저작자표시-비영리-변경금지 2.0 대한민국

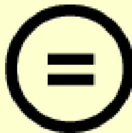
이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

Dissertation for the Degree of Doctor

Development of Reference Databases for Short Tandem Repeats (STRs) and mitochondrial HV1 region in Dogs

Supervisor: Seongbeom Cho, D.V.M., Ph.D.

Submitted by

Seohyun Moon

December, 2016

Department of Veterinary Public Health

Graduate School of Seoul National University

한국 내 개의 STRs (Short Tandem Repeats) 및 mitochondrial HV1 지역 분석을 통한 집단 유전학적 데이터베이스 구축에 대한 연구

지도교수 조 성 범

이 논문을 수의학 박사학위 논문으로 제출함
2016년 12월

서울대학교 대학원
수의학과 수의공중보건학 전공

문 서 현
문서현의 박사학위 논문을 인준함

2016년 12월

위 원 장	____________________	(인)
부 위 원 장	____________________	(인)
위 원	____________________	(인)
위 원	____________________	(인)
위 원	____________________	(인)

ABSTRACTS

Development of Reference Databases for Short Tandem Repeats (STRs) and mitochondrial HV1 region in Dogs

(Supervisor: Seongbeom Cho, D.V.M., Ph.D.)

Seohyun Moon

Department of Veterinary Public Health

Graduate School of Seoul National University

Dogs have intimate and close relationships with human for long time. Due to the large number of dogs in human, they are often involved in crimes. Occasionally, canine biological evidence such as saliva, bloodstains and hairs is found in crime scenes. Accordingly, canine DNA can be used as forensic evidence. The use of short tandem repeats (STRs) loci from biological evidence is valuable for forensic investigations. In Korea, canine STR

profiling-related crimes are being successfully analyzed, leading to solve crimes such as animal cruelty, dog-attacks, murder, robbery, and missing or abandoned dogs. However, the probability of random DNA profile matches cannot be analyzed because of a lack of canine STR database. Therefore, in this study, 10 STR loci were analyzed in 600 dogs in Korea (344 dogs belonging to30 different purebreds and 256 crossbred dogs) to estimate canine forensic genetic parameters. Among purebred dogs, a separate statistical analysis was conducted for five major subgroups, 97 Malteses, 47 Poodles, 31 Shih-tzus, 32 Yorkshire terriers, and 25 Pomeranians. Allele frequencies, expected (Hexp) and observed heterozygosity (Hobs), fixation index (F), probability of identity (P(ID)), probability of sibling identity (P(ID)sib) and probability of exclusion (PE) were calculated. The Hexp values ranged from 0.901 (PEZ12) to 0.634 (FHC2079), while the P(ID)sib values were between 0.481 (FHC2079) and 0.304 (PEZ12) and the P(ID)sib was about 3.35×10^{-5} for the combination of all 10 loci. In forensics, profiling of short tandem repeats (STRs) is an ideal tool to identify individuals. However, this evidence, like single, shed hairs from dogs contains only limited amounts of DNA. The complete STR profiles cannot be obtained from this degraded and limited DNA. So, high copy number of mitochondrial genomes per cell is popular for forensic analysis, especially, when samples

contain small amount, degraded or limited DNA. The canine mitochondrial DNA length of sequence is 16727 base-pairs. Like human, canine mitochondrial DNA (mtDNA) contains hypervariable regions (HVRs): hypervariable region 1 (HV1) and hypervariable region 2 (HV2). The HV1 of canine mitochondrial DNA (mtDNA) is highly polymorphic. Thus it is possible that the severely degraded DNA sample can be amplified. To build a canine population database of domestic dogs in Korea, the 612bp hypervariable region 1 (HV1) sequences from 158 dogs (*Canis lupus familiaris*) in Korea were analyzed. In this study, 23 haplotypes from 35 single nucleotide polymorphisms (SNPs) were identified. The four most common HV1 haplotypes (n = 112 dogs) represented 70.9% of the total samples. Haplotype frequency is consistent with previous studies and exclusion capacity of mtDNA population is 0.851. This study is to create the first reference database of canine STR and mtDNA in Korea. Consequentially, this STR and mtDNA population data would improve the genetic evidential power of canine crimes in forensic caseworks by two ways.

Keywords: STRs (Short Tandem Repeats), allele frequency, individual identification, canine mitochondrial DNA, hypervariable region 1 (HV1), haplotypes, single nucleotide polymorphisms (SNPs), exclusion capacity

Student number: 2012-30451

TABLE OF CONTENTS

ABSTRACTS III
TABLE OF CONTENTS VII
LIST OF TABLES X
LIST OF FIGURE XI
LIST OF ABBREVIATIONS XII

General Introduction 13

Chapter I. Literature Review 17

1.1. DNA profiling 18
1.2. Short Tandem Repeats (STRs) analysis 19
1.2.1. Genetic Markers and Repeated DNA Sequences 19
1.2.2. STRs Used in Forensic DNA Typing 20
1.3. Mitochondrial DNA Analysis 22
1.4. DNA database 25
1.4.1. Values of database 25
1.4.2. National DNA Databases around the World 26
1.4.3. mtDNA population databases 28

1.4.4. FBI mtDNA Database 29

1.5. Non-human DNA typing 31

1.5.1. Domestic Animal DNA Testing 31

1.5.2. Canine mtDNA analysis 34

1.5.3. Other Domesticated Animals. 35

Chapter II. Development of a Reference Database for Short Tandem Repeats (STRs) in dogs 37

2.1. Abstract 38

2.2. Introduction 40

2.3. Materials and Methods 43

2.3.1. Population 43

2.3.2. DNA Extraction 43

2.3.3. PCR Amplification and Quality Control 44

2.3.4. Typing and Analysis of Data 45

2.4. Results 47

2.5. Discussion 49

Chapter III. Development of a Reference Database for mitochondrial HV1 region in dogs 73

3.1. Abstract 74
3.2. Introduction 76
3.3. Materials and Methods 79
3.3.1. DNA samples 79
3.3.2. PCR amplification and Sequencing.. 79
3.3.3. Data analysis 80
3.4. Results 82
3.4.1. HV1 haplotypes 82
3.4.2. Statistical data analysis 83
3.5. Discussion 85
General Conclusions 105
References.. 107
국문 논문 초록.. 124

LIST OF TABLES

Table 2-1. Allele frequency and forensically informative statistical values of the 10 analyzed STR loci in the "All breeds" samples. 52
Table 2-2. Allele frequency and forensically informative statistical values of the 10 STR analyzed loci in the Maltese, Poodle, Shih tzu, Yorkshire terrier, Pomeranian and the "All breeds" samples. 57
Table 2-3. *F*-statistical (F_{ST}, F_{IS}, F_{IT}) values for each analyzed STR locus among five Korean dog populations. 71
Table 3-1. List of 158 samples from 23 breeds and mongrels. 90
Table 3-2. Sequence variation, haplotypes and breeds of 158 Korean domestic dogs. 91
Table 3-3. Canine mtDNA HV1 haplotypes distribution of 158 dogs analyzed in this study. 99
Table 3-4. Genetic distance for the frequencies of 158 dogs population (values of population pairwise Fst). 102
Table 3-5. AMOVA design and result of 158 Korean domestic dogs 103

LIST OF FIGURE

Figure 3-1. Pie chart illustrating the distribution of the domestic dog HV1 haplotypes observed in this study data set (n = 158). 101

LIST OF ABBREVIATIONS

STRs	Short tandem repeats
SNPs	Single nucleotide polymorphisms
VNTRs	Variable number of tandem repeats
PCR	Polymerase chain reaction
RFLP	Restriction fragment length polymorphism
LCN	Low copy number
RFU	Relative fluorescence unit
HWE	Hardy-Weinberg equilibrium
Hobs	Observed heterozygosity
Hexp	Expected heterozygosity
PIC	Polymorphic information content
PD	Power of discrimination
RMP	Random match probability
PE	Power of exclusion
mtDNA	Mitochondrial DNA
HVRs	Hypervariable regions
AMOVA	Analysis molecular variance
Fst	Fixation index
Ne	Number of effective allele
Na	Number of different allele
PI	Probability of identity
PI(sib)	Probability of identity (siblings)
I	Shannon's information index
uHe	Unbiased expected heterozygosity

General Introduction

DNA is present in every nucleated cell and is therefore present in biological materials left at crime scenes. DNA has been successfully isolated and analyzed from a variety of biological materials such as blood, saliva, semen, body fluids, tissue and so on. In forensics, it is possible to identify individuals by characteristics of their DNA from biological materials in crime scenes. It is called DNA profiling, DNA fingerprinting, DNA testing, or DNA typing. A DNA profile is a small set of DNA variations that is very likely to be different in all unrelated individuals, thereby being as unique to individuals as are fingerprints. Short Tandem Repeats (STRs) analysis is a useful tool in forensic identification (Lygo et al. 1994, Butler 2006, Ganco et al 2009). For human identification purposes, it is important to have DNA markers that exhibit the highest variation to discriminate between samples. DNA regions with repeat units that are two base-pairs to seven base-pairs in length are called microsatellites, or most usually short tandem repeats (STRs). Unrelated people almost certainly have different numbers of repeat units. STRs can be used to discriminate between unrelated individuals. The STRs have become popular DNA repeat

markers because they are easily amplified by the polymerase chain reaction (PCR) without the problems of differential amplification. This is because both alleles from a heterozygous individual are similar in size since the repeat size is small (Butler, 2005). Among the various types of STR systems, tetra-nucleotide repeats have become more popular than di- or tri-nucleotides. These STR loci are targeted with sequence-specific primers and amplified using PCR. The DNA fragments that result are then separated and detected using capillary electrophoresis (CE). The power of STR analysis comes from looking at multiple STR loci simultaneously. The pattern of alleles can identify an individual quite accurately. Thus STR analysis provides an excellent identification tool. The more STR regions that are tested in an individual the more discriminating the test becomes.

The other hand, for highly degraded samples, it is sometimes impossible to get a complete profile of full STRs. In these situations, mitochondrial DNA (mtDNA) is sometimes typed due to there being many copies of mtDNA in a cell, while there may only be 1-2 copies of the nuclear DNA. The mtDNA can be obtained from such material as hair shafts, old bones and teeth. There are two hypervariable regions in mtDNA control region; HV1 and HV2 regions of the mtDNA are

amplified and then sequence each region and compare single-nucleotide differences to a reference. Because mtDNA is maternally inherited, directly linked maternal relatives can be used as match references. In general, a difference of two or more nucleotides is considered to be an exclusion purpose. The DNA profiling of STRs and mtDNA sequencing is also commonly applied to canine DNA studies. Todays, various researches of canine STRs and mtDNA typing database are being constructed worldwide.

As to DNA database of worldwide, different STR-based DNA-profiling systems are in use. In United States, systems that amplify the human CODIS 13 core loci are almost universal. The true power of STR analysis is in its statistical power of discrimination. Because the 13 loci that are currently used for discrimination in CODIS are independently assorted, the product rule for probabilities can be applied. This has resulted in the ability to generate match probabilities of 1 in a quintillion ($1x10^{18}$) or more.

In the first study, we reviewed literature about fundamentals of forensic DNA typing such as DNA profiling, short tandem repeats (STRs), mitochondrial DNA analysis, worldwide DNA database and non-human DNA typing. It could be a basis for understanding DNA

profiling and needs for researching canine DNA typing.

In the second study, the first database of canine STR in Korea was made by 600 dogs STR typing. It could improve the reliability of canine STR analysis by determining the allele frequency and forensic informative values from 600 purebred and crossbred dogs. The results revealed that 10 canine STR markers could be suitable for individual identification in forensic cases and enable estimation of the matching probability and discrimination of forensic work.

And in the third chapter, the mtDNA database of Korean canine dogs was represented. The mtDNA analysis is conducted when the complete STR profiles cannot be obtained from this degraded and limited DNA. In forensic analysis, even single dog hair could be valuable evidence that links victims and suspects. In this study, to build a canine population database of domestic dogs in Korea, the 612-bp hypervariable region 1 (HV1) sequences of from 158 dogs (*Canis lupus familiaris*) were analyzed. The obtained haplotypes were consistent with previous population studies in other countries (Angelby et al. 2005, Webb et al. 2009, and Desmyter et al. 2009). Also, the exclusion capacity of canine mtDNA in the Korean population represented enough informative in forensic analysis.

Chapter I.

Literature Review

1.1. DNA profiling

DNA profiling (also called DNA fingerprinting, DNA testing, or DNA typing) is a forensic technique used to identify individuals by characteristics of their DNA. A DNA profile is different in all unrelated individuals, thereby being as unique to individuals as are fingerprints. It is also called DNA fingerprinting or DNA typing. It is first developed and used in 1985 and modern process of DNA profiling was developed in 1988 by Alec Jeffrey. DNA profiling is used in identifying individuals in parentage testing and criminal investigation. DNA fingerprinting has also been widely used in the study of animal populations and has revolutionized the field of zoology. Although 99.7 % of human DNA sequences are the same in every person, enough of the DNA is different that it is possible to distinguish one individual from another except identical twins. DNA profiling uses repeat sequences that are highly variable in particular short tandem repeats (STRs). The process begins with a sample of an individual's DNA. A common method of collecting a sample is the use of a buccal swab, which is easy, non-invasive and cheap. In forensic caseworks, samples usually collected of blood, saliva, semen, or other biological fluid or tissue from personal items (e.g. a toothbrush, razor). A sample is then

analyzed to create the individual's DNA profile using various steps of techniques. The DNA profile is then compared against another sample to determine whether there is a genetic match.

1.2. Short Tandem Repeats (STRs) analysis

1.2.1. Genetic Markers and Repeated DNA Sequences

Since it has been estimated that over 99.7% of the human genome is the same from individual to individual, regions that differ need to be found in the remaining 0.3% in order to tell people apart at the genetic level. Eukaryotic genomes are full of repeated DNA sequences (Ellegren 2004). These repeated DNA sequences come in all sizes and are typically designated by the length of the core repeat unit and the number of contiguous repeat units or the overall length of the repeat region. Long repeat units may contain several hundred to several thousand bases in the core repeat. The small size of STR alleles (100 base-pairs to 400 base-pares) compared to mini-satellite VNTR alleles (400 base-pairs to 1000 base-pairs) make the STR markers better

candidates for use in forensic applications where degraded DNA is common. PCR amplification of degraded DNA samples can be better accomplished with smaller product sizes. A biological phenomenon known as "stutter" results when STR alleles are PCR amplified. Stutter products are amplicons that are typically one or more repeat units less in size than the true allele and arise during PCR because of strand slippage (Walsh et al. 1996). Stutter product amounts vary depending on the STR locus and even the length of the allele within the locus but are usually less than 15% of the allele product quantity with tetra-nucleotide repeats. With di- and tri-nucleotides, the stutter percentage can be much greater (30% or more) making it difficult to interpret sample mixtures. In addition, the four-base spread in alleles with tetra-nucleotides makes closely spaced heterozygotes easier to resolve with size-based electrophoretic separations compared to alleles that could be two or three bases different in size with di-nucleotide and tri-nucleotide markers, respectively.

