EXAMINATION OF NOVEL LOCI TO ENHANCE PATERNAL LINEAGE RELATIONSHIP ANALYSIS

INTERNSHIP PRACTICUM REPORT

Presented to the Graduate Council of the
Graduate School of Biomedical Sciences
University of North Texas
Health Science Center at Fort Worth
in Partial Fulfillment of the Requirements

For the Degree of

FORENSIC GENETICS

By

Laura Evans, B.S.
Fort Worth, Texas

May 2010

TABLE OF CONTENTS

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

The United States population is composed of a wide array of ethnicities that harbor different degrees of genetic variation. The Native American population has been shown to have reduced genetic diversity compared to other populations. The Alaskan Native American populations in particular have not been greatly studied for genetic diversity at the forensically-relevant genetic markers, leading to limitations in determining the strength of paternity, kinship, and identity testing results with statistical confidence. These historically isolated populations consist of several groups of people distinguished primarily by their geographical and linguistic affiliations. These groups include: Inupiat, Yupik, Aleut, Tlingit, Haida, Tsimshian, Eyak, and North Athabaskan. Native Americans account for 19% of Alaska's population which necessitates the need to generate STR (short tandem repeat) allele frequency databases in order to estimate the degree of relatedness among individuals and the weight of evidence in forensic casework (1).

Forensically-relevant genetic markers are used to assess the variation that exists among individuals. Some DNA loci contain repetitive sequences in the non-coding region. Loci with repeat units that are two to six base pairs in length are known as short tandem repeats (STRs). The number of repeats can be highly variable among individuals, which makes STRs useful for human identification purposes (2).

Autosomal STR loci can be amplified using commercially available kits. The AmpFℓSTR Profiler Plus and AmpFℓSTR Cofiler kits (Applied Biosystems, Foster City, CA) amplify 13 STR loci: D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317,

D7S820, TH01, TPOX, CSF1PO, and D16S539. The PowerPlex® 16 HS system (Promega Corp., Madison, WI) amplifies 15 STR loci plus Amelogenin which is a sex determination marker. These loci include: CSF1PO, D13S317, D16S539, D18S51, D21S11, D3S1358, D5S818, D7S820, D8S1179, FGA, TH01, TPOX, vWA, Penta D, and Penta E (3).

Y-STRs are short tandem repeats that occur specifically on the Y-chromosome, therefore are unique to the male population. They are inherited from father to son unchanged with the exception of the occasional mutation (13). Y-STRs are not specific to an individual but are more common to a family lineage. The forensically important Y-STR loci reside in the non-recombining portion of the hemizygous Y chromosome. The Y-STR loci, in combination, currently available through commercial kits are highly informative. Additional loci could be beneficial for differentiation within related or similar populations (4).

Knijff et al. (1997) used Y-STRs in order to compare closely related populations which could not be separated otherwise. Haplotype data was collected from males from four distinct European male groups: Dutch, Germans, Swiss, and Italian. Among a total number of 322 males, 211 different 7-locus Y-STR haplotypes were observed. The haplotype diversity for all four populations was close to 99% which indicated that most haplotypes were observed only once in the data set. Because of no recombination Y-STR haplotypes can be used to differentiate (or resolve), closely related populations at a significant level. They concluded that Y-STR loci were extremely useful for the purpose of comparing closely related populations (5).

Y-STR loci can be amplified using commercially available kits. The PowerPlex® Y system (Promega Corp., Madison, WI) amplifies 12 Y-STR loci: DYS19, DYS385 a/b, DYS389 I/II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, and DYS439. The AmpFℓSTR® Yfiler™ PCR Amplification Kit (Applied Biosystems, Foster City, CA) amplifies

17 loci: DYS456, DYS389I, DYS390, DYS389II, DYS458, DYS19, DYS385 a/b, DYS393, DYS391, DYS439, DYS635, DYS392, Y GATA H4, DYS437, DYS438, and DYS448 (9).

Budowle et al. (2002) conducted population studies on three Native Alaskan populations using short tandem repeat (STR) loci. Amplification by PCR was carried out using the AmpFℓSTR Profiler Plus and AmpFℓSTR Cofiler kits (Applied Biosystems, Foster City, CA). Allele distributions for 13 STR loci D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, D7S820, TH01, TPOX, CSF1PO, and D16S539 were determined for three Native Alaskan population groups: Athabaskans, Inupiats, and Yupiks. Genetic diversity was lower in the Native Alaskan populations compared with Caucasians and Africans, but all loci were highly polymorphic in all three Native Alaskan groups. The F_{ST} estimate overall 13 STR loci was 0.0309 for the Native Alaskan populations, which is consistent with the recommendations for isolated populations by the National Academy of Sciences Report (1996). It was determined that Athabaskans were more closely related to Apaches and Navajos than the other Native Alaskan groups. The F_{ST} estimate for Athabaskans, Apaches, and Navajos was 0.0180 and for Inupiats and Yupiks it was 0.0167 (1).

