POPULATION AND MUTATION ANALYSIS OF 14 Y- CHROMOSOMAL STR LOCI IN THE PHILIPPINES

A Thesis Proposal
Presented to
The Faculty of Biology Department
De La Salle University
Taft Avenue, Manila

In Partial Fulfillment
of the Requirements for the Degree
of Master Science in Biology

by
Sheila Marie T. Angustia

April 2011

ABSTRACT

SHEILA MARIE T. ANGUSTIA. Population and Mutation Analysis of 14 Y Chromosomal STR Loci in the Philippines. De La Salle University-Manila.

Adviser : Dr. Ma. Carmen Ablan-Lagman

In forensic genetics, Y chromosomal STR markers are typically used for patrilineage tracking and forensic investigation including victim recognition in mass disaster and offender identification. Though underutilized at the moment, these markers have shown to be advantageous in analyzing male DNA in mixtures containing high female DNA fraction and suitable for deficiency paternity testing. To be useful however, STR markers must first be evaluated in the population for which they will be used.

In this study, 14 Y chromosomal STR markers (DYS19, DYS385 I/II, DYS389 I/II, DYS390, DYS391, DYS392, DYS 393, DYS 437, DYS 438, DYS 439, DYS 321, Y-A71, and Y-A72) will be assessed for its variability and mutability in Philippine population. Locus information, population variation data and mutability information will be gathered to come up with reliable forensic interpretation by improving forensic probability calculation and understanding of diversity data. Information obtained from this research will verify the usefulness and informative power of this set of markers for routine casework.

CHAPTER I

INTRODUCTION

Forensic DNA analysis is one of the most demanding investigative tools being used due to its social relevance. Advancements in this field were the outcome of the continuous development of functional methods in molecular biology. While autosomal microsatellite analysis serves as the most popular technique, other molecular techniques such as Y chromosomal microsatellite (Y STR), mitochondrial DNA (mtDNA) and single nucleotide polymorphism (SNPs) analyses are being explored. In line with these developments, forensic laboratories are being encouraged to use new techniques that would aid them in their quest for truth and justice.

In the Philippines, Philippine National Police Crime Laboratory (PNPCL), an operational support unit of the Philippine National Police (PNP) is mandated by the law to provide "examination, evaluation and identification of physical evidences involved in crimes with primary emphasis on their medical, chemical, biological and physical nature" as stated in Republic Act No. 8551 otherwise known as "An Act Providing for the Reform and Reorganization of the Philippine National Police and for other purposes. The PNP Crime Laboratory DNA Analysis Section was established in 2000 to perform DNA examination on evidence as well as testify to both criminal and civil cases in court.

Currently, PNPCL uses 18 autosomal STR markers (D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, THO1, TPOX, FGA, F13B, FES/FPS, CSF1PO, VWA, LPL, Penta D and Penta E) for its routine forensic casework. However, due to variable nature of forensic evidence as well as the limitation set by autosomal STR

technique in sexual assault cases involving mixed DNA profiles as well as deficient kinship investigation, the laboratory endeavors capability enhancement thru inclusion of other forensic DNA method particularly Y chromosomal microsatellite (Y STRs) analysis in routine casework typing.

The male specificity of Y chromosomal STR markers renders its significance as supplementary markers relevant to processing of sexual assault evidence which typically involves DNA mixtures containing low fraction of male DNA versus female DNA (Butler, 2006). In addition, owing to its paternal inheritance, it can appropriately be used in deficiency paternity analysis (i.e. father is not available and inferences are made by reference from male lineage such as brother (Gill et al, 2001).

In forensic case investigation, validity, the probability that an examination will accurately identify true matches and true nonmatches is an important aspect. Any method that will be implemented must be borne out of careful and thorough study. Thus, addition of Y chromosomal markers to routine DNA typing has to undergo careful evaluation.

Because Y chromosome is inherited largely as a single non recombining molecule, it is necessary to assess the variability of allelic combinations (Ballard, 2005) in the particular population for which they will be used. Differences in genetic diversity between populations will necessitate microgeographic sampling when calculating probabilities for forensic application. Also, reliable locus specific mutation rates must be determined to carefully choose loci that would minimize error rate in kinship analysis and sample identification while obtaining maximum discriminatory power as well as improve forensic probability calculations and interpretations of diversity data (Burgarella, 2011).

Furthermore, incorporation of Y STR analysis to routine case examination requires development of DNA typing procedures which ensures that reliable results obtained from particular set of protocols and associated materials (Butler *et. al* 2004). Indeed, methods have to be optimized and tested for its applicability to forensic casework.

Objectives of the Study

The main objective of this study is to determine allelic frequency, genetic diversity and mutation rate of 14 Y STR loci in the Philippine populace to gain information on this particular population relevant to forensic casework.

Specifically, this study aims to evaluate the variability of 14 Y STR markers in the Philippine population to ascertain its usefulness in paternity and identity examination by generating allelic frequencies and measurement of genetic diversities per locus. It also endeavors to determine the occurrence of spontaneous mutations between father and son pairs relevant to kinship analysis and interpretation. Finally, this research intends to develop cost effective protocols for routine DNA examination.