1.2.2. STRs Used in Forensic DNA Typing

The system of DNA profiling used today is based on short tandem repeats (STRs). This method uses highly polymorphic regions that have short repeated sequences of DNA. The most common is four bases repeated, but there are other lengths in use, including three and five bases. STRs have become popular DNA repeat markers because they are easily amplified by the polymerase chain reaction (PCR) without the problems of differential amplification. This is because both alleles from a heterozygous individual are similar in size since the repeat size is small. The number of repeats in STR markers can be highly variable among individuals, which makes these STRs effective for human identification purposes. For human identification purposes it is important to have DNA markers that exhibit the highest possible variation or a number of less polymorphic markers that can be combined in order to obtain the ability to discriminate between samples. Unrelated people almost certainly have different numbers of repeat units. STRs can be used to discriminate between unrelated individuals. Forensic specimens are often challenging to PCR amplify because the DNA in the samples may be severely degraded. Mixtures are prevalent as well in some forensic samples, such as those obtained from sexual assault cases containing biological material from both the perpetrator

and victim. Each STR is polymorphic, but the number of alleles is very small. Typically each STR allele will be shared by around 5 - 20% of individuals. The power of STR analysis comes from looking at multiple STR loci simultaneously. The pattern of alleles can identify an individual quite accurately. Thus STR analysis provides an excellent identification tool. The more STR regions that are tested in an individual the more discriminating the test becomes.

1.3. Mitochondrial DNA Analysis

Conventional STR typing systems do not work in every instance. Ancient DNA specimens or samples that have been highly degraded often fail to produce results with nuclear DNA typing systems. However, recovery of DNA information from damaged DNA is sometimes possible with mitochondrial DNA (mtDNA). While a nuclear DNA test is usually more valuable, and mtDNA result is better than no result at all. Because there are hundreds if not thousands of copies of mtDNA in each cell, the probability of obtaining a DNA typing result from mtDNA is higher than that of polymorphic markers found in nuclear DNA, particularly in cases where the amount of

extracted DNA is very small, as in tissues such as bone, teeth, and hair. When remains are quite old or badly degraded, often bone, teeth, and hair are the only biological sources left from which to draw a sample. In short, though nuclear DNA contains much more information, there are only two copies of it in each cell (one maternal and one paternal) while mtDNA provides a bit of useful genetic information hundreds of times per cell. Because of their higher numbers, some mtDNA molecules are more likely to survive than nuclear DNA. The vast majority of the human genome is located within the nucleus of each cell. However, there is a small, circular genome found within the mitochondria, the energy-producing cellular organelle residing in the cytoplasm. The number of mtDNA molecules within a cell can range from hundreds to thousands. On average there are 4 to 5 copies of mtDNA molecules per mitochondrion with a measured range of 1 to 15 (Satoh & Kuroiwa 1991). Because each cell can contain hundreds of mitochondria (Robin & Wong 1988), there can be up to several thousand mtDNA molecules in each cell. Mitochondrial DNA has approximately 16,569 base pairs with the total number of nucleotides in a specific mtDNA genome (mtGenome) varying due to small insertions or deletions. Most of the mtGenome codes for 37 gene products used in the oxidative

phosphorylation process or cellular energy production. The 37 transcribed "genes" of mtDNA found in the "coding region" include 13 proteins, 2 ribosomal RNAs (rRNA), and 22 transfer RNAs (tRNA). There is also a 1122bp "control" region that contains the origin of replication for one of the mtDNA strands but does not code for any gene products and is therefore referred to sometimes as the "non-coding" region. Most of the focus in forensic DNA studies to date has involved two hypervariable regions within the control region commonly referred to as HVI (HV1) and HVII (HV2). Occasionally a third portion of the control region, known as HV3, is examined to provide more information regarding a tested sample. Human mitochondrial DNA is considered to be inherited strictly from our mothers. Mitochondria with their mtDNA molecules are passed directly to all offspring independent of any male influence. Thus, barring mutation, a mother passes along her mtDNA type to her children, and therefore siblings and maternal relatives have an identical mtDNA sequence. Mitochondrial DNA analysis is commonly performed using the Sanger sequencing chemistry. This DNA sequencing is performed in both the forward and reverse directions so that the complementary strands can be compared to one another for quality control purposes. Mitochondrial DNA analysis

typically involves materials, where little DNA is present to begin with. Teeth, hair, and bones such as ribs and long bones (e.g., femur and humerus) are often materials used for mtDNA analysis in forensic cases. The mtDNA must be carefully extracted from these materials and often purified away from polymerase chain reaction (PCR) inhibitors that can be coextracted.

1.4. DNA database

1.4.1. Values of database

DNA databases in many cases enable successful conclusion to forensic cases without initial suspects and connection of serial crimes involving biological evidence. Two primary indices exist with forensic DNA databases that are searched against one another: (1) DNA profiles from offenders who have been convicted or in some cases arrested for a crime, and (2) DNA profiles from crime scene evidence. There are now several DNA databases in existence around the world. Some are private, but most of the largest databases are government controlled. The United States maintains the largest DNA database, with the Combined

DNA Index System (CODIS) holding over 5 million records as of 2007. The United Kingdom maintains the National DNA Database (NDNAD), which is of similar size, despite the UK's smaller population. The size of this database, and its rate of growth, is giving concern to civil liberties groups in the UK, where police have wide-ranging powers to take samples and retain them even in the event of acquittal. When a match is made from a National DNA Databank to link a crime scene to an offender having provided a DNA Sample to a databank that link is often referred to as a cold hit. A cold hit is of value in referring the police agency to a specific suspect but is of less evidential value than a DNA match made from outside the DNA Databank.

1.4.2. National DNA Databases around the World

Combined DNA Index System (CODIS) is the United States national DNA database created and maintained by the Federal Bureau of Investigation. CODIS consists of three levels of information; Local DNA Index Systems (LDIS) where DNA profiles originate, State DNA Index Systems (SDIS) which allows for laboratories within states to share information, and the National DNA Index System (NDIS) which

allows states to compare DNA information with one another. The CODIS software contains multiple different databases depending on the type of information being searched against. Examples of these databases include, missing persons, convicted offenders, and forensic samples collected from crime scenes. Each state, and the federal system, has different laws for collection, upload, and analysis of information contained within their database. However, for privacy reasons, the CODIS database does not contain any personal identifying information, such as the name associated with the DNA profile. The UK National DNA Database (NDNAD) was launched by the United Kingdom's Home Office (Werrett, 1997). This database originally stored data from only six STR loci from the Second Generation Multiplex (SGM) consisting of FGA, TH01, VWA, D8S1179, D18S51, and D21S11. In 1999, an expansion was made to 10 STR loci (the six SGM loci plus D3S1358, D16S539, D2S1338, and D19S433) using the SGM Plus kit from Applied Biosystems. From 1995 to 2000, more than 500,000 DNA profiles were entered into the database and more than 50,000 criminal investigations were aided. As of 2010, the NDNAD contains more than four million profiles and regularly aids UK law enforcement person resolving thousands of crimes each year. With around 50

million people in the UK, their DNA database of greater than four million profiles represents the highest proportion of its population.

1.4.3. mtDNA population databases

Population databases play an important role in estimating the expected frequency of mtDNA haplotypes that are observed in casework when a suspect's mtDNA sequence matches that of an evidentiary sample. A great deal of effort has been expended to gather information from thousands of maternally unrelated individuals in various population groups around the world. Having high-quality information in the database is also important in order to make a reliable estimate of the frequency for a random match. The mtDNA typing results on samples from unknown sources are most useful if they are evaluated in comparison to a known sample or a database. Databases of more than 1000unrelated individuals now exist and have been compiled from multiple population groups (Budowle et al. 1999, Röhl et al.2001, Monson et al. 2002). The size of the database is important because without recombination between mtDNA molecules, an mtDNA sequence is treated as a single locus (i.e., haplotype instead of

genotype). The largest compiled database described to date contains HV1 and HV2 sequences from 14,138 individuals (Röhl et al. 2001). This information was collated from 103 mtDNA publications prior to January 2000, 13 data sets published in 2000 and 2001, and two unpublished data sets.

1.4.4. FBI mtDNA Database

The FBI has compiled the mtDNA Population Database also known as CODISmt database (Monson et al. 2002) for the purpose of being able to determine a legally defensible frequency estimate. The CODISmt database has a forensic and a published literature component to it (Miller & Budowle 2001) in order to separate data obtained from laboratories following validated forensic protocols and academic research laboratories where data quality is not reviewed as carefully prior to publication. The forensic database contains 4839 mtDNA profiles from 14 different populations. These samples have been sequenced and the electropherograms carefully reviewed across positions 16024 to 16365 for HV1 and positions 73 to 340 for HV2. An additional 6106 published profiles have been compiled from the

literature with annotated population information (Miller et al. 1996, Miller & Budowle 2001). For classification of mtDNA profiles, a standard 14-character nucleotide sequence identifier was assigned to each profile where the first three characters represent the country of origin, the second three characters the group or ethnic affiliation, and the final six characters are sequential acquisition nnumbers (Miller & Budowle 2001, Monson et al. 2002). Both of these databases were publicly released in April 2002 in a Microsoft Access format and can be downloaded from the FBI website along with the “MitoSearch” analysis tool (Monson et al. 2002). MitoSearch can examine the population data sets for specific mtDNA sequences, which are entered based on differences from the Cambridge Reference Sequence. The software returns the number of times that the specified profile appears in each population group.

1.5. Non-human DNA typing

1.5.1. Domestic Animal DNA Testing

Budowle et al. (2005) note that genetic analysis with animal DNA samples can help resolve criminal and civil cases as well as aiding kinship analysis with applications such as determining the sire of an offspring when a female has been exposed to multiple males. The American Pet Products Association reported in their 2007 and 2008 national pet owners' survey that over 71 million U.S. households own a pet. Their survey found 88 million cats and 75 million dogs in these households, which make up almost two-thirds of all U.S. residences. Since many of these domestic animals shed hair, these hairs could be picked up or left behind at the scene of a crime by a perpetrator. An assailant may unknowingly carry clinging cat hairs from a victim's cat away from the scene of a crime, or hair from the perpetrator's cat may be left at the scene. The Veterinary Genetics Laboratory (VGL) at the University of California-Davis has been performing forensic animal DNA analyses since 1996. The VGL website notes that there are three types of animal DNA evidence: (1) the animal as victim, (2) the animal

as perpetrator, and (3) the animal as witness. Animal abuse cases or the theft of an animal can sometimes be benefited by the power of DNA testing. The remains of a lost pet can be positively identified through genetic analysis. Typically genetic markers like short tandem repeats (STRs) and mitochondrial DNA (mtDNA) are examined in much the same way as with human DNA. When animals are involved in an attack on a person, DNA typing may be used to identify the animal perpetrator (e.g., a Pit Bull). If the victim is deceased, then DNA evidence may be the only witness that an animal in custody committed the crime. Animal DNA testing can "exonerate" innocent animals so that they are not needlessly destroyed. Animal DNA has been used successfully to link suspects to crime scenes. A study on the transfer of animal hair during simulated criminal behavior found that hundreds of cat hairs or dog hairs could be transferred from the homes of victims to a burglar or an aggressor (D'Andrea et al. 1998). In fact, the number of hairs found was so high that the authors of this study felt that it is almost impossible to enter a house where a domestic animal lives without being "contaminated" by cat and/or dog hairs even when the owner describes his or her animal as a poor source of hair (D'Andrea et al. 1998). Due to the fact that shed hairs often do not contain roots, nuclear

DNA may not be present in sufficient quantities for STR typing. Mitochondrial DNA may be a more viable alternative for many of these types of shed hair transfers. Dog DNA may be involved in situations where the animal hair acts as a silent witness to connecting a perpetrator to a crime scene, evidence from dogs is more frequently linked to situations where the animal is the perpetrator. Rottweilers, German Shepherds, Doberman Pinschers, and Pit Bulls can be trained as security animals and may attack, injure, or even kill people. For example, with a canine population in Australia of around 4 million, there are an estimated 100,000 dog attacks each year and many of them go unsolved (Clarke & Vandenberg 2010). Early dog STR assays included many di-nucleotide repeat loci but tetra-nucleotide loci with lower stutter have been the focus of more recent efforts. However, different groups have targeted different sets of loci with almost no overlap between them (Halverson & Basten 2005, Berger et al. 2008, van Asch et al. 2009, Tom et al. 2010). There is a need for standardization on the loci and allele nomenclatures used (Berger et al. 2009). In an early effort, 15 canine STR loci were characterized with sequenced alleles (Eichmann et al. 2004). The PEZ locus names come from Perklin-Elmer Zoogen, a company that developed the StockMark

kits for Applied Biosystems back in the mid-to-late 1990s. The FH locus names came from the Fred Hutchinson Cancer Research Center (Seattle, Washington). A set of 10 dinucleotide repeat STRs has been used to aid investigations in illegal animal deaths (Padar et al. 2001) and a dog attack that resulted in the death of a seven-year old boy (Padar et al. 2002). In addition, it was demonstrated that DNA profiling of human blood recovered from a dog's fur can associate or exonerate the animal from connection to an attack (Brauner et al. 2001).

1.5.2. Canine mtDNA analysis

Canine mitochondrial DNA possesses two hypervariable regions (HV1 and HV2) similar to the human mtDNA. Savolainen et al. (1997) found 19 sequence variants across a 257 basepairs segment of the hypervariable region 1 of the mtDNA control region in 102 domestic dogs of 52 different breeds. They concluded that on average 88 out of 100 tested animals could be excluded with this mtDNA sequence analysis. By way of comparison in 100 British white Caucasians an exclusion capacity of 0.97 was observed (Piercy et al.1993). While domesticated dog mtDNA is not as variable as human mtDNA, it can

still provide helpful clues in forensic cases (Savolainen and Lundeberg 1999, Schneider et al. 1999). Efforts have been made to standardize the nomenclature for the canine mtDNA control region (Pereira et al. 2004) and informative sequence variants outside of the control region have also been identified (Webb & Allard 2009).