Budowle et al. (2005) used the PowerPlex® Y system (Promega Corp., Madison, WI) in order to type 12 Y-STR loci from a total of 2,443 male individuals distributed across the five North American population groups: African American, Caucasian, Asian, Hispanic, and Native American. All population samples were highly polymorphic across all loci. The Native American population groups demonstrated the lowest genetic diversity (Apaches, 97.0%; Navajo, 98.1%). Haplotype diversities were greater than 99.6% for the African Americans, Hispanics, Caucasians, and Asians (4).

Decker et al. (2007) looked at the impact of additional Y-STR loci in order to resolve common haplotypes and increase the power of discrimination between closely related

individuals. New Y-STR loci were investigated on a common set of 656 male samples representative of the major U.S. population groups. The 17 Yfiler™ (Applied Biosystems, Foster City, CA) loci along with 20 new Y-STR loci were evaluated on their ability to resolve the most common haplotype observed in 656 U.S. population samples. With the addition of these loci, all samples with the most common haplotype were able to be completely resolved and the power of discrimination between closely related individuals was improved (6).

Autosomal STR and Y-STR loci were used to evaluate the genetic variation that exists within one population of Native American Alaskans for this project. The autosomal STR loci used in this study were amplified using the PowerPlex® 16 HS system (Promega Corp., Madison, WI). The Y-STR loci used in this study were amplified using the AmpFℓSTR® Yfiler™ PCR Amplification Kit (Applied Biosystems, Foster City, CA).

The objectives of this study were to obtain autosomal STR data using the PowerPlex® 16 HS system (Promega Corp., Madison, WI) and to obtain Y-STR data using the AmpFℓSTR® Yfiler™ PCR Amplification Kit (Applied Biosystems, Foster City, CA) from a population of Native American Alaskans. In addition, the degree of genetic diversity was determined for assessing the utility of these genetic markers for forensic identification in Native Alaskans.

CHAPTER II

MATERIALS AND METHODS

Deoxyribonucleic acid (DNA) from 41 buccal swabs containing saliva and 7 (3 mm) punches containing blood was obtained from a Native American Alaskan population. These samples were extracted using the EZ1 Advanced XL instrument (Qiagen, Inc., Valencia, CA). The DNA was isolated from the lysates by binding to the silica surface of magnetic particles in the presence of a chaotropic salt. The magnetic particles were separated from the lysates using a magnet. The DNA was then washed and eluted (7).

The DNA from the 3 mm punches containing blood was extracted using the EZ1 Advanced XL instrument. First, 190 µl of diluted Buffer G2 were added to each of the 7 samples. Then 10 µl of proteinase K were added and mixed thoroughly by vortexing for 10 seconds. The samples were incubated at 56˚C for 15 minutes and then incubated at 95˚C for 5 minutes. The sample tubes were placed in the microcentrifuge. The DNA Purification ("Tip Dance") protocol was followed (7). The purification procedure took 17 minutes and 40 µl were the elution volume. The purified DNA was stored at 2-8˚C until use (7).

The DNA from the swabs containing saliva was extracted using the EZ1 Advanced XL instrument. First, 290 µl of diluted Buffer G2 were added to each of the 41 samples. Then 10 µl of proteinase K were added and mixed thoroughly by vortexing for 10 seconds. The samples was incubated at 56˚C for 15 minutes and then incubated at 95˚C for 5 minutes. The sample tubes were placed in the microcentrifuge. Spin baskets containing the swabs were placed into 1.5 ml tubes in order to spin down the samples (quick-spin). The tubes were subjected to centrifugation at 4.8 rpm for 2 minutes. The spin baskets were removed and the

liquid from the 1.5 ml tubes was placed back into the sample tubes. The swabs were placed into dry 1.5 ml labeled tubes and stored at 2-8˚C. The DNA Purification (Trace) protocol was then followed (7). The purification procedure took 16 minutes and 40 µl were the elution volume. The purified DNA was stored at 2-8˚C until use (7).