Scope and Limitations

The current study will be limited to the analysis of 14 Y-STR loci (DYS19, DYS385 I/II, DYS389 I/II, DYS390, DYS 391, DYS 392, DYS 393, DYS 438, DYS 437, DYS439, YA7-1, YA7-2 and DYS321) in the Philippine population. DNA profiles of 350 unrelated Filipinos will be used for population studies while 150 father and son pairs will be included in mutational analysis. Development of laboratory protocols will

depend on the technical capabilities of the PNP Crime Laboratory DNA Analysis Section. All personal information of cases used in this study will be kept confidential as it may be prejudicial to the case and in deference with legal proceedings.

CHAPTER II

REVIEW OF RELATED LITERATURE

DNA IN FORENSICS

The human genus has 3 billion base pairs. Basically, the permutation of these building blocks along the length of the genome determines each person's genetic characteristics. While any two human genomes are similar for the majority of the bases (99.9%); DNA samples from two unrelated individual differs at an average of only one base per thousand (0.1%) (NIJ 2000). Nevertheless, this proportion has 3 million base pairs enough to bring distinctiveness to each person such, with the exception of identical twins; no two individuals have the same genetic permutations.

In forensic identification, scientists scan DNA regions (locus) that vary from person to person and use the gathered data to create a DNA profile of that individual (DNA fingerprint). Because there is an extremely small chance that another person has the same DNA profile for a particular set of 13 regions (USDOE, 2009), this criterion is being implemented by some laboratories. Other forensic use of DNA involves crime offenders' identification, mass disaster victims' recognition as well as paternity and kinship verification.

DNA PROFILING

Genetic identity testing involves identifying the patterns of genetic material that are unique to almost every individual. The terms DNA fingerprinting, DNA profiling,

and DNA typing referring to the technique used in genetic identity testing process are used interchangeably despite some technical differences (Saad 2004).

Prior to use of DNA fingerprinting, blood surface markers including ABO typing were utilized to identify people (NIJ 2000). However, because criminal cases require higher standard of proof, more powerful tools were necessary.

Researches in biomedical field led to the recognition of certain DNA regions that were highly variable between individuals (NFSTC). In 1980, Wyman and White found a hypervariable (highly polymorphic), non-coding locus in the human genome which happens to be a restriction enzyme site. They demonstrated the use of restriction enzyme to cut strands of DNA and produce fragments of variable lengths (Wyman et. al 1980). With this study, Wyman and White was able to describe the first polymorphic restriction fragment length polymorphism (RFLP) marker and proposed methods for mapping the human genome by using restriction enzymes (NFSTC).

In 1985, the relevance of DNA analysis by using RFLP technology to criminal investigation was first described by Alec Jeffreys and his colleagues (NIJ 2000, Butler 2001). In his researches, he found out that in certain regions of the DNA, the number of repeats varies from one individual to another (NIJ 2000, US DNA Initiative). By developing a technique to examine these regions termed as variable number tandem repeats or VNTR regions, Dr. Jeffreys were able to perform human identity tests (US DNA initiative). This technique was called restriction fragment length polymorphism (RFLP) as it uses restriction enzyme to cut the regions of DNA surrounding the VNTRs (NIJ 2000). With such procedure, Dr. Jeffreys was able to demonstrate the significance

of DNA fingerprinting on paternity testing and human identification (Jeffreys 1985). Moreover, the same procedure was used to exonerate one suspect implicated into two rape cases that happened three years apart (1983 and 1986) in the village of Enderby, Leicestershire, UK. He was able to prove that another man, Colin Pitchfork, committed the said crimes by showing that this Colin's profile matches with the evidence obtained from the two rape victims (NIJ 2000, US DNA Initiative, EMFPU 2009, Derksen 2003).

The introduction of DNA profiling recognized the importance of physical evidence specifically biological crime scene evidence as an independent and objective proof that will link the perpetrator or the victim to a crime, disprove alibi, and build on important investigative leads (Catalin *et. al*). In addition, DNA collection and analysis gives the criminal justice field a powerful tool for reducing the impact of mistaken eyewitness identifications which results to convicting the guilty and exonerating the innocent (Klobuchar et. al 2005, NIJ 2002).

DNA techniques were preferably used over the earlier systems due to several reasons: (1) DNA techniques is based on the genetic make-up itself while serological test "involves a gene product and therefore maybe only an indirect reflection of DNA condition" ; (2) DNA methods avoids complications of dominance and recessiveness wherein conditions of heterozygozity (Aa) cannot be distinguished phenotypically with conditions of dominance(AA); (3) DNA is more thermally stable than protein which facilitates ease in development of protocols as well as processing of evidence; (4) DNA lasts longer than protein when collected and stored properly thereby reinforcing the validity of DNA profiling techniques even on evidence from "cold cases" or old cases

previously thought unsolvable (NIJ 2002); and, (5) DNA renders higher variability compare to protein markers.

STR ANALYSIS

As mentioned previously, DNA regions in the genome vary in length (Goodwin 2007). Individuals have tandem repeats, small number of repeated base sequences which vary in numbers along the stretch of DNA. Variability arises from the difference in the number of repeating units that result in alleles of different lengths. For this reason, tandem repeat polymorphism also called length polymorphism (Goodwin 2007) was explored as the basis for analysis and identity in forensics.

Minisatellites called variable number tandem repeats (VNTR) and microsatellites called short tandem repeats (STR) are two principal categories of tandem repeats found in the genome. Although both of which are composed of repeating units of bases, minisatellites are longer than microsatellites.