1.5.3. Other Domesticated Animals.

While DNA testing of household animals like dogs and cats can help solve crimes, other domesticated animals—particularly animals used for recreation or sources of food—may be DNA tested for identification purposes. DNA tests have been developed for horses (Dimsoski 2003, van de Goor et al. 2010, Chen 2010), cattle (van de Goor et al. 2009), pigs (Robino et al. 2008), and sheep (Heaton et al. 2010). These genetic identification tests can be used to track the source of tainted meat products such as those obtained from cattle suffering from "madcow disease." Horse DNA testing can be important for confirming genetic pedigrees and is required for registering some breeds including American Quarterhorses and many racehorses (Bowling et al. 1997, Tozaki et al. 2001). Testing of non-human DNA

samples is not routinely performed in public forensic laboratories and thus these types of studies have to be outsourced to academic labs or speciality laboratories that focus on testing specific species (Ogden 2010). As noted earlier, the Veterinary Genetics Laboratory (VGL) at the University of California-Davis has been performing forensic animal DNA analyses since 1996. As of October 2010, the VGL offers genetic analyses for parentage verification, genetic disease screening, and diagnostic testing on alpaca, beefalo, bison, cat, cattle, deer, dog, elk, goat, horse, llama, pig, sheep, water buffalo, and yak samples. The U.S. Fish and Wildlife Service Forensic Laboratory in Ashland, Oregon conducts species identification as well as other DNA testing to aid fish and wild life forensic investigations. Feline STRs and mtDNA testing is performed by Quest Gen Forensics which also does canine STR and mtDNA testing to aid forensic investigations

Chapter II

Development of a Reference Database for Short Tandem Repeats (STRs) in dogs

2.1. Abstract

Dogs have intimate and close relationships with human for long time. Due to the large number of dogs in human populations, they are often involved in crimes. Occasionally, canine biological evidence such as saliva, bloodstains and hairs is found in crime scenes. Accordingly, canine DNA can be used as forensic evidence. The use of short tandem repeats (STRs) loci from biological evidence is valuable for forensic investigations. In Korea, canine STR profiling-related crimes are being successfully analyzed, leading to solve crimes such as animal cruelty, dog-attacks, murder, robbery, and missing or abandoned dogs. However, the probability of random DNA profile matches cannot be analyzed because of a lack of canine STR database. Therefore, in this study, 10 STR loci were analyzed in 600 dogs in Korea (344 dogs belonging to 30 different purebreds and 256 crossbred dogs) to estimate canine forensic genetic parameters. Among purebred dogs, a separate statistical analysis was conducted for five major subgroups, 97 Malteses, 47 Poodles, 31 Shih-tzus, 32 Yorkshire terriers, and 25 Pomeranians. Allele frequencies, expected (Hexp) and observed heterozygosity (Hobs), fixation index (F), probability of identity

(P(ID)), probability of sibling identity (P(ID)sib) and probability of exclusion (PE) were calculated. The Hexp values ranged from 0.901 (PEZ12) to 0.634 (FHC2079), while the P(ID)sib values were between 0.481 (FHC2079) and 0.304 (PEZ12) and the P(ID)sib was about 3.35×10^{-5} for the combination of all 10 loci. In forensics, profiling of short tandem repeats (STRs) is an ideal tool to identify individuals. This study is to create the first reference database of canine in Korea and it can improve the genetic evidential power of canine crimes in forensic caseworks.

2.2. Introduction

Dogs are one of the oldest and most intimate companion animals for humans in Korea. According to the Korean Pet Association, over five million dogs live in Korea and 10 million households own one or more pet animals. Because dogs live close to humans, they are often involved in forensic cases such as dog attacks, murder, animal abuse, missing dogs and robbery (Dayton et al. 2009). As a result, canine biological evidence found at crime scenes could become essential to solve criminal cases (Halverson and Basten, 2005, Butler 2006). Short Tandem Repeats (STRs) analysis is a useful tool in forensic identification (Lygo et al. 1994, Butler 2006, Ganco et al 2009). In cases in which there are genetic matches between known and unknown samples, STR profiling can play an important role in evaluating the value of the match (Balding and Nichols, 1994, Calboli et al. 2008, Denise et al. 2009). Despite the importance of canine STR profiles at the crime scene, canine DNA evidence is still underestimated in forensic investigations (Binns et al. 1995, Wictum et al. 2013). However, several countries including Austria (2005), the United States (2009), Hungary (2011), and the United Kingdom (2012) have started

constructing canine STR databases (Eichmann et al. 2005, Kanthaswamy et al. 2009, Ogden et al. 2012, Zenke et al. 2011). In Korea, canine STR profiling related crimes have been successfully analyzed, resulting in various crimes including animal cruelty, dog-attacks, murder, robbery, and missing or abandoned dogs being solved. However, the probability of a random DNA profile match cannot be calculated because of the lack of canine STR allele data for the dog population. In Korea, media-inspired copycat crimes of animal cruelty are on the rise. As a result, the demand for establishing related laws is rapidly increasing. In addition, dog-bite injuries can be fatal, especially among the young and elderly, and serious injuries can impact people for their entire lifetime. Moreover, the number of abandoned dogs is rapidly increasing, which has resulted in problems such as dealing with abandoned dogs. Currently, evidence from dog related crime cases can be admitted for trial in Korea. To improve the reliability of canine STR analysis, we determined the allele frequency and forensic informative values from 600 purebred and crossbred dogs (Halverson and Basten, 2005). The results revealed that 10 canine STR markers could be suitable for individual identification in forensic cases and enable estimation of the matching probability and discrimination of forensic

work.

2.3. Materials and Methods

2.3.1. Population

A total of 600 dogs were evaluated to investigate genetic variation in dogs in Korea. Of these, 344 dogs were of 30 different pure breeds and 256 were crossbred dogs. Among the purebred dogs, a separate statistical analysis of five major subgroups (97 Malteses, 47 Poodles, 31 Shih-tzus, 32 Yorkshire terriers, 25 Pomeranians) was carried to evaluate intra-breed variation.

2.3.2. DNA Extraction

Samples were obtained from local animal clinics and the Veterinary Medical Teaching Hospital at Jeju National University in Korea. DNA was extracted from buccal swabs and blood samples using a QIAamp® DNA Micro Kit (Qiagen, Santa Clara, CA). All samples from animal clinics were obtained after getting the owners' consent for DNA analysis. Extracted DNA was quantified on agarose gel. The quantity of DNA was set to 10ng/μl compared to K562 concentration by serial dilution (40ng/μl, 20ng/μl,10ng/μl, 5ng/μl and 2.5ng/μl).A multiplex

kit was optimized for 10ng/µl to 100ng/µl DNA to PCR amplification.

2.3.3. PCR Amplification and Quality Control

The 10 STR loci (PEZ1, FHC2054, FHC2010, PEZ5, PEZ20, PEZ12, PEZ3, PEZ6, PEZ8, and FHC2079) were amplified in a GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA) using the StockMarks® Dog Genotyping Kit according to the manufacturer's instructions. 10 STR markers were co-amplified in the PCR multiplex reaction. The total volume was 10µl including StockMarks® PCR Buffer 1.4µl, 25mM $Mgcl_2$ 0.36µl, dNTP mix 2.2µl, AmpliTag Gold Polymerase 0.36µl, amplification primer mix 2.8µl, deionized water 1.9µl, and 1µl of template DNA. The amplification process comprised an initial denaturation step at 95°C for 10 minutes followed by 20 cycles of 95°C for 30 seconds, 58°C for 30s and extension at 72°C for 60s, 15 cycles of 95°C for 30s, 56°C for 30s and extension at 72°C for 60s and then final extension at 72°C for 30 min. A multiplex kit was optimized for 10ng/µl to 100ng/µl DNA to PCR amplification. Canine control DNA (included in the StockMarks® Dog Genotyping Kit) was used as a positive control and deionized water as a negative control.

2.3.4. Typing and Analysis of Data

Capillary electrophoresis was performed on an ABI 3500 Genetic Analyzer (Applied Biosystems, Foster City, CA) using POP 7, 50 cm capillary arrays and default instrument settings (Applied Biosystems, Foster City, CA). The injection parameters on the ABI 3500 were 1.6kV and 8s. Data were analyzed using the GeneMapper v1.2 software. The fragment sizes of the allelic ladder were measured using GeneScan 500 ROX internal sizing standards. An allelic ladder is not provided with the StockMarks® Dog Genotyping Kit. Therefore, profiles from positive control canine DNA included with the kit were used to offset the allele bins instead. The positive control provided with the StockMarks® Dog Genotyping Kit was used to calibrate the allele sizes. PCR amplification of the canine control DNA template of electrophoresis conducted at different times determined the reproducibility and precision of the data. The dog parentage test and analysis of low copy number of DNA template were successfully resolved with the StockMarks® Dog Genotyping Kit. Occasionally, intermediate sized loci were formed due to imperfect repeat tandems. However, the genotyping sizes were calculated automatically using the GeneMapper software based on size estimation of the fluorescent

labeled DNA fragments. Fragment sizes were also compared with previous data (Gango et al. 2009), and the allelic frequencies of each marker were calculated using the GenAlEx 6 software (Peakall and Smoluse, 2006). The data were analyzed using GeneMapper v1.2 with detection of peak amplitude thresholds set to 100 RFUs (Relative Fluorescence Units). For peak quality, the homozygous min peak height was 200 and the heterozygous min peak height was 100. The Hardy-Weinberg equilibrium, estimated coefficients of inbreeding (F_{IS}) within breeds, fixation indices (F_{ST}) among breeds, and total inbreeding (F_{IT}) were determined using an analysis of variance framework implemented by the GenAlEx 6 software.

2.4. Results

Population genetic statistics parameters, such as allele frequencies, expected (Hexp) and observed heterozygosity (Hobs), fixation index (F), probability of identity (P(ID)), probability of sibling identity (P(ID)sib) and probability of exclusion(PE) were determined using the GenAlEx 6 software. The allele frequencies for each marker of 600 dogs (344 purebred and 256 crossbred dogs) are shown in Table 2-1.

Heterozygosity values were used to estimate the allele diversity or variation of forensic markers. The expected heterozygosity (Hexp) values ranged from 0.901 (PEZ12) to 0.634 (FHC2079). Of the 10 STR loci, the highest observed heterozygosity (Hobs) value was observed in PEZ12 (0.788) and the lowest in FHC2079 (0.375) loci.

P(ID)sib values were between 0.455 (PEZ5) and 0.304 (PEZ12). The estimated value of P(ID)sib was approximately 3.35×10^{-5} for the combination of all 10 loci.

Allele frequency and forensically informative statistical values of the 10 STR from intra-breed analysis of five breeds (Maltese, Poodle, Shih tzu, Yorkshire terrier, Pomeranian) were evaluated. The results are shown in Table 2-2.

The relative distances among breeds were examined by calculating the F_{ST} values. All pairwise comparisons among breeds showed significant F_{ST} values ($F_{ST} > 0$, P=0.01). Genetic diversity was also determined by calculating the level of inbreeding F_{ST}, F_{IS} and F_{IT} estimates for each analyzed STR locus among the five Korean dog population samples are shown in Table 2-3. Most breeds showed deviations from the Hardy-Weinberg equilibrium across loci as indicated by inbreeding coefficients, F_{IS}, significantly > 0 (Table 2-3).

The results show that the genetic diversity of 10 canine STR markers is sufficient to be a valuable tool in solving crime scene casework involving dog samples.

2.5. Discussion

DNA profiling of domestic dogs has become powerful evidence in forensics (Dayton et al. 2009). In Korea, canine STR profiling-related crimes have been successfully analyzed for over 10 years, resulting in various crimes including animal cruelty, dog-attacks, murder, robbery, and missing or abandoned dogs being solved. However, the probability of a random DNA profile match could not be measured because of the lack of canine STR population data, which sometimes caused delays in cases involving dogs. We conducted a canine population study of 600 unrelated dogs in Korea using 10 canine STR markers. Calculating match probability is an essential step in estimating the exact power of a DNA match (Berger et al. 2014). The probability of identity (PI), an individual identification estimator, is the probability that two individuals have the same genotype of multiple loci in a random population (DeNise et al. 2004). It is imperative to determine if the calculated P(ID)sib value is adequate for forensic applications. A low P(ID)sib value between 1.0×10^{-3} and 1.0×10^{-4} is considered reliable for individual identification in natural animal populations (Waits et al. 2001). The expected heterozygosity (Hexp) values ranged from 0.901

(PEZ12) to 0.634 (FHC2079), while the P(ID)sib values were between 0.481 (FHC2079) and 0.304 (PEZ12). Additionally, the estimated values of P(ID)sib were approximately 3.35×10^{-5} for the combination of all 10 loci. The allele frequencies and measures of locus informatives (number of alleles, expected heterozygosity, observed heterozygosity, and power of exclusion) of 558 dogs were calculated in the United States and are available in the Zoogen database (Halverson and Basten, 2005). Moreover, population parameters such as allele frequencies, Hardy-Weinberg-Equilibrium, expected and observed heterozygosity, fixation index, and probability of exclusion of 295 dogs samples from Austria and Germany were calculated (Berger et al, 2014). In the United Kingdom, the genetic diversity of 285 dogs from 13 popular breeds was analyzed (Ogden et al. 2012), while in Hungary, the allelic frequencies of 10 short tandem repeats (STRs) in 668 dogs were determined (Zenke et al. 2011). The genetic parameters (Hexp, Hobs, F, P(ID), P(ID)sib, and PE) and allele frequency values obtained in this study were very similar to those of other countries based on recently developed canine STR databases. The purpose of canine STR genotyping is to identify individual dogs. In forensic cases, estimation of population genetic parameters, including allele frequency, is used to

differentiate each individual as they can strengthen the weight of canine STR evidence. Problems related to dogs such as animal cruelty, attacks on people or animals, involvement at the crime scene, property damage and identifying lost pets have been rapidly increasing. In these kinds of crimes, a canine STR profiling-based database for Korea will be a valuable tool during investigation of crimes. In addition, the data could be a practical alternative to use of a dog microchip registration system to solve missing or abandoned dog cases. There has been no approach to measure canine genetic parameters of Korean domestic dogs to date. Therefore, these data can provide accurate estimates for forensic informative parameters of dog DNA in Korea. Construction of a canine STR population database in Korea will be a challenge to canine forensic applications.

Table 2-1. Allele frequency and forensically informative statistical values of the 10 analyzed STR loci in the "All breeds" samples.

