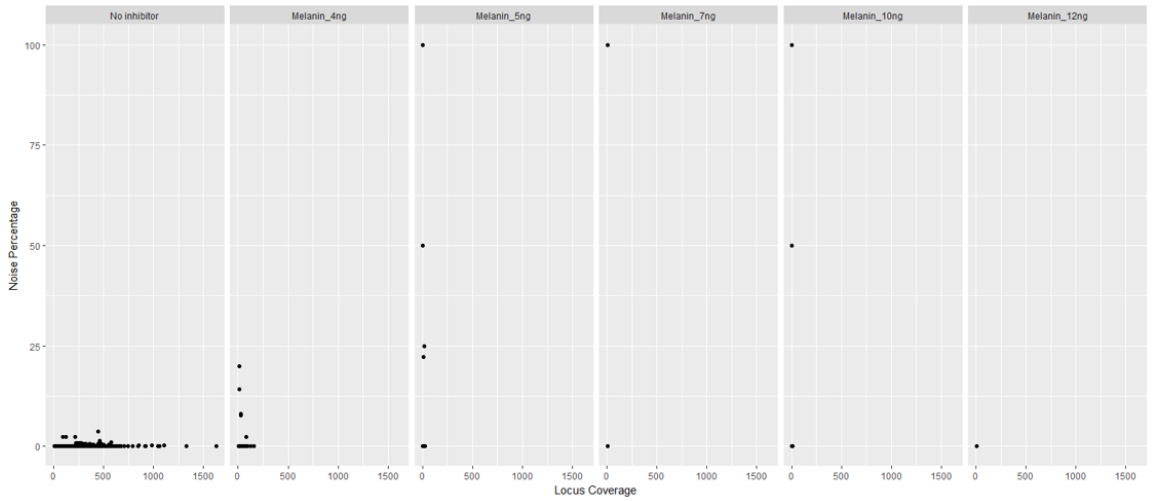
The noise percentages of STRs (ForenSeq™ DNA Signature Prep Kit) of reference samples and DNA samples spiked with five concentrations of calcium. X axis is locus coverage, Y axis is noise percentage.

Appendix 3.25



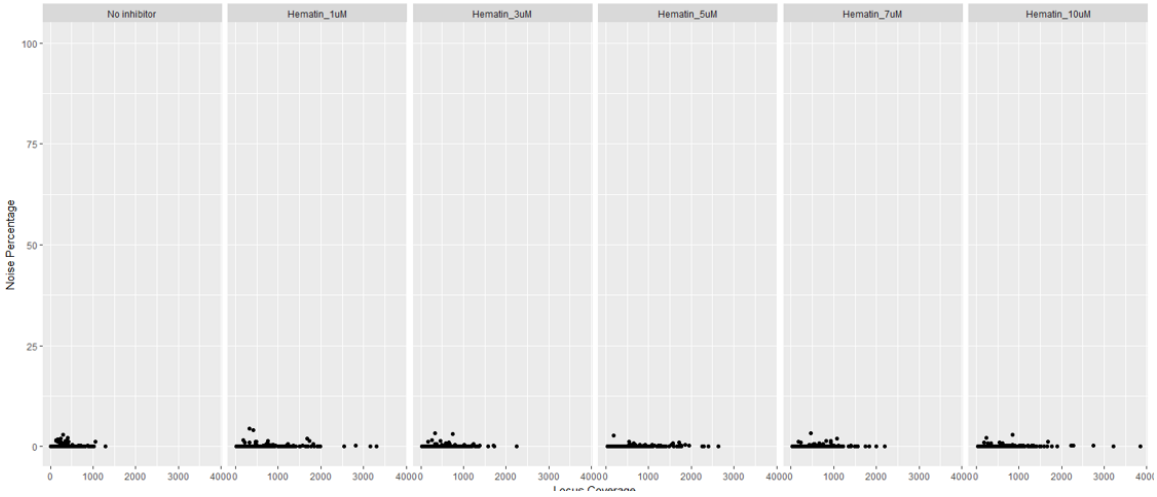
The noise percentages of SNPs (ForenSeq™ DNA Signature Prep Kit) of reference samples and DNA samples spiked with five concentrations of humic acid. X axis is locus coverage, Y axis is noise percentage.