The quantity of DNA recovered from the 48 samples was determined by using the Quantifiler™ Human DNA Quantification Kit DNA (Applied Biosystems, Foster City, CA). The samples were normalized to 0.75 ng/µl of DNA in a final volume of 50 µl.

The 48 samples were amplified using the reagents contained within the PowerPlex® 16 HS System (Promega Corp, Madison, WI). The PowerPlex® HS 5X Master Mix, PowerPlex® 16 HS 10X Primer Pair Mix, and 9947A DNA were thawed. A PCR amplification mix was made for 56 reactions which included the 48 samples, a reagent blank, a negative control, two 9947A positive controls, plus 4 reactions to compensate for pipetting error. The PCR amplification mix contained: 295 µl of PowerPlex® HS 5X Master Mix, 147.5 µl of PowerPlex® 16 HS 10X Primer Pair Mix, and 973.5 µl of deionized water. Each reaction contained: 5 µl of PowerPlex® HS 5X Master Mix, 2.5 µl of PowerPlex® 16 HS 10X Primer Pair Mix, and 16.5 µl of deionized water. The PCR amplification mix was vortexed for 5-10 seconds then pipetted into the wells on a 96-well amplification plate. A 1 µl volume of DNA from each of the 48 samples was pipetted into separate wells of a 96-well amplification plate. A 1 µl volume of reagent blank was pipetted into its corresponding well on the 96-well amplification plate. A 1 µl volume of deionized water was pipetted into the negative control well on the 96-well amplification plate. A 1 µl volume of 9947A was pipetted into the two corresponding wells on the 96-well amplification plate. The 96-well amplification plate was placed on the GeneAmp® PCR System 9700 thermal cycler

(Applied Biosystems, Foster City, CA). The PowerPlex® 16 HS protocol and a reaction volume of 25 µl were selected (8).

After the samples were amplified using the PowerPlex® 16 HS System (Promega Corp., Madison, WI), the DNA fragments were separated by capillary electrophoresis on an Applied Biosystems 3130xl Genetic Analyzer. A loading cocktail was prepared by combining and mixing Internal Lane Standard 600 and Hi-Di™ formamide. A 0.5 µl volume of ILS 600 and a 9.5 µl volume of Hi-Di™ formamide were used. A 34 µl volume of ILS and a 646 µl volume of Hi-Di™ formamide were combined for a total of 680 µl. A 9 µl volume of the ILS and Hi-Di™ formamide were added into a 96-well 3130xl plate. A 1 µl volume of amplified DNA from the samples was added into separate wells of a 96-well 3130xl plate. A 1 µl volume of the reagent blank, negative control, and 9947A from the amplification plate was added to the 96-well 3130xl plate. A 1 µl volume of allelic ladder was added into each of the 4 corresponding wells on the 3130xl plate. The septum was placed onto the plate. The plate was centrifuged briefly. The samples were denatured at 95˚C for 5 minutes and then put in an ice-water bath for 5 minutes. The plate was placed onto the Applied Biosystems 3130xl Genetic Analyzer (8).

The samples were also amplified using the reagents contained in the AmpFℓSTR® Yfiler™ PCR Amplification Kit (Applied Biosystems, Foster City, CA). The AmpFℓSTR Yfiler Kit PCR Reaction Mix, AmpFℓSTR Yfiler Kit Primer Set, and AmpliTaq Gold® DNA Polymerase were thawed then vortexed for 3-5 seconds and centrifuged briefly. A Yfiler Master Mix was made using: 408 µl of AmpFℓSTR Yfiler Kit PCR Reaction Mix, 204 µl of AmpFℓSTR Yfiler Kit Primer Set, and 21.1 µl of AmpliTaq Gold® DNA Polymerase. Each reaction

consisted of: 5.8 µl of AmpFℓSTR Yfiler Kit PCR Reaction Mix, 2.9 µl of AmpFℓSTR Yfiler Kit Primer Set, and 0.3 µl of AmpliTaq Gold® DNA Polymerase. A 9 µl volume of Yfiler Master Mix was pipetted into each of the wells on the 96-well amplification plate. A 1 µl volume of reagent blank was pipetted into the corresponding well on the amplification plate. A 1 µl volume of deionized water was pipetted into the corresponding well for the negative control on the amplification plate. A 1 µl volume of 007 positive control DNA was pipetted into the two corresponding wells on the amplification plate. A 1 µl volume of DNA from each of the samples was pipetted into separate wells in the amplification plate. The Yfiler amplification plate was subjected to centrifugation and placed on the GeneAmp® PCR System 9700 thermal cycler (Applied Biosystems, Foster City, CA). The AmpFℓSTR® Yfiler™ protocol and a reaction volume of 10 µl were selected (9).