ALLELE	PEZ1	PEZ2054	FHC2010	PEZ5	PEZ20	PEZ12	PEZ3	PEZ6	PEZ8	FHC2079
96	-	-	-	-	-	-	0.001	-	-	-
98	-	-	-	0.017	-	-	-	-	-	-
100	-	-	-	0.008	-	-	0.001	-	-	-
102	0.024	-	-	0.507	-	-	0.001	-	-	-
103	-	-	-	0.003	-	-	-	-	-	-
104	0.001	-	-	-	-	-	-	-	-	-
106	0.084	-	-	0.164	-	-	-	-	-	-
107	0.003	-	-	0.004	-	-	0.005	-	-	-
108	0.004	-	-	-	-	-	0.004	-	-	-
109	0.001	-	-	-	-	-	0.001	-	-	-
110	0.157	-	-	0.185	-	-	0.002	-	-	-
111	0.028	-	-	0.011	-	-	0.001	-	-	-
112	0.002	-	-	-	-	-	0.001	-	-	-
113	-	-	-	-	-	-	0.005	-	-	-
114	0.231	-	-	0.086	-	-	0.020	-	-	-
115	0.008	-	-	0.015	-	-	0.006	-	-	-
117	0.001	-	-	-	-	-	0.019	-	-	-
118	0.219	-	-	-	-	-	0.260	-	-	-
119	0.038	-	-	-	-	-	0.284	-	-	-
120	-	-	-	-	-	-	0.036	-	-	-
121	-	-	-	-	-	-	0.026	-	-	-
122	0.052	-	-	-	-	-	0.001	-	-	-
123	0.026	-	-	-	-	-	0.043	-	-	-
124	-	-	-	-	-	-	0.060	-	-	-
125	0.001	-	-	-	-	-	-	-	-	-
126	0.033	-	-	-	-	-	0.039	-	-	-
127	0.063	-	-	-	-	-	0.043	-	-	-
128	-	-	-	-	-	-	0.023	-	-	-
129	-	-	-	-	-	-	0.023	-	-	-
130	0.001	-	-	-	-	-	0.039	-	-	-
131	0.024	-	-	-	-	-	-	-	-	-
132	-	-	-	-	-	-	0.005	-	-	-

ALLELE	PEZ1	PEZ2054	FHC2010	PEZ5	PEZ20	PEZ12	PEZ3	PEZ6	PEZ8	FHC2079
133	-	-	-	-	-	-	0.033	-	-	-
134	-	-	-	-	-	-	0.002	-	-	-
135	0.001	-	-	-	-	-	0.003	-	-	-
136	-	-	-	-	-	-	0.005	-	-	-
139	-	-	-	-	-	-	0.001	-	-	-
140	-	-	-	-	-	-	0.003	-	-	-
143	-	-	-	-	-	-	0.002	-	-	-
144	-	-	-	-	-	-	0.001	-	-	-
146	-	0.057	-	-	-	-	-	-	-	-
149	-	-	-	-	-	-	0.001	-	-	-
150	-	0.169	-	-	-	-	-	-	-	-
151	-	0.047	-	-	-	-	-	-	-	-
153	-	-	-	-	-	-	0.003	-	-	-
154	-	0.077	-	-	-	-	-	-	-	-
155	-	0.136	-	-	-	-	-	-	-	-
158	-	0.005	-	-	-	-	-	-	-	-
159	-	0.077	-	-	-	-	-	-	-	-
162	-	0.001	-	-	-	-	-	-	-	-
163	-	0.124	-	-	-	-	-	-	-	-
164	-	-	-	-	-	-	-	0.001	-	-
167	-	0.168	-	-	-	-	-	0.003	-	-
169	-	-	-	-	-	-	-	0.003	-	-
170	-	-	-	-	0.002	-	-	-	-	-
171	-	0.094	-	-	0.183	-	-	0.063	-	-
172	-	-	-	-	0.035	-	-	-	-	-
173	-	-	-	-	-	-	-	0.004	-	-
175	-	0.043	-	-	0.329	-	-	0.176	-	-
176	-	-	-	-	0.023	-	-	0.002	-	-
177	-	-	-	-	-	-	-	0.006	-	-
178	-	-	-	-	0.001	-	-	0.023	-	-
179	-	0.003	-	-	0.180	-	-	0.216	-	-
180	-	-	-	-	0.011	-	-	0.061	-	-
181	-	-	-	-	0.002	-	-	0.013	-	-
182	-	-	-	-	0.006	-	-	0.094	-	-

ALLELE	PEZ1	PEZ2054	FHC2010	PEZ5	PEZ20	PEZ12	PEZ3	PEZ6	PEZ8	FHC2079
183	-	-	-	-	0.164	-	-	0.082	-	-
184	-	-	-	-	-	-	-	0.038	-	-
185	-	-	-	-	0.001	-	-	0.002	-	-
186	-	-	-	-	0.002	-	-	0.044	-	-
187	-	-	-	-	0.041	-	-	0.073	-	-
188	-	-	-	-	0.001	-	-	0.012	-	-
190	-	-	-	-	0.001	-	-	0.012	-	-
191	-	-	-	-	0.016	-	-	0.048	-	-
192	-	-	-	-	-	-	-	0.001	-	-
194	-	-	-	-	-	-	-	0.009	-	-
195	-	-	-	-	0.002	-	-	0.010	-	-
196	-	-	-	-	-	-	-	0.001	-	-
197	-	-	-	-	-	-	-	0.002	-	-
198	-	-	-	-	0.001	-	-	0.003	-	-
199	-	-	-	-	0.002	-	-	0.001	-	-
220	-	-	0.005	-	-	-	-	-	-	-
222	-	-	-	-	-	-	-	-	0.023	-
223	-	-	0.013	-	-	-	-	-	-	-
224	-	-	0.054	-	-	-	-	-	-	-
225	-	-	-	-	-	-	-	-	0.091	-
226	-	-	-	-	-	-	-	-	0.099	-
227	-	-	0.083	-	-	-	-	-	-	-
228	-	-	0.373	-	-	-	-	-	-	-
229	-	-	-	-	-	-	-	-	-	-
230	-	-	-	-	-	-	-	-	0.067	-
231	-	-	0.052	-	-	-	-	-	0.153	-
232	-	-	0.176	-	-	-	-	-	-	-
233	-	-	-	-	-	-	-	-	0.137	-
234	-	-	-	-	-	-	-	-	0.038	-
235	-	-	0.088	-	-	-	-	-	-	-
236	-	-	0.146	-	-	-	-	-	0.001	-
237	-	-	-	-	-	-	-	-	0.218	-
239	-	-	0.008	-	-	-	-	-	0.019	-
240	-	-	0.003	-	-	-	-	-	0.001	-

ALLELE	PEZ1	PEZ2054	FHC2010	PEZ5	PEZ20	PEZ12	PEZ3	PEZ6	PEZ8	FHC2079
241	-	-	-	-	-	-	-	-	0.119	-
242	-	-	-	-	-	-	-	-	0.003	-
243	-	-	-	-	-	-	-	-	0.001	-
245	-	-	-	-	-	-	-	-	0.023	-
246	-	-	-	-	-	-	-	-	0.002	-
249	-	-	-	-	-	-	-	-	0.008	-
256	-	-	-	-	-	0.002	-	-	-	-
257	-	-	-	-	-	0.001	-	-	-	-
260	-	-	-	-	-	0.033	-	-	-	-
261	-	-	-	-	-	0.005	-	-	-	-
262	-	-	-	-	-	0.003	-	-	-	-
264	-	-	-	-	-	0.128	-	-	-	-
265	-	-	-	-	-	-	-	-	-	0.005
267	-	-	-	-	-	0.003	-	-	-	-
268	-	-	-	-	-	0.152	-	-	-	0.005
269	-	-	-	-	-	0.021	-	-	-	0.515
271	-	-	-	-	-	0.103	-	-	-	-
272	-	-	-	-	-	0.083	-	-	-	-
273	-	-	-	-	-	0.009	-	-	-	0.163
275	-	-	-	-	-	0.165	-	-	-	-
276	-	-	-	-	-	0.045	-	-	-	-
277	-	-	-	-	-	0.007	-	-	-	-
279	-	-	-	-	-	0.092	-	-	-	-
280	-	-	-	-	-	0.003	-	-	-	-
282	-	-	-	-	-	0.002	-	-	-	-
283	-	-	-	-	-	0.034	-	-	-	-
284	-	-	-	-	-	0.003	-	-	-	-
285	-	-	-	-	-	-	-	-	-	0.001
286	-	-	-	-	-	0.003	-	-	-	-
287	-	-	-	-	-	0.009	-	-	-	-
289	-	-	-	-	-	0.003	-	-	-	0.020
290	-	-	-	-	-	0.033	-	-	-	-
291	-	-	-	-	-	0.002	-	-	-	-
292	-	-	-	-	-	0.001	-	-	-	-

ALLELE	PEZ1	PEZ2054	FHC2010	PEZ5	PEZ20	PEZ12	PEZ3	PEZ6	PEZ8	FHC2079
293	-	-	-	-	-	0.009	-	-	-	0.016
294	-	-	-	-	-	0.002	-	-	-	-
296	-	-	-	-	-	0.014	-	-	-	-
297	-	-	-	-	-	0.019	-	-	-	0.005
298	-	-	-	-	-	0.001	-	-	-	-
300	-	-	-	-	-	0.008	-	-	-	-
304	-	-	-	-	-	0.004	-	-	-	-
305	-	-	-	-	-	0.001	-	-	-	-
308	-	-	-	-	-	0.001	-	-	-	-

Table 2-2. Allele frequency and forensically informative statistical values of the 10 STR analyzed loci in the Maltese, Poodle, Shih tzu, Yorkshire terrier, Pomeranian and the "All breeds" samples.

Allele	PEZ 1					
	Maltese	Poodle	Shih tzu	Yorkshire terrier	Pomeranian	All breeds
102	0.025		0.016		0.02	0.024
104						0.001
106	0.031	0.011	0.661	0.016	0.06	0.084
107		0.011				0.003
108						0.004
109				0.016		0.001
110	0.186	0.106	0.065	0.063	0.22	0.157
111	0.021		0.016	0.016		0.028
112						0.002
114	0.237	0.202	0.016	0.453	0.16	0.231
115	0.01	0.021				0.008
117						0.001
118	0.304	0.372	0.145	0.203	0.14	0.219
119	0.046	0.021	0.081	0.063		0.038
122	0.041	0.032		0.078	0.02	0.052
123	0.01	0.032		0.063	0.04	0.026
125						0.001
126	0.005					0.033
127	0.082	0.191		0.016	0.24	0.063
130						0.001
131				0.016	0.1	0.024
135						0.001
N	97	47	31	32	25	600
Na	12	10	7	11	9	22
Ne	5.1	4.336	2.128	3.765	5.981	6.908
I	1.899	1.726	1.133	1.726	1.928	2.233
H_{OBS}	0.67	0.66	0.548	0.813	0.68	0.662
H_{EXP}	0.804	0.769	0.53	0.734	0.833	0.855

uH_{EXP}	0.808	0.778	0.539	0.746	0.85	0.856
F	0.166	0.143	-0.034	-0.106	0.183	0.226
PI	0.064	0.084	0.25	0.097	0.049	0.036
$PI_{(SIB)}$	0.364	0.386	0.547	0.407	0.346	0.331
PE	0.623	0.569	0.329	0.542	0.668	0.716

Allele	PEZ 2054					
	Maltese	Poodle	Shih tzu	Yorkshire terrier	Pomeranian	All breeds
146		0.191		0.25		0.057
150	0.062	0.085	0.21	0.328	0.2	0.169
151	0.015	0.021	0.048	0.016	0.04	0.047
154				0.031		0.077
155	0.175	0.117	0.306	0.094	0.16	0.136
158						0.005
159	0.129	0.085	0.032	0.016	0.04	0.077
162						0.001
163	0.418	0.117	0.081	0.031	0.06	0.124
167	0.155	0.33		0.156	0.16	0.168
171	0.031	0.053	0.29	0.047	0.2	0.094
175	0.01		0.032	0.031	0.14	0.043
179	0.005					0.003
N	97	47	31	32	25	600
Na	9	8	7	10	8	13
Ne	3.989	5.247	4.29	4.785	6.345	8.426
I	1.641	1.842	1.62	1.823	1.932	2.238
H_{OBS}	0.804	0.809	0.71	0.688	0.8	0.778
H_{EXP}	0.749	0.809	0.767	0.791	0.842	0.881
uH_{EXP}	0.753	0.818	0.779	0.804	0.86	0.882
F	-0.073	0.001	0.075	0.131	0.05	0.117
PI	0.094	0.059	0.091	0.071	0.045	0.026
$PI_{(SIB)}$	0.399	0.36	0.389	0.372	0.34	0.316
PE	0.546	0.636	0.553	0.603	0.682	0.76

Allele	FHC 2010					
	Maltese	Poodle	Shih tzu	Yorkshire terrier	Pomeranian	All breeds
220						0.005
223				0.031		0.013
224		0.011	0.016	0.031	0.06	0.054
227	0.098	0.011	0.048	0.109		0.083
228	0.505	0.426	0.355	0.516	0.36	0.373
231	0.01	0.021	0.113	0.063		0.052
232	0.098	0.383	0.274	0.109	0.08	0.176
235	0.036	0.032	0.016	0.047		0.088
236	0.247	0.117	0.177	0.094	0.48	0.146
239	0.005					0.008
240					0.02	0.003
N	97	47	31	32	25	600
Na	7	7	7	8	5	11
Ne	2.967	2.914	4.029	3.261	2.7	4.73
I	1.34	1.271	1.555	1.581	1.169	1.824
H_{OBS}	0.536	0.723	0.581	0.594	0.52	0.617
H_{EXP}	0.663	0.657	0.752	0.693	0.63	0.789
uH_{EXP}	0.666	0.664	0.764	0.704	0.642	0.789
F	0.191	-0.101	0.228	0.144	0.174	0.218
PI	0.158	0.181	0.101	0.117	0.204	0.069
$PI_{(SIB)}$	0.458	0.467	0.399	0.433	0.486	0.373
PE	0.429	0.394	0.53	0.502	0.364	0.608

Allele	PEZ 5					
	Maltese	Poodle	Shih tzu	Yorkshire terrier	Pomeranian	All breeds
98					0.02	0.017
100						0.008
102	0.613	0.383	0.548	0.766	0.52	0.507
103						0.003
106	0.067	0.34	0.274	0.188	0.06	0.164
107						0.004
110	0.18	0.191	0.016	0.047	0.3	0.185
111	0.01					0.011
114	0.124	0.085	0.161		0.04	0.086
115	0.005				0.06	0.015
N	97	47	31	32	25	600
Na	6	4	4	3	6	10
Ne	2.332	3.263	2.486	1.604	2.706	3.068
I	1.123	1.261	1.045	0.662	1.246	1.426
H_{OBS}	0.515	0.702	0.581	0.344	0.56	0.588
H_{EXP}	0.571	0.694	0.598	0.376	0.63	0.674
uH_{EXP}	0.574	0.701	0.608	0.382	0.643	0.675
F	0.098	-0.012	0.029	0.087	0.112	0.127
PI	0.225	0.152	0.227	0.433	0.192	0.145
$PI_{(SIB)}$	0.521	0.441	0.508	0.67	0.483	0.449
PE	0.346	0.431	0.331	0.183	0.383	0.45