Appendix 3.26



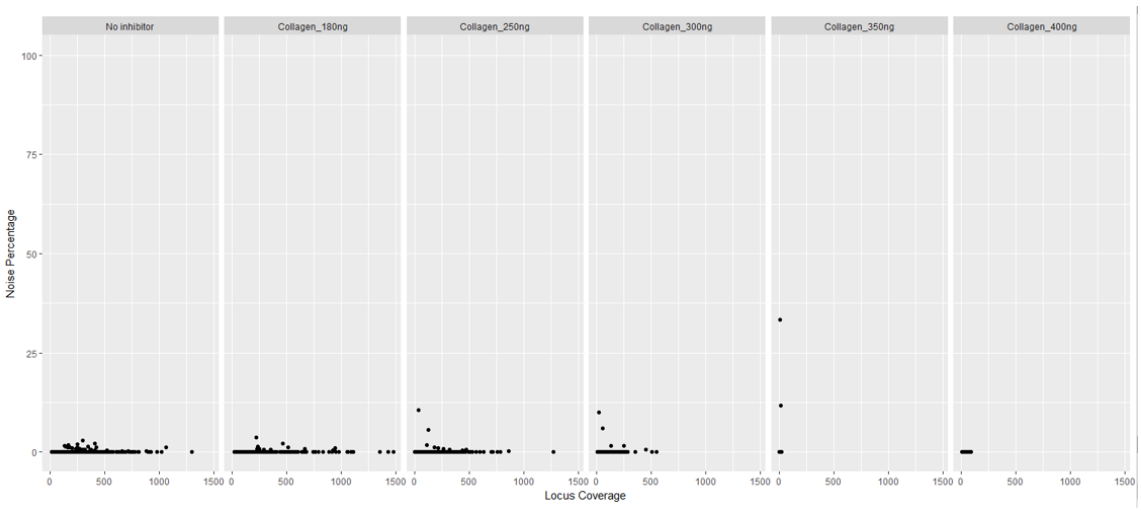
The noise percentages of SNPs (ForenSeq™ DNA Signature Prep Kit) of reference samples and DNA samples spiked with five concentrations of melanin. X axis is locus coverage, Y axis is noise percentage.

Appendix 3.27



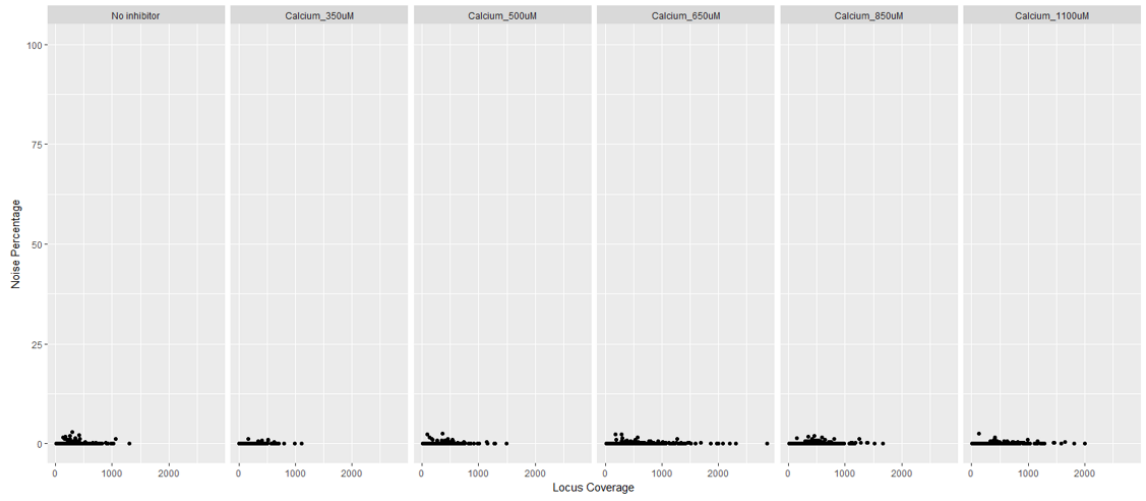
The noise percentages of SNPs (ForenSeq™ DNA Signature Prep Kit) of reference samples and DNA samples spiked with five concentrations of hematin. X axis is locus coverage, Y axis is noise percentage.

Appendix 3.28



The noise percentages of SNPs (ForenSeq™ DNA Signature Prep Kit) of reference samples and DNA samples spiked with five concentrations of collagen. X axis is locus coverage, Y axis is noise percentage.

Appendix 3.29



The noise percentages of SNPs (ForenSeq™ DNA Signature Prep Kit) of reference samples and DNA samples spiked with five concentrations of calcium. X axis is locus coverage, Y axis is noise percentage.