Microsatellites, also known as simple sequence repeats or Short Tandem Repeats (STRs) are consist of 1-5 bp repeated typically 5-30 times. Most microsatellite can efficiently be amplified by standard PCR since the repeat regions are shorter than 100 bp. It can show substantial polymorphism and are abundant throughout the human genome. In addition, microsatellites are particularly suitable for analyzing forensic specimens containing degraded and or limited amount of DNA (Tamaki and Jeffreys, 2005).

Analysis of short tandem repeat markers (STRs) has become one of the most powerful tools for DNA typing in forensic casework (Asamura *et. al* 2007). It allows reliable identification of individuals as well as classification of biological traces (Paabo

et. al 1989).With microsatellites, small PCR products can be sized with precision by polyacrylamide gel electrophoresis (PAGE). On the other hand, because spurious shadow or stutter bands are often observed at dinucleotide repeat loci making interpretations difficult, current typing systems use microsatellites with repeats 4bp long to reduce incidence of stuttering (Tamaki *et. al* 2005).

The STR approach allows very high throughput (more samples processed at a given time) via multiplex PCR (single tube PCR reactions that amplify multiple loci). In addition, fluorescent detection systems developed allowed substantial automation of gel electrophoresis turning DNA STR analysis and profile interpretation fast and reliable (Butler *et. al* 2001).

STR analysis is also more sensitive than other methods and can recover information even at the level of a single cell. The unambiguous assignment of alleles makes the methods suitable for the development of databases (Edwards *et. al* 1991)

The Federal Bureau of Investigation (FBI) uses a standard set of 13 specific STR regions for CODIS. CODIS is a software program that operates local, state, and national databases of DNA profiles from convicted offenders, unsolved crime scene evidence, and missing persons. The odds that two individuals will have the same 13-loci DNA profile is about one in a billion (US DOE 2009).

Y-Chromosome Analysis

Y chromosome is one of the smaller human chromosome with an estimated average size of 60 million base pairs (Mb). The heterochromatic region in the Y chromosome is located in the distal long arm (Yq) and constitutes more than half of the

chromosome in some normal males but is virtually undetectable in others (Gusmao *et. al* 1999). Short tandem repeat polymorphisms on the non recombining part of the Y chromosome are inherited in the form of compound Y haplotypes and can only evolve by mutation (Brinkmann, 1999).

Y-chromosomal polymorphic markers have been extensively investigated in forensic medicine and have been recognized as a powerful tool for analyzing mixed forensic stains and for paternity testing (Rodig *et. al* 2007, Goes *et. al* , 2005). These markers are inherited along a male lineage, and are characterized by a moderate number of polymorphic loci.

Because the Y-STRs reside on male specific DNA, female DNA is not reactive, therefore, the Y chromosome DNA testing enables examination of the male-specific portion of biological evidence. In cases of sexual assaults with a male perpetrator and a female victim, Y-STRs improve the chances of detecting low levels of the perpetrators DNA in a high background of the female victim's DNA (Johns *et al* 2006). In fact, Y chromosome STR has been shown to be a powerful tool in processing cases involving azoospermic semen (Dekairelle and Hoste 2000).

Studies have shown that Y STR analysis was successful in elucidating male profile (48%) despite a negative result for autosomal analysis depending on the on the amount of DNA available and the quality of DNA extracted (Dekairelle and Hoste 2000). Application of Y chromosomal STRs to forensic casework becomes more informative than autosomal STR typing (plus amelogenin) alone especially if male suspect is unknown or when mixed stains from more than one male have to be investigated

(Wurmb-Schwark *et.al* 2003). Sexual assault evidence having minor male DNA component may yield male specific profile that can be compared with the suspect (Betz et al., 2001). Furthermore, in forensic cases involving multiple male aggressors, autosomal STR analysis often provide inconclusive results while Y STR markers can be used to estimate the number of donors based on the presence of additional alleles present in the Y chromosome profile.

Y chromosome analysis can also benefit missing person investigations as it extends the range of potential reference samples. Since fathers pass their Y chromosome onto their sons unchanged (except for an occasional mutation), all males in a paternal lineage will possess a common Y chromosome haplotype and therefore can be used as reference samples in cases of deficient paternity testing (Goes *et. al* , 2005, Corach *et. al* , 2001).

A set of Y chromosome short tandem repeat (Y-STR) loci have been selected for use in human identification applications. Y-STR loci recommended by the SWGDAM (Scientific Working Group on DNA Analysis Methods) Y chromosome subcommittee in 2003 include the following: DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439, and the multi-copy locus DYS385 a/b (Butler, 2006). Commercial kits now exist that can co-amplify the entire core Y-STR loci in a single multiplex reaction (DNA Initiative).

Determination of Y STR haplotype frequencies in different populations is highly significant for correct interpretation of genetic profile matches in paternity and forensic casework (Goes *et. al* , 2005, Goedbloed et al, 2009). Occurrence of population

substructure renders the need for microgeographic sampling when calculating for match as well as paternity probabilities. Thus, Y STR markers included in forensic casework has to be evaluated for its variability and usefulness in the population for which they will be used by obtaining haplotype diversities and understanding population structure.