After the samples were amplified using the AmpFℓSTR® Yfiler™ PCR Amplification Kit (Applied Biosystems, Foster City, CA), the DNA fragments were separated by capillary electrophoresis using an Applied Biosystems 3130xl Genetic Analyzer. A loading cocktail was prepared by combining and mixing GeneScan-500 LIZ Size Standard and Hi-Di™ formamide. A 0.6 µl volume of GeneScan-500 LIZ Size Standard and a 9.2 µl volume of Hi-Di™ formamide were used as a master mix. A 40.8 µl volume of GeneScan-500 LIZ Size Standard and a 625.6 µl volume of Hi-Di™ formamide were combined for a total of 666.4 µl. A 9 µl volume of the GeneScan-500 LIZ Size Standard and Hi-Di™ formamide were added into a 96-well 3130xl plate. A 1 µl volume of amplified DNA from the samples was added to the 96-well 3130xl plate. A 1 µl volume of the reagent blank, negative control, and 007 control DNA from the

amplification plate were added to the 96-well 3130xl plate. A 1 µl volume of allelic ladder was added into each of the 4 corresponding wells on the 3130xl plate. The septum was placed onto the plate. The plate was subjected to centrifugation briefly. The samples were denatured at 95˚C for 5 minutes and then put in an ice-water bath for 5 minutes. The plate was placed onto the Applied Biosystems 3130xl Genetic Analyzer (9).

The electrophoretic data were analyzed using GeneMapper® ID v3.2 software (Applied Biosystems, Foster City, CA). Allele designations for each sample were determined by direct comparison of the sample fragments to allelic ladders developed for each kit by the manufacturer. Off-ladder variants were identified by evaluation of migration data to surrounding allelic ladder peaks and the internal lane standard.

Allele frequencies for each locus were determined for the 48 Native Alaskan samples amplified with the PowerPlex® 16 HS System (Promega Corp, Madison, WI) using the Genetic Data Analysis v 1.1 software. Expected heterozygosity (H_e) and observed heterozygosity (H_o) were calculated using the Genetic Data Analysis v 1.1 software (10). The power of discrimination (PD) was calculated using PowerStats V12 software (Promega Corp., Madison, WI) (11). The power of exclusion (PE) was calculated using the National Research Council (NRC) formula (12). The mean power of exclusion (MPE) was calculated using the Brenner and Morris formula (14). A Fisher Exact Test was used to test conformance to Hardy Weinberg Equilibrium (HWE) at all loci based on 3,200 shuffling experiments using the Genetic Data Analysis v 1.1 software (10).

Haplotype distributions were determined for the 48 Native Alaskan samples amplified with the AmpFℓSTR® Yfiler™ PCR Amplification Kit (Applied Biosystems, Foster City, CA). The frequency of occurrence was determined for each haplotype as well as the overall haplotype diversity.

CHAPTER III

RESULTS

A11 48 samples that were amplified using the PowerPlex® 16 HS System resulted in complete profiles. Allele frequencies for each of the following loci were determined for the 48 Native American Alaskan samples amplified with the PowerPlex® 16 HS System (Promega Corp., Madison, WI): CSF1PO, D13S317, D16S539, D18S51, D21S11, D3S1358, D5S818, D7S820, D8S1179, FGA, TH01, TPOX, vWA, Penta D, and Penta E (Tables 1A and 1B) (10). Power of exclusion (PE), mean power of exclusion (MPE), and power of discrimination (PD) were determined for all 15 STR loci (Tables 1A and 1B) (11,12). The values for PE ranged from 0.345 for TH01 to 0.902 for Penta E. The values for PD ranged from 0.550 for TH01 to 0.963 for Penta E. Observed and expected heterozygosities were determined for all 15 STR loci amplified using the PowerPlex® 16 HS System (Tables 1A and 1B). P-values were determined to detect any departures from Hardy Weinberg Equilibrium (HWE) expectations at the 15 loci using the Fisher Exact Test (Tables 1A and 1B) (10). All loci except for D13S317 yielded p-values > 0.05, and thus failed to reject the hypothesis that the population is in HWE.