VITA

EDUCATION

Sam Houston State University 8/2014 – Present
Forensic Science Doctoral Candidate
GPA: 3.9
Expected Graduation: December 2018

Truman State University 8/2010 – 5/2014
Bachelor of Arts in Biology
GPA: 3.25
Graduated: May 2014

PUBLICATIONS

Kyleen Elwick, Magdalena Bus, Jonathan King, Joseph Chang, Sheree Hughes-Stamm, Bruce Budowle (2018)
Utility of the Ion S5™ and MiSeq FGx™ Sequencing Platforms to Characterize Challenging Human Remains
Forensic Science International: Genetics (Under Review)

Xiangpei Zeng*, Kyleen Elwick*, Carrie Mayes, Jonathan King, Maiko Takahashi, David Gangitano, Bruce Budowle, Sheree Hughes-Stamm (2018)
Assessment of Impact of DNA Extraction Methods on Analysis of Human Remain Samples on Massively Parallel Sequencing Success
International Journal of Legal Medicine

Kyleen Elwick, Carrie Mayes, Sheree Hughes-Stamm. (2018)
Comparative Sensitivity and Inhibitor Tolerance of GlobalFiler® PCR Amplification and Investigator® 24plex QS Kits for Challenging Samples
Legal Medicine

Kyleen Elwick*, Xiangpei Zeng*, Jonathan King, Bruce Budowle, Sheree Hughes-Stamm. (2017)
Comparative Tolerance of Two Massively Parallel Sequencing Systems to Common PCR Inhibitors
International Journal of Legal Medicine

Amy Sorensen, Rachel Houston, Kyleen Elwick, Sheree Hughes-Stamm. (2017)
Evaluation of Four Commercial Quantitative Real-Time PCR (qPCR) kits with Inhibited and Degraded Samples
International Journal of Legal Medicine

PRESENTATIONS

International Symposium on Human Identification, Phoenix, Arizona, September 2018

Using the Ion S5™ and MiSeq FGx™ Systems to Identify Challenging Human Remains
Kyleen Elwick, Magdalena Bus, Jonathan King, Joseph Chang, Bruce Budowle, Sheree Hughes-Stamm

Gordon Research Seminar and Gordon Research Conference, Newry, Maine, June 2018
Assessment of Five Common Extraction Methods for the Analysis of Human Remains using the Ion S5™ and MiSeq FGx™ Systems
Kyleen Elwick, Carrie Mayes, Xiangpei Zeng, Jonathan King, Bruce Budowle, Sheree Hughes-Stamm

Human Identification Solutions, Rome, Italy, May 2018
Evaluation of Five Common Extraction Methods for Analysis of Human Remain Samples on Massively Parallel Sequencing Success
Kyleen Elwick, Carrie Mayes, Xiangpei Zeng, Jonathan King, Bruce Budowle, Sheree Hughes-Stamm

PITTCON Conference & Expo, Orlando, Florida, February 2018
Evaluation of Five Common Extraction Methods for Analysis of Human Remain Samples on Massively Parallel Sequencing Success
Kyleen Elwick, Xiangpei Zeng, Jonathan King, Bruce Budowle, Sheree Hughes-Stamm

American Academy of Forensic Sciences, Seattle, Washington, February 2018
A Novel Workflow for Identifying Phenotypic Polymorphisms in Detoxification Enzymes Associated with Drug Metabolism
Kari Graham, Ryan Gutierrez, Kyleen Elwick, Sheree Hughes-Stamm, Bobby LaRue

International Symposium on Human Identification, Seattle, Washington, September 2017
Comparative Tolerance of Two Massively Parallel Sequencing Platforms to Common PCR Inhibitors for Missing Person Cases
Kyleen Elwick, Xiangpei Zeng, Jonathan King, Bruce Budowle, Sheree Hughes-Stamm

International Symposium on Human Identification, Seattle, Washington, September 2017
Assessment of Impact of Extraction Methods on Analysis of Human Remain Samples on Massively Parallel Sequencing Success
Xiangpei Zeng, Kyleen Elwick, Jonathan King, Sheree Hughes-Stamm, Bruce Budowle

QIAGEN Asian Pacific Team in Singapore, Department of Chemistry Malaysia, & National Forensic Service Seoul Institute, August 2017

Evaluation of Four Commercial Quantitative Real-Time PCR (qPCR) kits with Inhibited & Degraded Samples
Amy Sorensen, Rachel Houston, Kyleen Elwick, Sheree Hughes-Stamm

International Society on Forensic Genetics, Seoul, Republic of Korea, August 2017
Comparative Tolerance of Two Massively Parallel Sequencing Platforms to Common PCR Inhibitors for Missing Person Cases
Kyleen Elwick, Xiangpei Zeng, Jonathan King, Bruce Budowle, Sheree Hughes-Stamm