In situations where direct comparison between evidence and a suspect is being made, mutation rates are not important. However, spontaneous mutation in the germline between putative father at any locus in kinship analysis may be misinterpreted due to differences in father and son pairs (Lee *et. al*, 2009). Thus, reliable mutability knowledge is needed to carefully choose loci to minimize error rate in kinship analysis and sample identification while obtaining maximum discriminatory power (Burgarella et al, 2011).

In addition, many regions of Y-chromosome are duplicated or even triplicated in some individuals thus complicating the potential mixture interpretation where multiple peaks can be interpreted as mixed profiles (Vieira *et. al,* 2008). Thus, in addition to haplotype frequencies estimates of mutation rates are essential in making decisions concerning paternity testing and forensic casework. The more accurate method of estimating Y STR mutation rates is the direct observation of transmission between father and son as long as large numbers of genetic transfers (meioses) are covered by testing a large number of father and son pairs (Ballantyne et. al, 2010). The proportion of allele mismatches in father-son transmission is currently the most widely used approach to obtain mutation rates for YSTRs. Owing to the low values of mutation rates, large number of father and son pairs must be genotyped to obtain accurate estimates (Burgarella et al, 2011).

DNA DATABASE

While forensic DNA profiling results was immediately accepted in court, the admissibility on the methods and the practices of computing DNA profile frequencies was legally challenged because of its failure to meet the Frye or other standards for admissibility of scientific evidence (NIJ 2000, Derksen 2003). The methods challenge focused on reliability and validity testing while statistics debate focused on the reliability of the assumption of independence for applying the product rule to derive estimates of DNA profile frequencies (Budowle 2000). In response to such, the technical standards for forensic DNA testing were improved greatly and databases were established to generate statistical frequencies that are more extensive and more representative (Budowle 2000, NIJ 2000). As a result of the forensic community's effort to support the technology (by research and data analysis), DNA typing has met both Frye and Daubert criteria for admissibility (Budowle 2000). In fact, on 1996, a study by the National Research Council (NRC 1996) mentioned that "The state of profiling technology and the methods for estimating frequencies and related statistics have progressed to the point where admissibility of properly collected and analyzed data should not be in doubt."

DNA PROFILING

DNA profiling is a process which begins when a minute sample of DNA is taken from human tissue and ends when a computerized alpha numeric value was obtained from the visualized output of the DNA analytical process (Interpol 2009, Leriche et. al 1998). The DNA profile obtained from a person (reference) and the DNA profile obtained from sample retrieved at the crime scene (questioned) can be evaluated to

establish association. Comparing these profiles can eliminate innocent people, but can also provide a strong indication of guilt (Interpol 2009).

DNA profiling involves a series of molecular biology techniques aimed at determining the sizes of discrete DNA fragments that contain hypervariable target sequences. Because the technology is relatively new, technical possibilities are still expanding such that a variety of 'standard' techniques are used at each step in DNA profiling. This implies that different methods are being implemented by different laboratories. However, despite having different platforms and technicalities, principles of analysis are necessarily the same.

Physical Evidence

The initial stages of DNA analysis are crucial to the successful resolution of criminal investigation. Contamination of DNA sample during collection produces ambiguous information while sample deterioration as a result of environmental insults results to incomplete or failure to produce a profile. Such conditions compromise the examination, thus proper collection and preservation is necessary (Butler 2001, Catalin et.al, Goodwin 2007). In addition, samples to be collected should be carefully chosen as well to prevent needless redundancy in the evidence for a case (Butler 2001).

All biological evidence is subject to deterioration. Careful collection and storage of this evidence will help ensure that evidence is preserved so that useful information can be obtained from the analysis. Contaminants such as dirt, grease, dyes from fabrics and other substances can seriously compromise the DNA typing process (Catalin *et. al*).

Many biological samples are ideal substrates for the growth of bacteria and other microorganisms, which can degrade DNA. In addition, exposure to ultraviolet light in the form of sunlight can induce pyrimidine dimers, which can inhibit PCR. Moreover, other PCR inhibitors can be introduced by the environment (e.g., humic acid in soil), by the substrate on which the sample is deposited (e.g., indigo dye from denim) or by the sample itself (e.g., hematin from blood samples). Therefore, ability to purify DNA free of these inhibitors can be critical to the success of DNA profiling (Promega).

Environmental insults will not change DNA allele A to B but they can adversely affect the ability of the scientist to obtain complete DNA profile from the sample (Catalin et. al). Therefore, evidence must be stored properly. Most biological sample is best preserved when stored dry and cold because these conditions reduce the rate of bacterial growth and degradation of DNA (Butler 2001).

Because extremely small samples of DNA can be used as evidence, greater attention to contamination issues is necessary when identifying, collecting and preserving DNA evidence. DNA evidence can be contaminated when DNA from another source gets mixed with the DNA relevant to the case (Catalin et. al, Goodwin 2007). Since PCR replicates or copies DNA in the evidence sample, the introduction of contaminants or other unintended DNA to an evidence sample can be problematic. Due to high sensitivity of profiling technique, the PCR process will copy whatever DNA is present in the sample; it cannot distinguish a suspects DNA and DNA from another source (Catalin et. al., 2007).

DNA Extraction

DNA extraction is the process of releasing DNA form the cell (NFSTC). Biological specimen contains other substances besides DNA. Macromolecules that protect DNA inside the cells as well as other cellular components interfere with DNA analysis process; thus, DNA molecules must be separated from other cellular material before they can be examined (Butler 2001). DNA extraction methods separate proteins and other cellular materials from DNA molecules.