Out of 48 samples amplified using the AmpFℓSTR® Yfiler™ PCR Amplification Kit, 5 samples resulted in incomplete profiles. Haplotype distributions were determined for the 48 Native Alaskan samples using the AmpFℓSTR® Yfiler™ PCR Amplification Kit (Applied Biosystems, Foster City, CA) (Tables 2A-2C and 3A-3C). There were a total of 36 different haplotypes among a group of 48 individuals. The haplotype diversity value was 0.989. This value is lower compared to haplotype diversity values determined for other population groups such as African Americans, Caucasians, Asians, and Hispanics which all had haplotype diversity

values greater than 0.996 (4). In order to determine haplotype diversity, the 5 samples that resulted in incomplete profiles were treated as separate haplotypes with a frequency of occurrence of 0.021. One haplotype was observed a total of 5 times and had a frequency of occurrence of 0.104. These five individuals could not be differentiated from one another because they had the same haplotype. Two haplotypes were observed 3 times each and had a frequency of occurrence of 0.063. Four haplotypes were observed 2 times each and had a frequency of occurrence of 0.042. The 29 other haplotypes were observed once and had a frequency of occurrence of 0.021 (Tables 2A-2C and 3A-3C).

Table 1A- Allele Frequencies and Statistical Parameters for a Native American Population (n=48)

Allele	CSF1PO	D13S317	D16S539	D18S51	D21S11	D3S1358	D5S818	D7S820
5								
6								
7		0.021					0.031	
8	0.01	0.135						0.292
9	0.083	0.104	0.063				0.01	0.083
9.3								
10	0.302	0.281	0.135				0.052	0.063
11	0.229	0.313	0.49				0.458	0.313
12	0.323	0.094	0.25	0.156			0.271	0.25
13	0.031	0.042	0.052	0.115			0.146	
14	0.021	0.01	0.01	0.156			0.031	
15				0.208		0.438		
16				0.063		0.438		
17				0.208		0.083		
18				0.042		0.042		
19				0.042				
20				0.01				
21								
22								
23								
24								
25								
26								
27								
28					0.01			
29					0.281			
30					0.229			
31					0.083			
31.2					0.177			
32.2					0.156			
33					0.01			
33.2					0.052			
PE	0.744	0.783	0.673	0.844	0.803	0.608	0.691	0.743
MPE	0.51	0.296	0.51	0.915	0.662	0.188	0.322	0.546
PD	0.875	0.913	0.844	0.934	0.919	0.784	0.859	0.881
He	0.751	0.791	0.68	0.853	0.811	0.615	0.698	0.752
Ho	0.75	0.604	0.75	0.958	0.833	0.500	0.625	0.771
P	0.339	0.041	0.967	0.631	0.448	0.14	0.216	0.771

PE, power of exclusion; MPE, mean power of exclusion; PD, power of discrimination; H_e, expected heterozygosity; H_o, observed heterozygosity; P, p-value for exact test

Table 1B- Allele Frequencies and Statistical Parameters for a Native American Population (n=48)

Allele	D8S1179	FGA	TH01	TPOX	vWA	Penta D	Penta E
5							0.01
6			0.031				
7			0.802				0.021
8	0.01		0.042	0.198			0.021
9			0.042	0.052		0.094	
9.3			0.083				
10	0.052			0.031		0.24	0.167
11				0.615		0.104	0.031
12	0.24			0.104	0.021	0.188	0.021
13	0.375					0.344	0.01
14	0.208				0.115	0.031	0.073
15	0.115				0.167		0.083
16					0.26		0.042
17					0.302		0.073
18					0.104		0.073
19		0.073			0.031		0.115
20		0.031					0.135
21		0.063					0.104
22		0.167					0.021
23		0.135					
24		0.135					
25		0.271					
26		0.094					
27		0.031					
28							
29							
30							
31							
31.2							
32.2							
33							
33.2							
PE	0.742	0.842	0.345	0.568	0.788	0.768	0.902
MPE	0.475	0.787	0.088	0.206	0.584	0.622	0.830
PD	0.876	0.944	0.550	0.762	0.907	0.896	0.963
He	0.75	0.851	0.349	0.575	0.796	0.777	0.912
Ho	0.729	0.896	0.354	0.521	0.792	0.813	0.917
P	0.143	0.881	0.363	0.53	0.51	0.384	0.254

PE, power of exclusion; MPE, mean power of exclusion; PD, power of discrimination; H_e, expected heterozygosity; H_o, observed heterozygosity; P, p-value for exact test