Allele	PEZ 20					
	Maltese	Poodle	Shih tzu	Yorkshire terrier	Pomeranian	All breeds
170				0.016		0.002
171	0.284	0.234	0.339	0.203	0.38	0.183
172	0.031	0.064	0.048		0.06	0.035
175	0.335	0.5	0.129	0.328	0.26	0.329
176	0.01	0.032	0.016	0.016		0.023
178	0.005					0.001
179	0.165	0.138	0.113	0.219	0.16	0.18
180	0.046		0.032	0.016	0.02	0.011
181						0.002
182	0.01			0.031		0.006
183	0.093	0.021	0.177		0.06	0.164
185	0.005					0.001
186			0.032			0.002
187	0.005	0.011	0.113	0.156		0.041
188	0.005					0.001
190				0.016		0.001
191	0.005				0.06	0.016
195						0.002
198						0.001
199						0.002
N	97	47	31	32	25	600
Na	13	7	9	9	7	20
Ne	4.312	3.034	5.181	4.481	4.019	4.875
I	1.722	1.376	1.865	1.68	1.596	1.843
H_{OBS}	0.608	0.532	0.419	0.625	0.56	0.588
H_{EXP}	0.768	0.67	0.807	0.777	0.751	0.795
uH_{EXP}	0.772	0.678	0.82	0.789	0.767	0.796
F	0.208	0.207	0.48	0.195	0.255	0.26
PI	0.088	0.151	0.06	0.083	0.098	0.07
$PI_{(SIB)}$	0.388	0.453	0.361	0.382	0.399	0.37
PE	0.561	0.439	0.633	0.569	0.538	0.606

Allele	PEZ 12					
	Maltese	Poodle	Shih tzu	Yorkshire terrier	Pomeranian	All breeds
256				0.016		0.002
257					0.02	0.001
260	0.026	0.032	0.032	0.031	0.04	0.033
261					0.04	0.005
262						0.003
264	0.103	0.202	0.194	0.203	0.08	0.128
267						0.003
268	0.191	0.128	0.065	0.172	0.26	0.152
269	0.041		0.097	0.031		0.021
271	0.01	0.149		0.016	0.16	0.103
272	0.088	0.191		0.063	0.2	0.083
273	0.01	0.011		0.031	0.08	0.009
275	0.119	0.106	0.129	0.281		0.165
276	0.067	0.128	0.032	0.078	0.04	0.045
277	0.01			0.031		0.007
279	0.201	0.011	0.065	0.047	0.04	0.092
280	0.005					0.003
282						0.002
283	0.062				0.02	0.034
284						0.003
286						0.003
287	0.041	0.011				0.009
289	0.005	0.021				0.003
290	0.005		0.274			0.033
291	0.005					0.002
292						0.001
293						0.009
294						0.002
296	0.01	0.011	0.048			0.014
297			0.065			0.019
298						0.001
300					0.02	0.008

304						0.004
305						0.001
308						0.001
N	97	47	31	32	25	600
Na	18	12	10	12	12	35
Ne	8.192	6.871	6.428	6.006	6.51	10.065
I	2.347	2.072	2.061	2.062	2.119	2.623
H_{OBS}	0.835	0.766	0.677	0.594	0.68	0.788
H_{EXP}	0.878	0.854	0.844	0.833	0.846	0.901
uH_{EXP}	0.882	0.864	0.858	0.847	0.864	0.901
F	0.049	0.104	0.198	0.288	0.197	0.125
PI	0.026	0.038	0.041	0.047	0.04	0.018
$PI_{(SIB)}$	0.318	0.332	0.338	0.345	0.337	0.304
PE	0.757	0.707	0.697	0.678	0.729	0.801

Allele	PEZ 3					
	Maltese	Poodle	Shih tzu	Yorkshire terrier	Pomeranian	All breeds
96						0.001
100					0.02	0.001
102						0.001
107				0.078		0.005
108				0.047		0.004
109				0.016		0.001
110				0.031		0.002
111						0.001
112						0.001
113						0.005
114			0.032		0.02	0.02
115	0.005		0.016			0.006
117	0.01	0.021	0.032	0.125		0.019
118	0.052	0.011	0.065	0.063		0.26
119	0.284	0.436	0.306	0.266	0.2	0.284
120	0.113	0.021	0.016		0.06	0.036
121	0.077	0.032			0.14	0.026
122						0.001
123	0.057	0.032	0.097	0.078	0.02	0.043
124	0.046	0.032	0.113	0.016	0.12	0.06
126	0.113	0.064	0.065	0.016	0.04	0.039
127	0.062	0.043	0.065		0.12	0.043
128	0.057	0.011		0.109	0.06	0.023
129	0.057	0.064	0.065	0.016		0.023
130	0.021	0.16	0.048	0.031	0.04	0.039
132		0.011		0.047		0.005
133	0.036	0.053	0.065	0.063	0.06	0.033
134			0.016			0.002
135	0.005					0.003
136					0.04	0.005
139		0.011				0.001

140					0.02	0.003
143					0.04	0.002
144	0.005					0.001
149						0.001
153						0.003
N	97	47	31	32	25	600
Na	16	15	14	15	15	36
Ne	7.56	4.293	7.04	7.969	9.328	6.108
I	2.333	1.983	2.287	2.362	2.44	2.356
H_{OBS}	0.588	0.574	0.645	0.531	0.8	0.502
H_{EXP}	0.868	0.767	0.858	0.875	0.893	0.836
uH_{EXP}	0.872	0.775	0.872	0.888	0.911	0.837
F	0.323	0.251	0.248	0.393	0.104	0.4
PI	0.028	0.072	0.031	0.026	0.021	0.042
$PI_{(SIB)}$	0.323	0.384	0.329	0.319	0.309	0.342
PE	0.747	0.601	0.732	0.757	0.786	0.693

Allele	PEZ 6					
	Maltese	Poodle	Shih tzu	Yorkshire terrier	Pomeranian	All breeds
164						0.001
167						0.003
169						0.003
171	0.046	0.085	0.016	0.156	0.1	0.063
173			0.032			0.004
175	0.268	0.096	0.113	0.328	0.1	0.176
176						0.002
177						0.006
178		0.021		0.016		0.023
179	0.294	0.138	0.065	0.297	0.38	0.216
180	0.216	0.011	0.081	0.016	0.04	0.061
181	0.005		0.016			0.013
182	0.046	0.011	0.113	0.078	0.14	0.094
183	0.031	0.043	0.129	0.063	0.02	0.082
184	0.015	0.106	0.081	0.016		0.038
185		0.011				0.002
186	0.015	0.021	0.242			0.044
187	0.026	0.309	0.016		0.06	0.073
188			0.032			0.012
190			0.016	0.016	0.02	0.012
191	0.015	0.096	0.016	0.016	0.14	0.048
192			0.016			0.001
194			0.016			0.009
195	0.015	0.043				0.01
196						0.001
197						0.002
198		0.011				0.003
199	0.005					0.001
N	97	47	31	32	25	600
Na	13	14	16	10	9	28
Ne	4.717	6.403	8.214	4.321	4.771	8.85

I	1.843	2.16	2.371	1.714	1.833	2.503
H_{OBS}	0.701	0.809	0.935	0.719	0.68	0.74
H_{EXP}	0.788	0.844	0.878	0.769	0.79	0.887
uH_{EXP}	0.792	0.853	0.893	0.781	0.807	0.888
F	0.11	0.042	-0.065	0.065	0.14	0.166
PI	0.075	0.039	0.026	0.087	0.066	0.022
$PI_{(SIB)}$	0.375	0.338	0.317	0.388	0.371	0.312
PE	0.594	0.702	0.761	0.563	0.615	0.777

Allele	PEZ 8					
	Maltese	Poodle	Shih tzu	Yorkshire terrier	Pomeranian	All breeds
222	0.005					0.023
225	0.247	0.064	0.016	0.031	0.12	0.091
226	0.103	0.032	0.306	0.297	0.12	0.099
229	0.155	0.032	0.016		0.12	0.067
230	0.113	0.457	0.258	0.141	0.2	0.153
233	0.082	0.128	0.097	0.281	0.06	0.137
234	0.015	0.106	0.016	0.016	0.12	0.038
236						0.001
237	0.119	0.117	0.129	0.219	0.16	0.218
238	0.062	0.011				0.019
240						0.001
241	0.062	0.043	0.129	0.016	0.08	0.119
242	0.01					0.003
243	0.021					0.001
245	0.005	0.011	0.032		0.02	0.023
246						0.002
249						0.008
N	97	47	31	32	25	600
Na	13	10	9	7	9	17
Ne	7.249	3.865	4.878	4.231	7.485	7.746
I	2.164	1.736	1.777	1.564	2.082	2.225
H_{OBS}	0.753	0.596	0.581	0.563	0.68	0.693
H_{EXP}	0.862	0.741	0.795	0.764	0.866	0.871
uH_{EXP}	0.867	0.749	0.808	0.776	0.884	0.872
F	0.127	0.196	0.27	0.263	0.215	0.204
PI	0.033	0.089	0.07	0.095	0.033	0.036
$PI_{(SIB)}$	0.327	0.402	0.37	0.392	0.325	0.331
PE	0.727	0.557	0.605	0.541	0.729	0.716

Allele	FHC2079					
	Maltese	Poodle	Shih tzu	Yorkshire terrier	Pomeranian	All breeds
265	0.005					0.005
268						0.005
269	0.634	0.489	0.742	0.313	0.14	0.515
273	0.222	0.287	0.226	0.25	0.6	0.271
277	0.119	0.213	0.032	0.438	0.24	0.163
285						0.001
289	0.005	0.011				0.02
293					0.02	0.016
297	0.015					0.005
N	97	47	31	32	25	600
Na	6	4	3	3	4	9
Ne	2.148	2.722	1.66	2.844	2.285	2.734
I	0.994	1.086	0.668	1.072	1.002	1.22
H_{OBS}	0.381	0.319	0.323	0.563	0.52	0.375
H_{EXP}	0.535	0.633	0.398	0.648	0.562	0.634
uH_{EXP}	0.537	0.639	0.404	0.659	0.574	0.635
F	0.286	0.496	0.188	0.133	0.075	0.409
PI	0.269	0.204	0.42	0.197	0.25	0.191
$PI_{(SIB)}$	0.55	0.485	0.656	0.475	0.531	0.481
PE	0.296	0.353	0.183	0.357	0.31	0.378

N = Number

Na = No. of Different Alleles

Ne = No. of Effective Alleles

I = Shannon's Information Index

Hobs = Observed Heterozygosity

Hexp = Expected Heterozygosity

uHe = Unbiased Expected Heterozygosity

F = Fixation Index

PI = Probability of Identity

$PI_{(SIB)}$ = Probability of Identity(Siblings)

PE = Probability of Exclusion

Table 2-3. F-statistical (F_{ST}, F_{IS}, F_{IT}) values for each analyzed STR locus among five Korean dog populations.

F_{ST}	PD /PM	MTS /POM	POM /STZ	YT /POM	MTS /PD	MTS /STZ	MTS /YT	PD /STZ	PD /YT	STZ /YT	All Breeds
PEZ 1	0.034	0.029	0.170	0.088	0.008	0.243	0.008	0.281	0.024	0.319	0.062
FHC 2054	0.057	0.105	0.013	0.051	0.090	0.136	0.019	0.128	0.027	0.103	0.055
FHC 2010	0.139	0.051	0.052	0.109	0.070	0.038	0.001	0.008	0.040	0.027	0.075
PEZ 5	0.064	0.014	0.043	0.113	0.088	0.051	0.031	0.036	0.098	0.062	0.065
PEZ 20	0.040	0.002	0.008	0.026	0.017	0.029	0.004	0.098	0.040	0.043	0.057
PEZ 12	0.017	0.038	0.068	0.061	0.042	0.063	0.013	0.075	0.020	0.053	0.079
PEZ 3	0.046	0.007	0.014	0.029	0.025	0.004	0.065	0.011	0.033	0.010	0.016
PEZ 6	0.071	0.044	0.054	0.034	0.100	0.084	0.000	0.083	0.015	0.092	0.091
PEZ 8	0.038	0.010	0.003	0.050	0.093	0.067	0.003	0.066	0.052	0.023	0.039
FHC 2079	0.136	0.261	0.355	0.118	0.019	0.006	0.021	0.067	0.064	0.235	0.098

F_{IS}	PD /POM	MTS /POM	POM /STZ	YT /POM	MTS /PD	MTS /STZ	MTS /YT	PD /STZ	PD /YT	STZ /YT	All Breeds
PEZ 1	0.171	0.178	0.156	0.048	0.166	0.139	0.122	0.100	0.078	-0.061	0.150
FHC 2054	0.033	-0.037	0.085	0.112	0.041	-0.029	0.047	0.042	0.077	0.120	0.156
FHC 2010	0.006	0.196	0.233	0.174	0.103	0.209	0.189	0.053	0.021	0.202	0.192
PEZ 5	0.042	0.109	0.108	0.119	0.064	0.088	0.111	0.015	0.050	0.068	0.179
PEZ 20	0.238	0.225	0.399	0.238	0.214	0.284	0.216	0.339	0.209	0.353	0.147
PEZ 12	0.149	0.086	0.230	0.264	0.073	0.092	0.118	0.153	0.192	0.258	0.159
PEZ 3	0.208	0.285	0.199	0.281	0.308	0.312	0.336	0.262	0.325	0.336	0.173
PEZ 6	0.088	0.125	0.060	0.116	0.094	0.072	0.116	0.011	0.104	0.013	0.157
PEZ 8	0.217	0.153	0.270	0.257	0.154	0.167	0.192	0.239	0.261	0.281	0.168
FHC 2079	0.373	0.249	0.145	0.127	0.369	0.275	0.289	0.416	0.351	0.169	0.185

F_{IT}	PD /POM	MTS /POM	POM /STZ	YT /POM	MTS /PD	MTS /STZ	MTS /YT	PD /STZ	PD /YT	STZ /YT	All Breeds
PEZ 1	0.199	0.202	0.299	0.132	0.172	0.348	0.128	0.353	0.100	0.278	0.203
FHC 2054	0.088	0.072	0.097	0.157	0.053	0.111	0.065	0.165	0.101	0.210	0.202
FHC 2010	0.144	0.237	0.273	0.264	0.167	0.239	0.190	0.060	0.061	0.223	0.252
PEZ 5	0.104	0.122	0.146	0.218	0.147	0.134	0.139	0.051	0.143	0.125	0.232
PEZ 20	0.268	0.224	0.404	0.258	0.227	0.305	0.219	0.404	0.241	0.380	0.195
PEZ 12	0.164	0.121	0.282	0.309	0.112	0.149	0.130	0.217	0.208	0.297	0.226
PEZ 3	0.245	0.289	0.210	0.302	0.325	0.315	0.379	0.270	0.348	0.343	0.186
PEZ 6	0.153	0.163	0.111	0.146	0.185	0.150	0.116	0.093	0.117	0.103	0.234
PEZ 8	0.247	0.162	0.268	0.295	0.233	0.223	0.190	0.289	0.299	0.298	0.201
FHC 2079	0.458	0.445	0.449	0.230	0.381	0.279	0.304	0.456	0.392	0.365	0.265

PD: Poodle, POM: Pomeranian, YT: Yorkshire terrier, MTS: Maltese, STZ: Shih tzu.