There are two goals in DNA extraction. First, the process must extract enough DNA from a sample to perform DNA profiling and second, the process must yield pure enough extracted DNA for subsequent analysis. The complexity of extraction process depends on the nature of sample (Goodwin 2007).

There are a variety of methods in extracting the DNA. The protocol depends on a number of factors including sample type and quantity, speed of extraction procedure, success rate of forensic sample and the cost of procedure.

DNA extraction procedure has three stages: 1) disruption of cellular membranes, resulting in cell lysis, (2) protein denaturation, and 3) DNA separation from denatured protein and other cellular components.

Some of the extraction methods commonly used in the laboratory are the following: Chelex 100 Resin, Silica based extraction, FTA paper and Phenol–Chloroform based extraction (Goodwin 2007). Differential extraction is a modified extraction procedure which enables the separation of the male and female DNA fraction.

Chelex method of extraction is more rapid than other methods of DNA extraction. Chelex 100 resin is composed of styrene-divinylbenzene copolymers containing paired iminodiacetate ions (Walsh 1991) which sequesters polyvalent ions such as Mg^{2+} and effectively removes them from the solution. Chelex extraction involves no organic solvents and do not require multiple tube transfers for most types of samples (Walsh 1991) thus, fewer opportunities for sample-to-sample contamination. However, it produces single stranded DNA as a result of the extraction process and therefore is only useful for PCR-based testing procedures (Butler 2001).

Organic Extraction involving the use of organic reagents for DNA extraction are well accepted in the forensic science community (NFSTC). Organic extraction methods are often preferred for the extraction of biological stains containing small amounts of DNA or degraded DNA. These methods could be considered less harsh than other methods, such as the use of Chelex beads, because no boiling step is required. Organic extraction is a conventional method that uses organic chemicals to isolate genomic DNA. The procedure can be described in four steps: (1) solubilization of the stain components, (2) denaturation and hydrolysis of proteins, (3) removal of denatured proteins, and (4) purification of DNA (Baechtel 1989).

Organic extraction involves the serial addition of several chemicals. First, sodium dodecyl sulfate (SDS) and proteinaise K are added to break open the cell walls and to break down the proteins that protect the DNA molecules while they are in chromosomes. Next the phenol chloroform mixture is added to separate the proteins from the DNA. The DNA is more soluble in the aqueous portion of the organic aqueous-mixture. When

centrifuged, the unwanted proteins and cellular debris are separated away from the aqueous phase and the double stranded molecules can be cleanly transferred for analysis (Butler 2001).

Another popular method for DNA involves the use of FTA paper, a cellulose-based paper specially treated to bind and nucleic acids and protect them from further degradation. Use of FTA paper simply involves adding a spot of blood to the paper and allowing the stain to dry. The cells are lysed upon contact with the paper and DNA from the white blood cells is immobilized within the matrix of the paper. A small punch of the paper is removed from the FTA card bloodstain and placed into a tube for washing. The bound DNA can then be purified by washing it with a solvent to remove heme and other inhibitors of the PCR reaction .This purification of the paper punch can be seen visually because as the paper is washed, the red color is removed with the supernatant. The clean punch is then added directly to the PCR reaction (Butler 2001).

A major advantage of FTA paper is that consistent results may be obtained without quantification. However, dry paper punches do not like to stay in their assigned tubes and due to static electricity can 'jump' between wells in a sample tray. Thus, this method is not as widely used today as was once envisioned. However, due to its preservation and storage capabilities, efforts have been made to use FTA cards for more widespread collection of crime scene evidence (Lorente *et al.* 2004).

Silica Based Extraction involves a process by which nucleic acids selectively absorb to a silica support, such as small glass beads, in the presence of high concentrations of chaotropic salts such as guanidine hydrochloride, guanidine

isothiocyanate, sodium iodide, and sodium perchlorate (Vogelstein and Gillespie 1979, Boom *et al.* 1990, Duncan *et al.* 2003). These chaotropic salts disrupt hydrogen-bonding networks in liquid water and thereby make denatured proteins and nucleic acids more thermodynamically stable than their correctly folded or structured counterparts (Tereba *et al.* 2004). If the solution pH is less than 7.5, DNA adsorption to the silica is typically around 95% and unwanted impurities can be washed away. Under alkaline conditions and low salt concentrations, the DNA will efficiently elute from the silica material.

PCR Amplification

The quality and quantity of DNA obtained from the extraction process of forensic samples is often inadequate, thus, DNA amplification thru polymerase chain reaction (PCR) is necessary. PCR is an enzymatic process by which a specific segment of DNA is replicated to produce many copies of target DNA in vitro (Reynolds et al 1991). In 1985, Kary Mullis described the PCR process (NFSTC).

Molecular replication process involves consecutive cooling and heating of samples in a specific thermal cycling pattern over 30 times (Butler 2001). In each cycle, a copy of DNA is produced from each template having the target sequence. Oligonucleotide primers determines the boundaries of DNA replication process by flanking on the region of interest (Butler 2001, Roby et. al).