International Society on Forensic Genetics, Seoul, Republic of Korea, August 2017
Worlds Converge™: A New Approach to Analyzing CE and MPS-based STRs
Kyleen Elwick, Carrie Mayes, Michelle Harrel, Sheree Hughes-Stamm

Association of Forensic DNA Analysts and Administrators, Austin, Texas, July 2017
Comparative Tolerance of Inhibited Samples for Short Tandem Repeat and Massively Parallel Sequencing Chemistries
Kyleen Elwick, David Gangitano, Sheree Hughes-Stamm

Association of Forensic DNA Analysts and Administrators, Austin, Texas, July 2017
Analysis of DNA from Post-blast Pipe Bomb Fragments for Identification and Determination of Ancestry
Esiri Tasker, Kyleen Elwick, Bobby LaRue, Charity Beherec, Rachel Houston, David Gangitano, Sheree Hughes-Stamm

3rd Human Identification Solutions (HIDS) Conference, Vienna, Austria, May 2017
Worlds Converge™: A New Approach to Analyzing CE and MPS-based STRs
Kyleen Elwick, Carrie Mayes, Michelle Harrel, Esiri Tasker, David Gangitano, Sheree Hughes-Stamm

3rd Human Identification Solutions (HIDS) Conference, Vienna, Austria, May 2017
NGM Detect-ing More from Skeletal Remains
Carrie Mayes, Kyleen Elwick, Michelle Harrel, David Gangitano, Sheree Hughes-Stamm

QIAGEN Investigator Forum, Prague, April 2017
Alternate Methods for the Collection, Preservation, and Processing of DNA Samples from Decomposing Human Cadavers: A DVI Strategy
Amy Sorensen, Rachel Houston, Kyleen Elwick, Kayla Ehring, David Gangitano, Sheree Hughes-Stamm

American Academy of Forensic Sciences, New Orleans, February 2017

Comparative Tolerance of Short Tandem Repeat and Massively Parallel Sequencing Chemistries to Inhibited Samples
Kyleen Elwick, David Gangitano, Sheree Hughes-Stamm

Summit Forum of Forensic Technology and Applications, Foshan, Guangzhou, China, November 2016
HID & MPS for Post-blast Bomb Fragments and Highly Inhibited Samples
Esiri Tasker, Kyleen Elwick, Bobby LaRue, Charity Beherec, Rachel Houston, David Gangitano, Sheree Hughes-Stamm

International Symposium on Human Identification, Minneapolis, September 2016
Comparative Evaluation of Three Commercial Quantitative PCR kits with Extremely Inhibited and Degraded Samples
Amy Sorensen, Kyleen Elwick, David Gangitano, Sheree Hughes-Stamm

International Symposium on Human Identification, Minneapolis, September 2016
Comparative Tolerance of Short Tandem Repeat and Massively Parallel Sequencing Chemistries to Inhibited Samples
Kyleen Elwick, David Gangitano, Sheree Hughes-Stamm

American Academy of Forensic Sciences, Las Vegas, February 2016
Optimization and Validation of the forensicGEM® Rapid Extraction Methods for High-Throughput Processing of Cotton Buccal Swabs
Kyleen Elwick, Sheree Hughes-Stamm, Kimberly Sturk Andreaggi, Michelle Peck

Maize Genetics Conference, Chicago, March 2013
Composition of nuclear mitochondrial DNA insertions on the short arm of chromosome 1 in B73
Kyleen Elwick, Ashley Lough

OTHER PRODUCTS

QIAGEN Application Note
Comparison of Four Commercial qPCR kits for Analyzing Inhibited & Degraded Forensic Samples

Forensic Magazine webinar sponsored by QIAGEN, June 2017
How Do Modern Quantification kits STACK-UP?

RESEARCH

National Institute of Justice grant 2015-DN-BX-K066, PI: Sheree Hughes Stamm 1/2016 – 8/2018
Enhanced Sample Preparation and Data Interpretation Strategies for Massively Parallel Sequencing for Human Identification in Missing Persons and DVI Casework

- Compared the tolerance of STR and MPS sequencing chemistries to common PCR inhibitors

- Evaluated bone and tissue samples spiked with high amounts of inhibitor and extracted with various methods
- Compared and assessed the sequencing chemistry and tolerance of challenging human remains with the S5 and MiSeq systems
- Experience collecting decomposed tissue samples at the Applied Anatomical Research Center (AARC)
- Learned how to properly use a bone saw and cut bone samples