Chapter III

Development of a Reference Database for mitochondrial HV1 region in dogs

3.1. Abstract

Dogs (*Canis lupus familiaris*) are common and widespread in human society. So, dog trace material is frequently found in forensic casework. Dog trace materials such as hairs, bloodstains, saliva, urine and feces are easily moved by contacting with carriers. Therefore, dog trace evidence can provide associate evidence connecting victims and suspects. Usually dog hairs are naturally shed, so most collected hairs found in crime scenes contain degraded nuclear DNA. This evidence, like single, shed hairs from dogs contains only limited amounts of DNA. The complete STR profiles cannot be obtained from this degraded and limited DNA. So, high copy number of mitochondrial genomes per cell is popular for forensic analysis, especially, when samples contain small amount, degraded or limited DNA. The canine mitochondrial DNA length of sequence is 16727 base-pairs. Like human, canine mitochondrial DNA (mtDNA) contains hypervariable regions (HVRs): hypervariable region 1 (HV1) and hypervariable region 2 (HV2). The HV1 of canine mitochondrial DNA (mtDNA) is highly polymorphic. Thus it is possible the severely degraded DNA sample can be amplified. To build a canine population database of

domestic dogs in Korea, the 612bp hypervariable region 1 (HV1) sequences of from 158 dogs (*Canis lupus familiaris*) in Korea were analyzed. In this study, 23 haplotypes from 35 single nucleotide polymorphisms (SNPs) were identified. The four most common HV1 haplotypes (n = 112 dogs) represented 70.9% of the total samples. Haplotype frequency is consistent with previous studies and exclusion capacity of mtDNA population is 0.851. The purpose of this study is to create the first reference database of canine mtDNA in Korea. Consequentially, mtDNA population data would improve the genetic evidential power of canine crimes in forensic caseworks.

3.2. Introduction

Dogs (*Canis lupus familiaris*) are living in human life together. Dog trace materials such as hairs, bloodstains, saliva, urine and feces are easily dispersed by contacting with human or other carriers. Usually dog hairs are naturally shed, so most collected hairs found in crime scenes contain degraded nuclear DNA. In forensics, profiling of short tandem repeats (STRs) from nuclear DNA is an ideal tool to identify individuals. However, this evidence such as single, shed hairs from dogs contains limited amounts of DNA. The complete STR profiles cannot be obtained from this degraded and limited DNA. In forensic analysis, even single dog hair could be valuable evidence that links victims and suspects. Actually, several articles related to dog trace evidence in forensic casework were published (Allen et al.1998, Schnider et al.1999 and Wetton et al. 2003).

Nuclear DNA exists in one copy per cell and it is easily degraded by environment due to a linear shape (Butler, 2005). On the contrary, circular maternally inherited mitochondrial DNA (mtDNA) is more stable to rapid environmental degradation. Mitochondria are found in numbers as high as 1000 per cell with as many as 10 genome copies per

mitochondrion (Nass 1969, Bohenhagen and Clayton 1974). Thus, high copy number of mitochondrial genomes per cell is popular for forensic analysis in case samples contain small amount, degraded or limited DNA (Budowle, 2005). The domestic dog mitochondrial genome consists of 16729 base-pairs were sequenced and published (Kim et al, 1998). The control region or D-loop of dog mtDNA is located between nucleotide position 15458 and 16727. The control region contains two hypervariable regions (HV1 and HV2) and a Variable Number of Tandem Repeat (VNTR) region. The HV1 region between nucleotide position 15458 and 16130 is the most polymorphic region and VNTR is a 10 base-paired tandem repeat, which is forensically not informative. The HV1 region has the highest mutation rate, so it is useful for DNA variation analysis.

To assess an evidential value of mtDNA match the haplotypes' random match probability must be evaluated. The random match probability is determined by calculating haplotype frequency. The more common a haplotype, the higher is probability between two dogs share this same haplotype incidentally. It means the evidential value of a match decrease (Verscheure 2013). Therefore, accurate estimation of haplotypes frequencies in genetic population is strongly required for

forensic applications. In recent years, studies of canine population database worldwide were conducted by Savolainen (1997, Sweden), Randi (2000, Italy), Branicki (2002, Poland), Gundry (2007, United States), Eichmann and Parson (2007, Austria), Hassell (2008, United Kingdom), Kanthaswamy (2008, United States), Toshinori (2013, Japan) and Desmyter (2014, Belgium).

Unfortunately, In Korea, there is only few genetic study of canine databases in forensics. The canine genetic research has focused on genetics of Korean traditional dog (RN Kim et al. 2012) or genetic diseases such as dislocation of hip joint (BH Choi 2010). In this study, to build a canine population database of domestic dogs in Korea, the 612-bp hypervariable region 1 (HV1) sequences of from 158 dogs (*Canis lupus familiaris*) were analyzed. The obtained haplotypes were consistent with previous population studies in other countries (Angelby et al. 2005, Webb et al. 2009, and Desmyter et al. 2009). Also, the exclusion capacity of canine mtDNA in the Korean population represented enough informative in forensic analysis.

3.3. Materials and Methods

3.3.1. DNA samples

A total of 158 domestic dogs, consisting 23 different breeds and mongrels, were collected with buccal swabs from veterinary practices and private owners in Korea (Table 3-1). The breeds of dogs were determined by owners. The 158 dogs are unrelated and domestic dogs in Korea. Unrelated familial relationship is important in mitochondrial DNA (mtDNA) studies because the mitochondrial DNA is maternally inherited. Familial relationships could affect to interpretations of mitochondrial DNA (mtDNA) nucleotide diversity. DNA was extracted from buccal swabs and blood samples using a QIAamp® DNA Micro Kit (Qiagen, Santa Clara, CA).

3.3.2. PCR amplification and Sequencing

The approximately 705 base-paired fragment was amplified from canine mitochondrial DNA (mtDNA) using the H15405 (5'-CCTAAGACTTCAAGGAAGAAGC-3') and L16110 (5'-CCCTAAAACTATATGTCCTGAA-3') primer set. The 705 base-

paired targeted the region of the dog mitochondrial genome hypervariable region 1 (HV1) nucleotide position 15405-16110. Each 25μL polymerase chain reaction (PCR) amplification mixture consisted with 1μL of DNA, 0.5μL of primer (20μM each), 2.5μL of Gold Star 10X Buffer, 0.25μL of Tag Gold (5U/μL) and water. The samples were amplified following conditions: 11 minutes at 95℃ for one cycle followed by 35 cycles of 30seconds at 94℃, 30s at 60℃, 90s at 72℃ and then a 4℃ hold. After amplification, PCR products were run on 2% agarose gel and stained with ethidium bromide to confirm PCR amplification. The PCR products were purified using QIAquick PCR purification kit (Qiagen, Santa Clara, CA). The purified amplicon was sequenced with BigDyeTerminator v3.1 Cycle Sequencing Kit (Applied BioSystems, Foster City, CA). The sequencing products were purified to remove unincorporated dye terminators and salt with DyeEx 2.0 Spin Kit (Qiagen, Santa Clara, CA). Sequencing was performed on an ABI 3500 Genetic Analyzer according to the manufacturer's instructions (Applied BioSystems, Foster City, CA).

3.3.3. Data analysis

The sequences were assembled in SeqScape version 2.0 instructions (Applied BioSystems, Foster City, CA). Sequence alignment was visualized and manually in MEGA version 5.02 (Tamura et al. 2011). Sequence variants and haplotypes of HV1 region were indicated relative to the reference dog sequence, a Sapsaree breed of Korea (Kim et al. GenBank Accession No. U96639). Exclusion capacity (PD) and random match probability (RMP) of canine mtDNA HV1 haplotype were calculated by $1 - \sum x_i^2$, $\sum x_i^2$ (where x_i^2 is the frequency of the *i*th haplotype) respectively. Fixation index (Fst) and an analysis molecular variance (AMOVA) to assess the genetic variation were evaluated by Arlequin 3.5 (Excoffier and Lischer, 2010).

3.4. Results

3.4.1. HV1 haplotypes

The 158 dogs sequenced exhibited 35 single nucleotide polymorphism (SNPs) and 23 HV1 haplotypes (Table 3-2). The 612 base-paired sequences were aligned and corresponded to nucleotides 15453-16112 of reference dog (Kim et al). Table 3-3 listed the haplotype frequency of the twenty-three different haplotypes. The haplotype frequency ranged from 0.006 (one representative) to 0.266 (forty-two representatives). Haplotype B1 was the highest haplotype frequency (frequency = 0.266) and included 46 numbers of dogs with 12 breeds. Haplotype B1 represented most breeds including Bichon Frise, Border Collie, Cockerspaniel, Dachshund, Greyhound, Maltese, Mix, Poodle, Poongsan, Shih tzu and English Cockerspaniel. Haplotype A18 was the second most frequent haplotypes of the data set, and the haplotype frequency was 0.203. Haplotype A18 represents 32 numbers of dogs, including Poodle, Schnauzer, Shih tzu, Spitz, Cockerspaniel and English Cockerspaniel. Haplotype A11 (frequency = 0.095) represented two Jindo dogs, a Malamute, four Malteses, two Pomeranians and three

Schnauzers. Especially, the breed of Yorkshire terrier including 7 numbers of dog was all detected only in haplotype A63 (haplotype frequency=0.044). It meant haplotype A63 occurred with the highest frequency in Yorkshire terrier breed (100 %). Haplotype B1 was common in Shih tzu (75 %). The breed of Maltese was found in 10 haplotypes, which was the highest number among the whole breeds. The four most common haplotypes (B1, A18, C1 and A11) were reported as 70.9 % of total 158 dogs. In total 56% of the Korean dog population sample belonged to haplogroup A, 29% to B and 15% to C (Figure 3-1). This portion showed a similar pattern with population of the United States (Webb and Allard, 2009) and Belgian population sample (Desmyter and Gijsbers, 2012). Representatives of haplogroup D, E and F were not indicated.

3.4.2. Statistical data analysis

The haplotype frequency was summed to determine the random match probability (RMP), which means the probability of random, unrelated dog would share an HV1 haplotype with a questioned dog. The exclusion capacity, or 1 - $\sum x_i^2$(where Xi is the frequency of the ith haplotype), means unrelated dog would have different HV1 haplotype

with a questioned dog. The exclusion capacity of the canine HV1 haplotype in this study is 0.851 (a random match probability of 0.149). This represents that 85 out of 100 unrelated different dogs and a questioned dog could be excluded in this study.

To assess population structure and distance among breeds the AMOVA analysis was performed. The fixation index (Fst) value in Table 3-4 and Table 3-5, it was represented the proportion of genetic variation within population. The result was very low, thus there’s no genetic relation among population. To evaluate the genetic relationship between dogs, the mtDNA haplotype data was calculated pairwise Fst values.

3.5. Discussion

The dog samples in this were collected randomly all over Korea. Considering of importance of assessing the dog mitochondrial DNA study, the 158 dogs were surely unrelated. Unrelated familial relationship is essential in mitochondrial DNA studies because the mitochondrial DNA is maternally inherited. Familial relationships could affect interpretations of mitochondrial DNA (mtDNA) nucleotide diversity.

The canine mitochondrial DNA length of sequence is 16727 base-pairs. Like the human, canine mitochondrial DNA (mtDNA) contains hypervariable regions (HVRs): hypervariable region 1 (HV1) and hypervariable region 1 (HV2). The HV1 of canine mitochondrial DNA (mtDNA) is highly polymorphic. Thus it is possible that severely degraded DNA samples can be amplified.

In this study, total of 158 dogs were sequenced and exhibited 35 single nucleotide polymorphism (SNPs) and 23 HV1 haplotypes. The 612 base-pair sequences were aligned and corresponded to nucleotides 15405-16110 of reference dog (Kim et al). The haplotype frequency ranged from 0.006 (one representative) to 0.266 (forty-two

representatives). The most common haplotype (B1) occurred in 27%, and followed in order of A18 (20%), C1 (15%) and A11 (9%) of total data set. In total 56% of the Korean dog population sample belonged to haplogroup A, 29% to B and 15% to C. This portion showed similar patterns to the United States' population sample (Webb and Allard, 2009) and Belgian population sample (Desmyter and Gijsbers, 2012). No representatives of haplogroup D, E and F were indicated. The haplotype frequency in this study is lined with previously published studies.

Many numbers of dog breed (Maltese or Poodle) were found in multiple haplogroups, and sometimes even belonged to different haplogroups. According to haplotype distribution, there weren't significantly observed correlations between breeds.

The exclusion capacity that is calculated in this study is consistent with those previously studied populations of other countries. The exclusion capacity and random match probability obtained from a rare haplotypes that may increase the evidential value and common haplotypes are applicable for exclusion purpose. The informative SNPs are derived from the set of sequences of 158 dogs that may be helpful distinguish dogs forensically.

To assess population structure among breeds the AMOVA analysis was conducted. The fixation index (Fst) value in Table 3-4 and Table 3-5, it is represented the proportion of genetic variation within population. The result was very low, so there was no genetic relation among population. The fixation index increases when dogs are grouped based on breeds, even they have identical sequences (Webb and Allard, 2009). They have similar sequences compare to randomly collected samples. This suggests that collecting samples regarding breed type is not crucial on mtDNA analysis. And it is strongly recommended that multiple individuals in same breed is needed for mtDNA analysis

Comparing the proportion of variable sites in the control region of canine mtDNA (5.4%) is approximately over twice of the entire mtGenome (2.2%) of sequences of 161 dogs of the study. This is the reason why we focus on analyzing canine mtDNA on hypervariable regions (Verscheure et al. 2014, 2015). The entire dog mtGenome is 16727 base-pair, which is about 16 times longer than control region. The entire mtGenome can provide additional discriminating sites, but it is known that these sites are rather found in control region than large parts of mtGenome (Webb et al. 2009, Imes et al. 2012).