DNA amplification procedures are based on the concept of strand complementarity. Replication process is achieved by repeatedly undergoing a 3-step process in a thermal cycler namely: (1) strand denaturation, (2) primer annealing and (3) primer extension to produce PCR product called amplicon (Roby et.al).

Each cycle results in the doubling of amplicons, thus exponential accumulation of product is achieved. After 30 cycles, approximately a billion copies of the target DNA template are generated. However, further increase in the cycle number decreases the efficacy of replication.

In PCR, more than one region can be amplified simultaneously by adding more than one primer set to the reaction. Such method is called multiplexing (Haas et. al 2003). When done at an optimum condition, this procedure saves time and reagents.

PCR is commonly performed with a sample volume in the range of 5–100 µL. While very low volumes causes evaporation and pipetting errors which leads to concentration issues, high volume affects thermal equilibrium which leads to unoptimized amplification condition. Thus, most PCR amplification protocol uses 20-50µL volume (Butler 2001).

PCR reaction cocktail is composed of several components. These include the following: water. primers, buffer, dinucleotide triphosphates (DNTPs), template.

A primer acts to identify or 'target' the portion of the DNA template to be copied. It is a chemically synthesized oligonucleotide that is added in a high concentration relative to the DNA template to drive the PCR reaction. DNTPs will serve as building blocks for DNA synthesis while DNA polymerase serves as the catalyst to drive the whole amplification process. The most commonly used thermal stable polymerase is *Taq*, which comes from a bacterium named *Thermus aquaticus* that inhabits hot springs (Butler 2001).

PCR has several qualitative characteristics namely; specificity, efficiency (yield of product), sensitivity and fidelity (error rate). Optimization ensures that PCR process is highly specific, high product yield and low error rate.

DNA Analysis

Amplification products are then separated and detected in order to characterize the STR region being examined. Analysis platforms used today include slab gel and capillary electrophoresis (CE). "Fluorescence detection methods have greatly aided the sensitivity and ease of measuring PCR-amplified STR alleles. After detecting the STR alleles, the number of repeats in a DNA sequence is determined, a process known as sample genotyping" (Butler, 2005).

CHAPTER III

MATERIALS AND METHOD

Evaluation of the variability of 14 Y STR markers in the Philippine population

Population profile in the Philippines using 14 Y chromosomal loci will be determined by typing biological samples from unrelated men from all over the Philippines. Information on their ethnicity will gathered to correlate information on the presence of rare alleles. From the analysis, information on repeat range, size range, allelic frequency, haplotype diversity and discrimination power for each locus will be gathered. Also, in line with the use of YSTR analysis for forensic data, haplotype frequencies will be obtained to determine the probability of uniqueness of a particular profile.

Specimen Collection

A total of three hundred fifty (350) biological samples (blood and buccal swabs) from unrelated male DNA samples were provided by the PNP Crime Laboratory.

DNA Extraction

All specimen will be extracted following the existing PNP Crime Laboratory protocol using chelex extraction and kit based extraction protocol..

Chelex extraction

Swab surface will be obtained and placed in 1.5 ml microcentrifuge tube. 200 uL 5% Chelex will be added and sample is placed in Eppendorf thermomixer set at 950 rpm, 56°C (±1°C) for 30 minutes. After which, sample will further be incubated in a thermomixer set at 950 rpm, 100°C for 8 minutes . Sample is then vortexed for 10

seconds and centrifuge at 13000 rpm for 3 minutes. Finally, supernatant is obtained, transferred in a clean sterile tube and stored at -20°C.

Silica kit based extraction

Silica based DNA extraction is performed using Purelink DNA Extraction kit following the manufacturer's manual.

DNA Amplification

Fourteen Y Chromosomal STR markers will be used for amplification: DYS19, DYS385 I/II, DYS389 I/II, DYS390, DYS391, DYS392, DYS 393, DYS 437, DYS 438, DYS 439, DYS 321, Y-A71. 10 µL PCR mix will be prepared as follows: 0.25 mM of primers, 0.2 U/µL Taq polymerase , 0.125 mM of each dNTP, 1.5 mM $MgCl_2$, 1µL 10 X PCR Buffer, 1 µL DNA template (DNA standard sample or 2ng/ml DNA reference standard), 0.06 mg/ml BSA and deionized H_2O. Amplifications will be performed in an Eppendorf or ABI Veriti thermal cycler under the following conditions: initial denaturation step (2 min at 96°C), 30 cycles of successive replication step (1 min at 94 °C, 1 min at 56°C and 1.5 min at 72°C), and final elongation step (10 min at 72 °C and 30 min at 4°C).

DNA Typing

The PCR products will be either be separated using a 5 % polyacrylamide gel and detected using ALF Express or run on ABI 3130xl genetic analyzer with 36 cm array and POP-4 polymer. Fragment sizes will be determined using the Genescan Analysis Software v.3.7 (Applied Biosystems) or Alfwin software. Allele designations will be based on comparison with allelic ladders included in Y filer kit using Genotyper 3.7

software for products run on ABI genetic analyzer and Alleloc software for products run on Alf Express software . Promega 9948 male DNA will be used as control sample.