Table 2A- Haplotype Distributions

Sample Name	DYS456	DYS389I	DYS390	DYS389II	DYS458	DYS19
A.1170.P	15	14	24	30	14	13
A.1317.S	15	14	24	30	14	13
A.1831.S	15	14	24	30	14	13
A.2513.S	15	14	24	30	14	13
A.3123.S	15	14	24	30	14	13
A.1869.S	15	14	24	30	14	13
A.2086.S	15	14	24	30	14	13
A.3678.S	15	14	24	30	14	13
A.1518.S	15	14	25	30	14	13
A.1657.S	15	14	25	30	14	13
A.2821.S	15	14	25	30	14	13
A.1161.S	16	14	24	31	16	13
A.1373.S	16	14	24	31	16	13
A.1321.S	15	14	24	30	14	13
A.1577.S	15	14	24	30	14	13
A.1896.S	14	13	23	29	17	15
A.1922.S	14	13	23	29	17	15
A.1786.S	15	14	24	30	15	13
A.2366.S	15	14	24	30	15	13
A.1036.P	15	12	22	29	16	15
A.1073.P	15	13	23	28	16	14
A.1084.P	15	14	24	?	14	?
A.1085.P	16	14	24	31	16	13
A.1135.P	15	14	23	30	15	13

Table 2B- Haplotype Distributions

Sample Name	DYS385	DYS393	DYS391	DYS439	DYS635	DYS392
A.1170.P	13, 19	14	10	13	22	15
A.1317.S	13, 19	14	10	13	22	15
A.1831.S	13, 19	14	10	13	22	15
A.2513.S	13, 19	14	10	13	22	15
A.3123.S	13, 19	14	10	13	22	15
A.1869.S	13, 21	14	10	11	22	15
A.2086.S	13, 21	14	10	11	22	15
A.3678.S	13, 21	14	10	11	22	15
A.1518.S	13, 21	14	10	11	22	15
A.1657.S	13, 21	14	10	11	22	15
A.2821.S	13, 21	14	10	11	22	15
A.1161.S	15, 17	14	10	13	22	14
A.1373.S	15,17	14	10	13	22	14
A.1321.S	13, 20	14	10	13	22	15
A.1577.S	13, 20	14	10	13	22	15
A.1896.S	12, 15	12	9	11	24	11
A.1922.S	12, 15	12	9	11	24	11
A.1786.S	13, 19	14	10	13	22	15
A.2366.S	13, 19	14	10	13	22	15
A.1036.P	14	14	10	11	20	11
A.1073.P	11, 14	13	11	11	23	13
A.1084.P	?	14	10	11	?	?
A.1085.P	15, 17	15	10	13	22	14
A.1135.P	13, 20	14	10	11	23	?

Table 2C- Haplotype Distributions

Sample Name	Y_GATA_H4	DYS437	DYS438	DYS448	# of Times Observed	Frequency
A.1170.P	10	15	11	18	5	0.104
A.1317.S	10	15	11	18	5	0.104
A.1831.S	10	15	11	18	5	0.104
A.2513.S	10	15	11	18	5	0.104
A.3123.S	10	15	11	18	5	0.104
A.1869.S	10	15	11	18	3	0.063
A.2086.S	10	15	11	18	3	0.063
A.3678.S	10	15	11	18	3	0.063
A.1518.S	10	15	11	18	3	0.063
A.1657.S	10	15	11	18	3	0.063
A.2821.S	10	15	11	18	3	0.063
A.1161.S	11	15	10	21	2	0.042
A.1373.S	11	15	10	21	2	0.042
A.1321.S	10	15	11	18	2	0.042
A.1577.S	10	15	11	18	2	0.042
A.1896.S	11	14	10	21	2	0.042
A.1922.S	11	14	10	21	2	0.042
A.1786.S	10	15	11	18	2	0.042
A.2366.S	10	15	11	18	2	0.042
A.1036.P	12	16	10	21	1	0.021
A.1073.P	13	15	12	19	1	0.021
A.1084.P	10	15	11	?	1	0.021
A.1085.P	12	15	10	21	1	0.021
A.1135.P	10	15	11	18	1	0.021