In a Korean dog population, samples had 23 haplotypes, and their

haplotype frequencies spread of 1 % - 26.7 %. Consequently, applying variable SNPs and haplotypes to forensic casework may increase the discriminating value of canine mtDNA.

The most dogs have frequent haplotypes in most dog population in the world. The high frequent haplotypes have a greater impact on the discrimination power than less frequent ones (Verscheure et al. 2014). Therefore, even if the population data set has many rare haplotypes, the discrimination value from mtDNA matches of HV1 region is limited (Savolainen et al. 1997, Wetton et al. 2003, Gundry et al. 2007, Imes et al. 2012).

Practically, in this study, the exclusion capacity is somewhat lower than previous populations of other countries. The result were in order of 0.94 (Sweden), 0.93 (United Kingdom), 0.90 (China), 0.86 (Germany) and 0.95 (Japan) (Reina et al. 2008). Generally, increased sample size makes the haplotypes frequency to be increased by finding rare haplotypes. The influence of increased sample size on the distribution of haplotype frequency was evaluated by Webb and Allard (2010). As a result, to improve discriminatory power, expanding the length of sequences and sample size are the key by finding increased number of polymorphic sites. Nevertheless, the data from this study suggests the

canine HV1 region variation of dog mtDNA in Korea is sufficiently helpful to discriminate and has high forensic value. Typically, the mtDNA sequencing data applying discrimination between breeds and individuals is somewhat limited independently. However, in a complementary view, mtDNA sequencing analysis and STRs analysis for crimes related dogs are very powerful tools to solve forensic caseworks in multiple approaches. Furthermore, constructing the first reference database of STRs and mitochondrial DNA of Korean dog will be valuable to strengthen the potency of canine evidence.

Table 3-1. List of 158 samples from 23 breeds and mongrels.

Breeds	n
BichonFrise	3
Border collie	4
Cocker Spaniel	2
Dachshund	3
English Cockerspaniel	3
Greyhound	2
Jindo	5
King Charles Spaniel	1
Malamute	4
Malinois	1
Maltese	33
Miniature Pinscher	2
Mix (Mongrels)	29
Pekinese	5
Pomeranian	8
Poodle	21
Poongsan	3
Retriever	1
Schnauzer	9
Shetland Sheepdog	1
Shih tzu	8
Spitz	2
White Terrier	1
Yorkshire Terrier	7
Total	158

Table 3-2. Sequence variation, haplotypes and breeds of 158 Korean domestic dogs.

Breed		15483	15490	15508	15513	15522	15523	15526	15548	15553	15595	15611	15612	15620	15627	15628	15632	15636	15639	15643	15650	15652	15665	15710	15781	15800	15811	15814	15815	15819	15841	15912	15955	16003	16025	16038
Kim *et al.* Reference	Haplotype	C	A	C	G	T	T	C	C	A	C	T	T	T	A	T	C	T	T	A	C	G	T	C	C	T	A	C	T	T	C	C	C	A	T	C
BichonFrise 01	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
BichonFrise 02	C1			T				T				C							G		C					C		T				T	T	G		
BichonFrise 03	C1			T				T				C							G		C					C		T				T	T	G		
Border Collie 01	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Border Collie 02	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Border Collie 03	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Border Collie 04	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Cocker Spaniel 01	A18																											T								
Cocker Spaniel 02	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Dachshund 01	A9														A				A									T								
Dachshund 02	A20																											T							C	
Dachshund 03	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
English Cockerspaniel	A18																											T								
English Cockerspaniel	A18																											T								
English Cockerspaniel	B1				A			T			T		C				T		G	G		A				C		T	C			T	T	G		
Greyhound 01	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Greyhound 02	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Jindo 01	A11																		A									T							C	
Jindo 02	A18																											T								

Breed		15483	15490	15508	15513	15522	15523	15526	15548	15553	15595	15611	15612	15620	15627	15628	15632	15636	15639	15643	15650	15652	15665	15710	15781	15800	15811	15814	15815	15819	15841	15912	15955	16003	16025	16038
Kim *et al.* Reference	Haplotype	C	A	C	G	T	T	C	C	A	C	T	T	T	A	T	C	T	T	A	C	G	T	C	C	T	A	C	T	T	C	C	C	A	T	C
Jindo 03	A165																		A									T								
Jindo 04	A18																											T								
Jindo 05	A11																		A									T							C	
KingCharlesSpaniel 01	A17													C	G				A									T					T			
Malamute 01	A29														G				A			A						T								
Malamute 02	A29														G				A			A						T								
Malamute 03	A29														G				A			A						T								
Malamute 04	A11																		A									T							C	
Malinois 01	A18																											T								
Maltese 01	A18																											T								
Maltese 02	C1			T				T				C							G		C					C		T				T	T	G		
Maltese 03	C1			T				T				C							G		C					C		T				T	T	G		
Maltese 04	C1			T				T				C							G		C			T		C		T				T	T	G		
Maltese 05	A11																		A									T							C	
Maltese 06	A11																		A									T							C	
Maltese 07	C1			T				T				C							G		C					C		T				T	T	G		
Maltese 08	B1							T			T		C						G	G		A				C		T	C			T	T	G		
Maltese 09	A165																		A									T								
Maltese 10	B1				A			T			T		C						G	G		A				C		T	C			T	T	G		
Maltese 11	A18																											T								

Breed	Haplotype	15483	15490	15508	15513	15522	15523	15526	15548	15553	15595	15611	15612	15620	15627	15628	15632	15636	15639	15643	15650	15652	15665	15710	15781	15800	15811	15814	15815	15819	15841	15912	15955	16003	16025	16038
Kim *et al.* Reference	Haplotype	C	A	C	G	T	T	C	C	A	C	T	T	T	A	T	C	T	T	A	C	G	T	C	C	T	A	C	T	T	C	C	C	A	T	C
Maltese 12	A3														G				A									T							C	
Maltese 13	A11																		A									T							C	
Maltese 14	A2	T													G				A									T								
Maltese 15	A11																		A									T							C	
Maltese 16	C1			T				T				C							G		C			T		C	G	T				T	T	G		
Maltese 17	B1							T			T		C						G	G		A				C		T	C			T	T	G		
Maltese 18	A20																											T							C	
Maltese 19	A29														G				A			A						T								
Maltese 20	A18																											T								
Maltese 21	C1			T				T				C							G		C					C		T				T	T	G		
Maltese 22	B1							T			T		C						G	G		A				C		T	C			T	T	G		
Maltese 23	A18																											T								
Maltese 24	B1							T			T		C						G	G		A				C		T	C			T	T	G		
Maltese 25	A18																											T								
Maltese 26	A39														G			C	A			A	C					T			T					
Maltese 27	C1			T				T				C							G		C			T		C	G	T				T	T	G		
Maltese 28	A18																											T								
Maltese 29	B1																											T								
Maltese 30	A2	T													G				A									T								
Maltese 31	C1			T				T				C							G		C					C		T				T	T	G		

Breed		15483	15490	15508	15513	15522	15523	15526	15548	15553	15595	15611	15612	15620	15627	15628	15632	15636	15639	15643	15650	15652	15656	15710	15781	15800	15801	15814	15815	15819	15841	15912	15953	16003	16025	16038
Kim *et al.* Reference	Haplotype	C	A	C	G	T	T	C	C	A	C	T	T	T	A	T	C	T	T	A	C	G	T	C	C	T	A	C	T	T	C	C	C	A	T	C
Maltese 32	B1							T					C						G			A				C		T	C			T	T	G		
Maltese 33	B1							T					C						G			A				C		T	C			T	T	G		
Miniature Pincher 01	A29														G				A			A						T								
Miniature Pincher 02	A29														G				A			A						T								
Mix 01	A18																											T								
Mix 02	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Mix 03	B1				A			T			T		C				T		G	G		A				C		T	C			T	T	G		
Mix 04	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Mix 05	A18																											T								
Mix 06	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Mix 07	C1			T				T				C							G		C			T		C		T				T	T	G		
Mix 08	A18																											T								
Mix 09	A11																		A									T					T			
Mix 10	A18																											T								
Mix 11	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Mix 12	A2	T													G				A									T								
Mix 13	C1			T				T				C							G		C					C		T				T	T	G		
Mix 14	A11																		A									T							C	
Mix 15	C1			T				T				C							G		C			T		C		T				T	T	G		
Mix 16	A18																											T								

Breed		15483	15490	15508	15513	15522	15523	15526	15548	15553	15595	15611	15612	15620	15627	15628	15632	15636	15639	15643	15650	15652	15665	15710	15781	15800	15811	15814	15815	15819	15841	15912	15955	16003	16025	16038
Kim *et al.* Reference	Haplotype	C	A	C	G	T	T	C	C	A	C	T	T	T	A	T	C	T	T	A	C	G	T	C	C	T	A	C	T	T	C	C	C	A	T	C
Mix 17	B1				A			T			T		C				T		G	G		A				C		T	C			T	T	G		
Mix 18	A224													C	G				A									T					T			
Mix19	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Mix 20	A18																											T								
Mix 21	A224													C	G				A									T					T			
Mix 22	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Mix 23	C1			T				T				C							G		C					C		T				T	T	G		
Mix 24	A3														G				A									T							C	
Mix 25	A11																		A									T							C	
Mix 26	A15																		A									T								
Mix 27	A18																											T								
Mix28	A18																											T								
Mix29	A18																											T								
Pekinese 01	A15																		A									T								
Pekinese 02	A65																		A						T			T								
Pekinese 03	A18																											T								
Pekinese 04	A15																		A									T								
Pekinese 05	A65																		A						T			T								
Pomeranian 01	C1			T				T				C							G							C		T				T	T	G		
Pomeranian 02	B41							T			T		C				T		G			A				C		T	C			T	T	G		

Breed		15483	15490	15508	15513	15522	15523	15526	15548	15553	15595	15611	15612	15620	15627	15628	15632	15636	15639	15643	15650	15652	15665	15710	15781	15800	15811	15814	15815	15819	15841	15912	15955	16003	16025	16038
Kim *et al.* Reference	Haplotype	C	A	C	G	T	T	C	C	A	C	T	T	T	A	T	C	T	T	A	C	G	T	C	C	T	A	C	T	T	C	C	C	A	T	C
Pomeranian 03	A18																											T								
Pomeranian 04	A11																		A									T							C	
Pomeranian 05	A11																		A									T							C	
Pomeranian 06	C1			T				T				C							G		C					C		T				T	T	G		
Pomeranian 08	C1			T				T				C							G		C					C		T				T	T	G		
Pomeranian 09	A18																											T								
Poodle 01	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Poodle 02	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Poodle 03	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Poodle 04	C1			T				T											G		C			T		C		T				T	T	G		
Poodle 05	C1			T				T											G		C			T		C		T				T	T	G		
Poodle 06	C1			T				T				C							G		C					C		T				T	T	G		
Poodle 07	C1			T				T				C							G		C					C		T				T	T	G		
Poodle 08	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Poodle 09	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Poodle 10	A18																											T								
Poodle 11	A9														G				A									T								
Poodle 12	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Poodle 13	C1			T				T				C							G		C					C		T				T	T	G		
Poodle 14	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		

Breed		15483	15490	15508	15513	15522	15523	15526	15548	15553	15595	15611	15612	15620	15627	15628	15632	15636	15639	15643	15650	15652	15665	15710	15781	15800	15811	15814	15815	15819	15841	15912	15955	16003	16025	16038
Kim *et al.* Reference	Haplotype	C	A	C	G	T	T	C	C	A	C	T	T	T	A	T	C	T	T	A	C	G	T	C	C	T	A	C	T	T	C	C	C	A	T	C
Poodle 15	A18																											T								
Poodle 16	A18																											T								
Poodle 17	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Poodle 18	C1			T				T				C							G		C					C		T				T	T	G		
Poodle 19	A73														G				A			A	C					T								
Poodle 20	A18																											T								
Poodle 21	A16													C	G				A									T								
Poongsan 01	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Poongsan 02	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		T
Poongsan 03	A149														G				A									T								
Retriever 01	A16													C	G				A									T								
Schnauzer 01	A17													C	G				A									T					T			
Schnauzer 02	A11																		A									T							C	
Schnauzer 03	A11																		A									T							C	
Schnauzer 04	B12							T		G	G		C				T		G	G		A				C		T	C			T	T	G		
Schnauzer 05	A17													C	G				A									T					T			
Schnauzer 06	A18																											T								
Schnauzer 07	A11																		A									T							C	
Schnauzer 08	A15																		A									T								
Schnauzer 09	A2	T													G				A									T								

Breed		15483	15490	15508	15513	15522	15523	15526	15548	15553	15595	15611	15612	15620	15627	15628	15632	15636	15639	15643	15650	15652	15665	15710	15781	15800	15811	15814	15815	15819	15841	15912	15955	16003	16025	16038
Kim *et al.* Reference	Haplotype	C	A	C	G	T	T	C	C	A	C	T	T	T	A	T	C	T	T	A	C	G	T	C	C	T	A	C	T	T	C	C	C	A	T	C
Sheepdog 01	A5														G				A									T					T			
Shih Tzu 01	A18																											T								
Shih Tzu 02	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Shih Tzu 03	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Shih Tzu 04	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Shih Tzu 05	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Shih Tzu 06	A18																											T								
Shih Tzu 07	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Shih Tzu 08	B1							T			T		C				T		G	G		A				C		T	C			T	T	G		
Spitz 01	A19																																			
Spitz 02	A18																											T								
White Terrier 01	A2	T													G				A									T								
Yorkshire Terrier 01	A63																											T					T			
Yorkshire Terrier 02	A63																											T					T			
Yorkshire Terrier 03	A63																											T					T			
Yorkshire Terrier 04	A63																											T					T			
Yorkshire Terrier 05	A63																											T					T			
Yorkshire Terrier 06	A63																											T					T			
Yorkshire Terrier 07	A63																											T					T			

Kim *et al*. reference is sequence of complete nucleotide sequences of the domestic dogs (*Canis lupus familiaris*) mitochondrial genome.

Table 3-3. Canine mtDNA HV1 haplotypes distribution of 158 dogs analyzed in this study.