Analysis of Data

Allele frequencies will be estimated by gene counting. The observed gene diversity (D) and haplotype diversity will be calculated according to Nei. Gene or haplotype diversity will be calculated according to the formula $h= n(1-\sum x_i^2) /(n-1)$, where n represents the number of chromosomes sampled and x_i is the frequency of ith haplotype or allele (Nei, 1987). The discrimination capacity will be estimated as percentage proportion of individual specific haplotypes (Gene et al, 1999). In Y linked systems, haplotype diversity value is identical to the discrimination index (DI) (Sensabaugh, 1982) and to the power of exclusion (POE) (Chakravarti and Li, 1983)

Determination of the spontaneous mutations occurrence between father and son pairs

Mutations will initially be identified by electrophoresis as allele length differences between father and sons. Mutations will be confirmed by reanalysis and DNA sequence analysis. Prior to sequencing reaction, samples carrying mutations will be purified using DNA purification kit. Mutation rates will be estimated as the number of mutations divided by the number of allele transmissions. Characteristics of mutations will be compared with observations from previous researches.

Specimen Collection

Biological samples (blood and buccal swabs) of 150 father and son pairs previously confirmed using autosomal STR analysis will be provided by the PNP Crime Laboratory.

DNA Extraction

All specimen used for mutational analysis will be extracted based on the existing PNP Crime Laboratory protocol using Purelink DNA Extraction kit following the manufacturer's manual.

DNA Amplification

Specimen from father and son pairs will be amplified using the same set of Y Chromosomal STR markers under the same condition as that of population study.

DNA Typing

The PCR products for mutational analysis will be typed following the same set of procedure for population analysis.

Analysis of Data

Mutation rates per locus will be estimated as the number of mutations divided by the number of allele transmissions. Combined mutation rates will also be determined as number of mutations observed divided by the total number meioses. Comparison of mutation rates with previous studies will be done using exact tests.

Development of cost effective protocols for routine DNA examination

In developing cost effective protocol for routine examination, particularly the markers suitable for chelex extraction – conventional PCR amplification – gel based

detection, 30 specimen from the population study subjected to chelex extraction and 30 specimen from the population study subjected to kit based extraction will be amplified using the protocol described above. PCR products will be separated using a 5% polyacrylamide gel and detected using ALF Express. Success rate in typing per locus will evaluated in order to determine the set of primers suitable for low cost DNA typing.

List of References

Basic Biology of DNA [Internet]. US DNA Initiative; c2009 [cited February 3, 2010] Available from: http//www.dna.gov/basics/analysis/history/

Baechtel, F. S. 1989. The extraction, purification and quantification of DNA. In *Proceedings of the International Symposium on the Forensic Aspects of DNA Analysis*, 25–28. Washington, D.C.: U.S. Government Printing Office.

Burgoyne, L.A. (1996) Solid medium and method for DNA storage. U.S. Patent 5 496 562.

Budowle B. 2000. History and future of DNA typing. Genetic Identity Conference Proceedings. Eleventh International Symposium on Human Identification. [cited February 28, 2010] available from: http://www.promega.com/geneticidproc/ussymp11proc/content/budowle2.pdf

Butler JM, Tomsey CS, Kline, MC. 2004. Can validation process in forensic DNA typing be standardized?. In: 15th International symposium on Human Identification; October 6, 2004. Available from: <http.www.promega.com/GENETICIDPROC/US symp15proc/oralpresentations/butler.pdf> Accessed February 3, 2010.

Catalin M, Andrei A, Mitrasca, O. Modern methods of collection and preservation of biological evidence for human identification by DNA analysis. University of Medicine and Pharmacy from Timiasoara. [cited February 28, 2010] Available from: http://www.oglethorpe.edu/faculty/k_aufderheide/forensic_science/web_documents/Catalin-Andrei-Mitrasca.pdf

Clark L. Organic extraction. National Science Forensic training Center (NFSTC); c2010 [cited March 9, 2010]. Available from: http://www.nfstc.org/pdi/Subject03/pdi_s03_m02.htm

Corach, D., Risso L.F., Marino, M., Penacino, G., SAla, A. 2001.Routine Y STR typing in forensic casework. Forensic Science International 118: 131-135.

Derksen LA. 2003. Agency and Structure in the History of DNA Profiling: The Stabilization and Standardization of a New Technology[dissertation]. San Diego: University of California. 342p. [cited March 20, 2010] Available from: http://web.viu.ca/derksenl/Publications/DNA%20Profiling%20History.htm

DNA Initiative. May 2001. *Understanding DNA Evidence: A Guide for Victim Serice Providers* : US Departmet of Justice, National Institute of Justice and Office for Victims of Crime [cited March 9, 2010]. Available from: http://www.dna.gov/audiences/victim/know/

DNA Profiling. [Internet] Interpol; c2009. [cited March 10, 2010]. Available from: http://www.interpol.int/Public/Forensic/dna/dnafaq.asp

Domingues, P.M., Gusmao, L., Da Silva, D.A., Amorim, A., Pereira, R.W. Carvalho, E.F. 2007. Sub-Saharan Africa descendants in Rio de Jeneiro (Brazil): pppulation and mutational data for 12 Y-STR loci. 2007. International Journal Journal of Legal Medicine. 121:238-241.

[EMFPU] East Midlands Forensic Pathology Unit. History of DNA Profiling. University of Leicester; c2009 [cited March 1, 2010] Available from http://www2.le.ac.uk/departments/emfpu/genetics/explained/profiling-history

Edwards A, Citivello A, Hammond, HA, Caskey CT. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet 49:746-756.