Table 3A- Haplotype Distributions

Sample Name	DYS456	DYS389I	DYS390	DYS389II	DYS458	DYS19
A.1215.P	16	14	24	31	16	13
A.1376.S	15	13	22	30	17	13
A.1390.S	17	13	25	30	17	13
A.1391.S	16	12	24	28	17	14
A.1516.S	15	14	24	30	14	13
A.1604.S	15	14	24	30	14	13
A.1662.S	15	14	23	31	16	13
A.1666.S	15	14	24	30	14	13
A.1700.S	16	13	23	31	16	13
A.1752.S	15	14	24	30	14	13
A.1758.S	15	13	24	29	16	14
A.1761.S	17	13	23	29	17	14
A.1897.S	17	13	25	29	18	14
A.1930.S	15	14	24	30	14	14
A.1968.S	15	13	24	29	14	13
A.1970.S	14	12	23	28	17	15
A.1999.S	14	13	23	30	14	15
A.2017.S	15	13	22	30	17	13
A.2024.S	15	14	24	30	14	13
A.2253.S	16	13	25	29	18	14
A.2298.S	17	13	25	30	17	14
A.2544.S	17	14	24	31	15	13
A.2694.S	15	14	24	30	17	15
A.2927.S	16	14	24	31	17	13

Table 3B- Haplotype Distributions

Sample Name	DYS385	DYS393	DYS391	DYS439	DYS635	DYS392
A.1215.P	15, 18	14	10	13	22	14
A.1376.S	15	13	10	12	22	14
A.1390.S	16, 18	14	10	12	22	15
A.1391.S	11, 14	13	12	13	23	?
A.1516.S	12, 19	15	10	11	22	16
A.1604.S	13, 20	14	10	13	22	?
A.1662.S	14, 17	13	10	12	23	14
A.1666.S	13, 17	14	10	11	22	14
A.1700.S	14, 17	13	10	11	22	14
A.1752.S	12, 19	15	10	11	22	15
A.1758.S	11, 14	13	11	13	24	13
A.1761.S	11, 14	13	11	12	24	13
A.1897.S	11, 13	13	11	12	23	14
A.1930.S	13, 21	14	10	11	22	15
A.1968.S	12, 19	15	10	11	22	15
A.1970.S	14	13	10	11	21	11
A.1999.S	14, 15	15	10	8	21	12
A.2017.S	14, 15	13	10	12	22	14
A.2024.S	12, 20	15	10	11	22	15
A.2253.S	11, 14	13	11	12	23	13
A.2298.S	11, 14	13	11	12	23	13
A.2544.S	16, 19	13	10	14	21	11
A.2694.S	11, 15	13	11	12	23	13
A.2927.S	16, 17	15	10	13	22	14

Table 3C- Haplotype Distributions

Sample Name	Y_GATA_H4	DYS437	DYS438	DYS448	# of Times Observed	Frequency
A.1215.P	11	15	10	21	1	0.021
A.1376.S	12	14	11	19	1	0.021
A.1390.S	11	15	10	21	1	0.021
A.1391.S	13	14	12	19	1	0.021
A.1516.S	10	15	11	18	1	0.021
A.1604.S	10	15	11	18	1	0.021
A.1662.S	11	15	11	20	1	0.021
A.1666.S	10	15	11	18	1	0.021
A.1700.S	12	14	11	?	1	0.021
A.1752.S	10	15	11	18	1	0.021
A.1758.S	11	15	12	19	1	0.021
A.1761.S	11	14	12	19	1	0.021
A.1897.S	11	14	12	18	1	0.021
A.1930.S	10	15	11	18	1	0.021
A.1968.S	10	15	11	18	1	0.021
A.1970.S	11	16	10	20	1	0.021
A.1999.S	10	14	10	20	1	0.021
A.2017.S	12	14	11	19	1	0.021
A.2024.S	10	16	11	18	1	0.021
A.2253.S	11	14	12	17	1	0.021
A.2298.S	12	15	12	19	1	0.021
A.2544.S	12	14	10	20	1	0.021
A.2694.S	12	15	12	19	1	0.021
A.2927.S	11	15	10	21	1	0.021

CHAPTER IV

DISCUSSION

D13S317, which had a p-value of 0.041, showed a significant departure from Hardy Weinberg Equilibrium (HWE). There are several factors that could have contributed to the outcome of this result. In order for a locus to show departure from HWE, the p-value has to be less than 0.05. This means that out of 20 loci, it is estimated that approximately one will show departure from HWE. Since this study examined 15 loci, one would expect for approximately one loci to show departure from HWE. One factor could have been the population sample size. This study examined a small subset of a larger sample. Since there were only 48 samples, it was possible that the data obtained resulted due to sampling error. In order to correct for this, more samples needed to be examined. There could have been a false level of homozygosity due to allelic drop out. A true heterozygote would show two electrophoretic peaks. When allelic drop out occurs, only one peak may be visible and therefore appears as a homozygote. Allelic drop out could have occurred due to a primer binding site mutation for that locus (15). When the alleles drop out, it appears that there is an excess of homozygotes when actually there is not.