WORK EXPERIENCE

McNair Scholars Program, Huntsville, TX 8/2015 – 8/2018

Graduate Mentor

- Conduct monthly meetings with 30 scholars to ensure they make progress in school and in research
- Prepare and conduct weekly workshops including: GRE Prep, Getting into Graduate School, Paying for Graduate School, How to Find Research Articles, and various other subjects
- Conduct summer workshops on public speaking and oral presentation skills

Armed Forces DNA Identification Laboratory, Dover, DE 6/2015 – 8/2015

Intern

- Worked on forensicGEM Optimization and Validation
- Worked with 3130xl, 3500xl, QIAxcel, GeneAmp 9700, ABI Veriti
- Worked with PP16, Fusion, Yfiler, Identifiler, forensicGEM
- Performed control region sequencing for mtDNA

Truman State University Department of Public Safety, Kirksville, MO 8/2013 – 5/2014

Ticket Writer

- Wrote tickets to students/faculty for parking violations on campus

Admittance Personnel

- Locked and unlocked classrooms and offices for students/faculty for activities and organizational meetings
- Worked closely with officers on self-defense and office work

Truman State University Biology Department, Kirksville, MO 1/2012 – 5/2012

Undergraduate Teaching Assistant

- Graded tests, quizzes, and assignments
- Was available for any questions or help students needed to complete assignments
- Prepared lab equipment for student labs

Truman State University Biology Department, Kirksville, MO 8/2011 – 12/2011

Animal Lab Assistant

- Prepared and transferred animals to clean cages
- Maintained the well-being of the animals
- Cleaned and sanitized dirty cages

PROFICIENCIES

Bone Preparation

- Dremel tool, SPEX CertiPrep 6750 Freezer/Mill Cryogenic Grinder

Extraction Protocols

- DNA IQ, DNA Investigator, PrepFiler BTA, Chelex, various organic extractions, purifications (QIAquick, Minelute, microcon filters, amicon filters)

Extraction Instrumentation

- EZ1 Advanced XL, QIAcube, AutoMate Express

DNA Quantitation Protocols

- Quantifiler Trio, Investigator Quantiplex Pro, PowerQuant, InnoQuant HY, SYBR Green, Qubit dsDNA HS Assay, Agilent 1000, 7500, and 12000 kits

Quantitation Instrumentation

- 7500 Real-Time PCR System, StepOne Real-Time PCR System, Qubit 2.0 Fluorometer, 2100 Bioanalyzer, Rotor-Gene Q

PCR Amplification Protocols

- GlobalFiler, Investigator 24plex QS, YFiler, PowerPlex Fusion

PCR Instrumentation

- GeneAmp 9700, Veriti, ProFlex

General Robots

- QIAgility

STR Genotyping Instrumentation

- 3500 Genetic Analyzer, 3130 Genetic Analyzer

Sequencing Protocols

- Chemistry: Ion AmpliSeq chemistry, Precision ID chemistry, ForenSeq DNA Signature Prep kit, NEBNext Fast DNA Fragmentation & Library Prep for Ion Torrent, Ion AmpliSeq and Precision ID DL8 Library Prep
- Primer panels: Identity panel, Early Access Beta-testing Degradation panel, Early Access Mixture ID panel, Early Access Mito panel

Sequencing Instrumentation

- Ion Chef System, Ion PGM System, Ion S5 System, MiSeq FGx

Software

- GeneMapper and GeneMapper ID-X, Converge 2.0

Computer Applications

- Linux, Windows, and Unix experience
  - SAMtools, VCFtools, Structure, R and R Studio, Statistica, STRait Razor, PGD Spider, Arlequin, MitoSave, IGV, Excel

CONTINUING EDUCATION

STRmix training – 10/2017

SHSU Internal Review Board Citi Training – 1/2016

Applied Anatomical Research Center – Volunteer/Sample Collection – 9/2015 – 12/2018

NSC Certified Bloodborne and Airborne Pathogen Training – 9/2015

To Hell and Back: **The Ethics of Stewardship and the Stewardship of Ethics –** 9/2015

Certified OSHA Training – 8/2014

Certified Bloodborne Pathogen Training – 8/2014

Independent Research using bioinformatics to analyze maize – Researcher – 8/2012 – 8/2014

MEMBERSHIPS, HONORS, AND ACTIVITIES

International Society for Forensic Genetics – Member - 3/2017 – Present
American Academy of Forensic Sciences – Member – 2/2015 – Present
Society of Forensic Science – Sam Houston State University Chapter – Graduate Student Organization Coordinator – 12/2014 – 12/2015, Member – 9/2014 – 12/2018
Sam Houston State University Saturdays at Sam – Volunteer – 11/2014, 4/2015, 11/2015, 4/2016, 11/2016