Federal Bureau of Investigation. Quality assurance standards for forensic DNA testing laboratories. Forensic Science Communication [Online]. Available from: <http:www.fbi.gov/hq/lab/fsc/backissu/july2000/codispre.htm Accessed February 10, 2010.

Figarelli D. DNA Amplification. National Science Forensic training Center (NFSTC); c2010 [cited March 9, 2010]. Available from: http://www.nfstc.org/pdi/Subject04/pdi_s03_m02.htm

Goedbloed, M., Vermeulen, M., Fang R. N., Lembring M.Wolstein A. 2009. Comprehensive Mutation Analysis of 17 Y chromosomal short tanden repeats polymorphisms included in the AmpFLSTR Y Filer PCR Amplification Kit. International Journal Legal Medicine 123:471-482.

Goodwin W, Linacre A, Had S. 2007. An Introduction to Forensic Genetics. John Wiley and Sons inc. West Sussex England pp. 30-45

Goes, A.C., Carvalho, E.F., Gomes, I, Da Silva, D.A. Gil, E.H., Amorim, A, GUsmao, L. Population and Mutation Analysis of 17 Y STR loci from Rio de Janeiro (Brazil). 2005. International Journal of Legal Medicine, 119:70-76.

Haas, S. A., M. Hild, A. P. Wright, T. Hain, D. Talibi, and M. M. Vingron. 2003. Genome-scale design of PCR primers and long oligomers for DNA microarrays. *Nucleic Acids Res* 31 (19): 5576–81

Human Genome Project Information: DNA Forensics [internet]. US Department of Energy Office of Science, Office of Biological and Environmental Research, Human

Genome Program; 2009 [cited March 1, 2010]. Available from: http//www.ornl.gov/sci/techresources/Human_Genome/home.shtml

Jeffreys AJ, Wilson V, Thein SL. 1985. Individual-specific 'fingerprints' of human DNA. Nature. 316(6023):76-9

Klobuchar A, Caliguiri HL. 2005. Protecting the innocent/convicting the guilty: Hennepin county's pilot project in blind sequential eyewitness identification. Law Review 32(1):1-25. [cited March 12, 2010] http://www.wmitchell.edu/lawreview/Volume32/Issue1/1Klobuchar.pdf

Lee, H.W., Park, M.J., Chung U., Lee, H.Y., Yang W.I., Cho S. Shin, K.J. 2007. Haplotypes and mutational Analysis of 22 Y-Chromosomal STRs in Kerean father-som pairs. International Journal of Legal Medicine 121: 128-135.

Leriche A, Vanek D, Schmitter H, Schleenbecker U, Woller J, Montagna P, Garofano L, Sprangers W, Wolf E, Moe N, Matusek J, Heranz JAD, Rojo-Gambin AMG, Utrilla L, Grahamslaw W, Fifka P, Branchflower M. 1998. Final Report of the Interpol European Working Party on DNA Profiling. Genetic Identity Conference Proceedings. Second European Symposium on Human Identification

[NCFDE] National Commission on the Future of DNA Evidence. 2000. The future of forensic DNA testing: predictions of the research and development working group: US Department of Justice, National Institute of Justice. NCJ 183697.

[NIJ] National Institute of Justice. July 2002. Using DNA to solve cold cases. National Institute of Justice Report. US Department of Justice. NCJ 194197. [cited March 1, 2010]. Available from: http://www.ncjrs.gov/pdffiles1/nij/194197.pdf

Nicklas J, Buel E. 2003. Quantification of DNA in forensic samples. Analytical and Bioanalytical Bhemistry. 376: 1160. [cited March 12, 2010] Available from: http://cat.inist.fr/?aModele=afficheN&cpsidt=15064380

President's DNA Initiative. April 2005. Identifying victims using DNA: a guide for families: US Department of Justice, National Institute of Justice . NCJ 209493.

Roby R, Figarelli D. The extraction, purification and quantification of DNA. STR Analysis. National Science Forensic training Center (NFSTC). c2010 [cited March 9, 2010]. Available from: http://www.nfstc.org/pdi/Subject04/pdi_s03_m05.htm

Rodig, H., Grum, M., Grimmecke, H.D. 2007. Population Study and evaluation of 20 Y chromosome STR loci in Germans. International journal of Legal Medicine. 121:24-27.

Saad R. 2004.Discovery, development, and current applications of DNA identity testing. In: Baylor University Medical Center; November 23, 2004. p 1-12.

Tamaki K, Jeffreys, A J. 2005. Human tandem repeats in forensic DNA typing. Legal Medicine. 7: 244-250.

Wyman A, White R.1980. A highly polymorphic locus in human DNA. Proc. Nati. Acad. Sci. USA. 77(11): 6754-6758 [cited March 12, 2010] Available from http://www.pnas.org/content/77/11/6754.full.pdf

Promega. 2009. DNA-Based Human Identification: Protocols and Applications guide. Promega Corporation. [cited March 9, 2010]. Available from: http://www.promega.com/paguide/chap14.htm#title2

Viera, J., Billerbeck, A., Iwamura, E., Mendonca, B., Gusmao, L. Otto, P. 2008. Genetics and Molecular Biology. 31 (3): 651-656.

Walsh PS, Metzger DA, Higuchi R. 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques. 10(4):506-13