REFERENCES

1. Budowle, Bruce, Chidambaram, Abirami, Strickland, Leanne, Beheim, Chris W.,Taft, George M., and Ranajit Chakraborty. Population studies on three Native Alaska population groups using STR loci. *Forensic Science International*. 129 (2002), 51-57.

2. Hammond, H.A., Jin, L., Zhong, Y., Caskey, C.K., and R. Chakraborty. Evaluation of 13 short tandem repeat loci for use in personal identification application. *American Journal of Human Genetics.* 1994. 55(1): 175-89.

3. Birch, D.E., Kolmodin, L., Wong, J., Zangenberg, G.A., Zoccoli, M.A., McKinney, N., Young, K.K.Y., and Laird, W.J. (1996) *Nature*, 381, 445-446.

4. Budowle, Bruce, Adamowicz, Mike, Aranda, Xavier G., Barna, Charles, Chakraborty, Ranajit, Cheswick, Dan, Dafoe, Bradley, Eisenberg, Arthur, Frappier, Roger, Gross, Ann Marie, Ladd, Carll, Lee, He-Suk, Milne, Scott C., Meyers, Carole, Prinz, Mechthild, Richard, Melanie L., Saldanha, Gabriela, Tierney, Amy A., Viculis, Lori, and Benjamin E. Krenke. Twelve short tandem repeat loci Y chromosome haplotypes: Genetic analysis on populations residing in North America. *Forensic Science International.* 28 May 2005. 150(1), 1-15.

5. Kniff, P. de, Kayser, M.,Caglia, A.,Corach, D.,Fretwell, N.,Gehrig,C., Graziosi, G., Heidorn, F.,Herrmann, S.,Herzog, B.,Hidding, M., Honda, K., Jobling, M., Krawezak, M., Leim, K., Meuser, S., Meyer, E., Oesterreich, W., Pandya, A., Parson, W., Penacino, G.,Perez-Lezaun, A., Piccinini, A., Prinz, M., Schmitt, C.,Schneider, P.M., Szibor, R., Teifel-Greding, J., Weichhold,G., and L. Roewer. Chromosome Y microsatellites: population Genetic and evolutionary aspects. *Int J Legal Med* (1997) 110:134-140.

6. Decker, A.E., Kline, M.C., Vallone, P.M., and J.M. Butler. The impact of additional Y-STR loci on resolving common haplotypes and closely related individuals. *Forensic Science International:Genetics* (2007), 1(2): 215-217.

7. Qiagen, Inc. (2009) *EZ1® DNA Investigator Handbook.* Valencia, CA: Qiagen, Inc.

8. Promega Corporation *PowerPlex® 16 HS System Technical Manual.* Madison,Wisconsin: Promega Corporation.

9. Applied Biosystems *AmpFℓSTR® Yfiler™ PCR Amplification Kit User's Manual*. Foster City, California: Applied Biosystems.

10. Lewis, P. O., and Zaykin, D. 2001. Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.1. Free program distributed by the authors over the internet from http://hydrodictyon.eeb.uconn.edu/people/plewis/software.php.

11. Promega Coorporation *Powerstats V12*. Madison, Wisconsin: Promega Corporation.

12. National Research Council. 1996. The Evaluation of Forensic DNA Evidence. Chapter 4, pp. 89-124.

13. Ge, Jianye, Budowle, Bruce, Aranda, Xavier G., Planz, John V., Eisenberg, Arthur J., and Ranajit Chakraborty. Mutation rates at Y chromosome short tandem repeats in Texas populations. *Forensic Science International: Genetics*. June 2009. 3(3), 179-184.

14. Brenner, C., and J. Morris. Paternity index calculations in single locus hypervariable DNA probes. International symposium on human identification. 1989.

15. Leibelt, Craig, Budowle, Bruce, Collins, Patrick, Daoudi, Yasser, Moretti, Tamyra, Nunn, Gary, Reeder, Denis, and Rhonda Roby. Identification of a D8S1179 primer binding site mutation and the validation of a primer designed to recover null alleles. *Forensic Science International*. 5 May 2003. 133(3), 220-227.