Haplotype	Breeds (n)	Total (n)	Frequency
A2	Maltese (2), Mix (1), Schnauzer (1), White Terrier (1)	5	0.032
A3	Maltese (1), Mix (1)	2	0.013
A5	Sheepdog (1)	1	0.006
A9	Dachshund (1), Poodle (1)	2	0.013
A11	Jindo (2), Malamute (1), Maltese (4), Mix (3), Pomeranian (2), Schnauzer (3)	15	0.095
A15	Mix (1), Pekinese (2), Schnauzer (1)	4	0.025
A16	Poodle (1), Retriever (1)	2	0.013
A17	King Charles Spaniel (1), Schnauzer (2)	3	0.019
A18	Jindo (2), Malinois (1), Maltese (6), Mix (9), Pekinese (1), Pomeranian (2), Poodle (4), Schnauzer (1), Shih tzu (2), Spitz (1), Cockerspaniel (1), English Cockerspaniel (2)	32	0.203
A19	Spitz (1)	1	0.006
A20	Maltese (1), Dachshund (1)	2	0.013
A29	Malamute (3), Maltese (1), Miniature Pincher (2)	6	0.038
A39	Maltese (1)	1	0.006
A63	Yorkshire Terrier (7)	7	0.044
A65	Pekinese (2)	2	0.013
A73	Poodle (1)	1	0.006
A149	Poongsan (1)	1	0.006
A165	Jindo (1), Maltese (1)	2	0.013
A224	Mix (2)	2	0.013
B1	BichonFrise (1), Border Collie (4), Cocker spaniel (1), Dachshund (1), Greyhound (2), Maltese (8), Mix (8), Poodle (8), Poongsan (2), Shih tzu (6), English Cockerspaniel (1)	42	0.266
B12	Schnauzer (1)	1	0.006
B41	Pomeranian (1)	1	0.006

C1	BichonFrise (2), Maltese (5), Mix (4), Pomeranian (3), Poodle (6), Maltese (3)	23	0.146
Total	23 breeds and 29 mongrels	158	1.000

Figure 3-1. Pie chart illustrating the distribution of the domestic dog HV1 haplotypes observed in this study data set (n = 158).

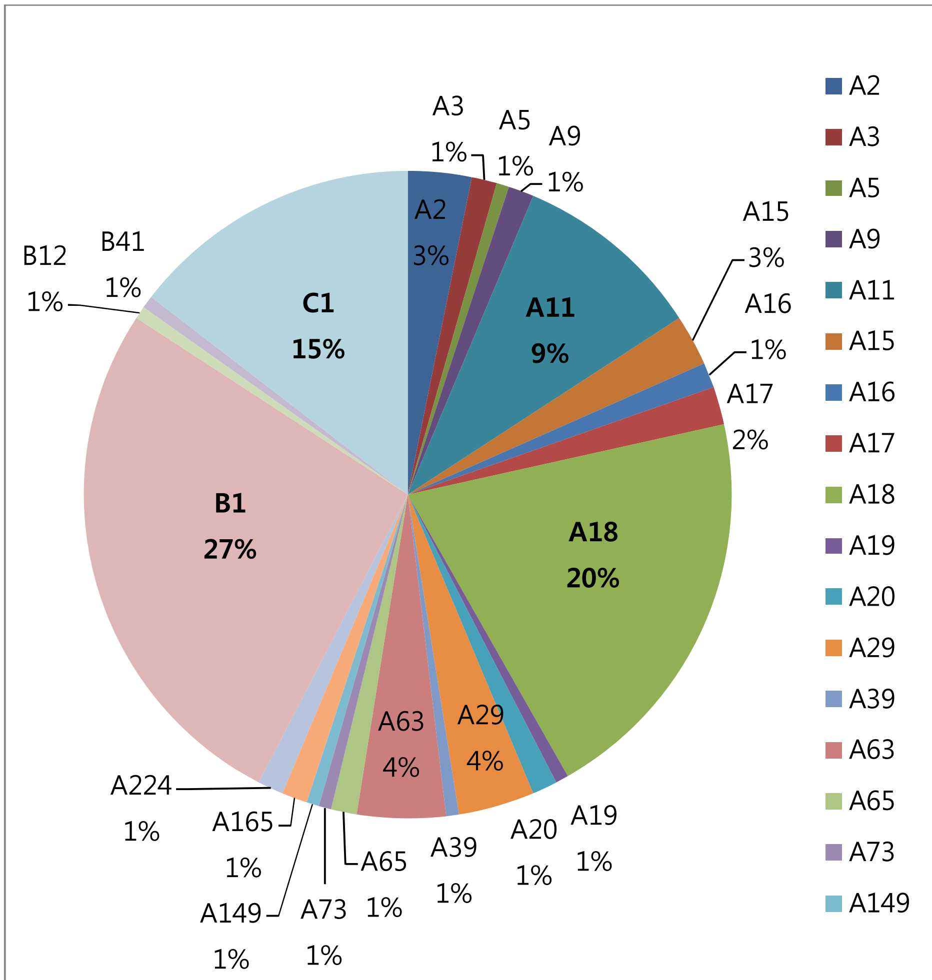


Table 3-4. Genetic distance for the frequencies of 158 dogs population (values of population pairwise Fst).

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	0																							
2	0.648	0																						
3	0.009	0.384	0																					
4	0.186	0.563	-0.484	0																				
5	0.212	0.561	-0.59	-0.352	0																			
6	0.463	0	0	0.347	0.344	0																		
7	0.514	0.765	0.0865	-0.08	0.005	0.681	0																	
8	0.5	1	-0.411	-0.45	-0.18	1	0	0																
9	0.717	0.938	0.3669	0.0789	0.258	0.909	0.068	0.538	0															
10	0.5	1	-0.714	-0.611	-0.62	1	-0.333	1	0.6	0														
11	-0.01	0.204	-0.245	-0.154	-0.13	0.098	0.004	-0.356	0.038	-0.449	0													
12	0.71	1	0.2941	0.1054	0.26	1	0.237	1	0.404	1	0.016	0												
13	-0.1	0.001	-0.29	-0.167	-0.16	-0.15	-0.056	-0.728	-0.065	-0.757	0.013	-0.182	0											
14	0.654	0.871	0.2857	0.0686	0.161	0.823	-0.014	0.375	0.243	0.117	0.044	0.495	-0.039	0										
15	-0.13	0.017	-0.289	-0.164	-0.16	-0.17	-0.059	-0.858	-0.06	-0.897	0.082	-0.228	-0.056	-0.042	0									
16	-0.01	0.223	-0.266	-0.041	-0.05	0.111	0.247	0.104	0.304	0.0709	0.014	0.32	0.019	0.309	0.114	0								
17	0.092	0.111	-0.511	-0.164	-0.13	-0.2	0.298	0	0.434	0.0322	-0.058	0.368	-0.14	0.439	-0.131	-0.078	0							
18	0.542	1	-0.411	-0.526	-0.23	1	-0.217	1	0.333	1	-0.37	1	-0.726	0.215	-0.863	0.136	0.032	0						
19	0.292	0.481	-0.021	-0.084	-0.01	0.375	-0.058	-0.599	-0.03	-0.563	0.087	-0.048	0.011	-0.035	0.02	0.272	0.148	-0.716	0					
20	0.609	1	-0.2	-0.074	-0.04	1	0.222	1	0.777	1	-0.15	1	-0.616	0.531	-0.811	0.306	0.25	1	-0.245	0				
21	0.051	0.456	-0.343	-0.076	-0.1	-0.14	0.204	-0.091	0.248	-0.145	0.045	0.19	-0.014	0.272	0.025	0.023	-0.215	-0.07	0.211	0.068	0			
22	0.626	0.979	0	-0.133	-0.16	0.96	-0.1	0.777	0.582	0.333	-0.111	0.888	-0.239	0.171	-0.282	0.204	0.322	0.714	-0.175	0.777	0.099	0		
23	0.542	1	-0.411	-0.705	-0.23	1	-0.217	1	0.333	1	-0.381	1	-0.719	0.215	-0.858	0.136	0.032	1	-0.675	1	-0.07	0.714	0	
24	0.369	0.595	0.0567	0.0112	0.056	0.493	0.058	-0.45	0.127	-0.321	0.085	0.146	-0.004	0.086	0	0.282	0.238	-0.321	0.015	0.013	0.22	-0.043	-0.321	0

Table 3-4. Continued.

No. (table above)	Breeds
1	Bichon Frise
2	Border collie
3	Cocker Spaniel
4	Dachshund
5	English Cockerspaniel
6	Greyhound
7	Jindo
8	King Charles Spaniel
9	Malamute
10	Malinois
11	Maltese
12	Miniature Pinscher
13	Mix (Mongrels)
14	Pekinese
15	MPomeranian
16	Poodle
17	Poongsan
18	Retriever
19	Schnauzer
20	Shetland Sheepdog
21	Shih tzu
22	Spitz
23	White Terrier
24	Yorkshire Terrier

Table 3-5. AMOVA design and result of 158 Korean domestic dogs ($p < 0.001$).

Source of Variation	Degree of Freedom	% Variation
Among samples sets	23	-2.00
Within sample sets	134	102.00
Total	157	100
	Fst = 0 (-0.01997)	

General Conclusions

From constructing canine STR typing of Korean 600 dogs, population genetic statistics showed forensically informative results. Allele frequencies, expected (Hexp) and observed heterozygosity (Hobs), probability of identity (P(ID)), and probability of exclusion(PE) were determined using the GenAlEx 6 software. The expected heterozygosity (Hexp) values ranged from 0.901 (PEZ12) to 0.634 (FHC2079). Of the 10 STR loci, the highest observed heterozygosity (Hobs) value was observed in PEZ12 (0.788) and the lowest in FHC2079 (0.375) loci. The results show that the genetic diversity of 10 canine STR markers is sufficient to be a valuable tool in solving crime scene casework involving dog samples. Secondly, development of a reference database for mitochondrial HV1 region in dogs was conducted with 158 domestic dogs. The 158 dogs sequenced exhibited 35 single nucleotide polymorphism (SNPs) and 23 HV1 haplotypes. The 612 base-paired sequences were aligned and corresponded to nucleotides 15453-16112 of reference dog (Kim et al). The haplotype frequency ranged from 0.006 (one representative) to 0.266 (forty-two representatives). Haplotype B1 was the highest haplotype frequency (frequency = 0.266). The four most common haplotypes (B1, A18, C1

and A11) were reported as 70.9 % of total 158 dogs. In total 56 % of the Korean dog population sample belonged to haplogroup A, 29 % to B and 15 % to C This portion showed a similar pattern with population of the United States (Webb and Allard, 2009) and Belgian population sample (Desmyter and Gijsbers, 2012). Representatives of haplogroup D, E and F were not indicated. The haplotype frequency was summed to determine the random match probability (RMP), which means the probability of random, unrelated dog would share an HV1 haplotype with a questioned dog. The exclusion capacity means unrelated dog would have different HV1 haplotype with a questioned dog. The exclusion capacity of the canine HV1 haplotype in this study is 0.851. The purpose of this study is to create the first reference database of canine STR and mtDNA in Korea. Consequentially, this STR and mtDNA population data would improve the genetic evidential power of canine crimes in forensic caseworks by complementary ways.

.

References

Allen. M., Engstrom. A. S., Meyers. S., Handt. O., Saldeen. T., von Haeseler. A. 1998. Mitochondrial DNA sequencing of shed hairs and saliva on robbery caps: sensitivity and matching probability. *J. Forensic Sci.* 43: 453-464.

Andrea. L. H., Theresa. F.S. Jessica A. S., Debra A. G., Wendy T. G., Venkat. S. M. S., David. G. S., Kristen. M., Webb. B. S., Allard. M. W., and Kanthaswamy. S. 2008. Forensic utility of the mitochondrial hypervariable region 1 of domestic dogs, in conjunction with breed and geographic information. *J. Forensic Sci.* 53: 81-89.

AnglebyH., and Savolainen. P. 2005. Forensic informativity of domestic dog mtDNA control region sequences. *Forensic Sci. Int*. 154: 99-110.

Balding. D. J., and Nichols. R. A. 1994. DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands. *Forensic Sci. Int.* 64: 125-140.

Berger, B., et al. 2008. Forensic canine STR analysis. In H. M. Coyle (Ed.), *Nonhuman DNA typing: Theory and casework applications* (Chapter 4, pp. 45–68). Boca Raton: CRC Press.

Berger, C., et al. 2009. Canine DNA profiling in forensic casework: the tail wagging the dog. *Forensic Sci. Rev.* 21: 1–14.

Berger. B., Berger. C., Hecht. W., Hellmann. A., Rohleder. U., Schleenbecker. U., and Parson. W. 2014. Validation of two canine STR multiplex-assays following the ISFG recommendations for non-human DNA analysis. *Forensic Sci. Int. Genet.* 8: 90-100.

Binns. M. M., Holmes. N.G., Marti. E., and Bowen. N. 1995. Dog parentage testing using canine microsatellites. *J. Small Anim. Pract.* 36: 493-497.

Bogenhagen. D., and Clayton. D. 1974. The number of mitochondrial deoxyribonucleic acid genomes in mouse L and human HeLa cells. Quantitative isolation of mitochondrial deoxyribonucleic acid. *J. Biol. Chem*. 249: 7991-7995.

Bowling, A. T., et al. 1997. Validation of microsatellite markers for routine horse parentage testing. *Anim. Genet.* 28: 247–252.

Branicki. W., Kalista, K., Kupiec, T., Wolánska-Nowak, P., Żołędziewska, M., and Lessig, R. 2005. Distribution of mtDNA haplogroups in a population sample from Poland. *J. Forensic Sci.* 50: 1-2.

Brauner, P., et al. 2001. DNA profiling of trace evidence-mitigating evidence in a dog biting case. *J. Forensic Sci.* 46: 1232–1234.

Budowle, B., et al. 1999. Mitochondrial DNA regions HVI and HVII population data. *Forensic Sci. Int.* 103: 23–35.

Budowle. B., Garofano. P., Hellman. A., Ketchum. M., Kanthaswamy., and Parson. W. 2005. Recommendation for animal DNA forensic identity testing. *Int. J. Legal. Med.* 119: 295-302.

Butler. J. M. 2005. *Forensic DNA typing*: Biology, Technology, and Genetics of STR Markers., second ed. Massachusetts, Burlington.

Butler. J. M. 2006. Genetics and genomics of core short tandem repeat loci used in human identity testing. *J. Forensic Sci.* 51: 253-265.

Calboli. F. C., Sampson. J., Fretwell. N., and Balding. D. J. 2008. Population structure and inbreeding from pedigree analysis of purebred dogs. *Genet.* 179: 593-601.

Chen, J. W., et al. 2010. Identification of racehorse and sample contamination by novel 24-plex STR system. *Forensic Sci. Int. Genetics*, 4: 158–167.

Choi. B. H. 2010. 개의 선천성 고관절 탈구 (고관절 이형성증) 관련 25개의 유전자 마커 개발에 대한 연구. 국립축산과학원 (National Institute of Animal Science).

Clarke, M., & Vandenberg, N. 2010. Dog attack: the application of canine DNA profiling in forensic casework. *Forensic Sci. Med. and Path.* 6: 151–157.

Dayton. M., Koskinen. M. T., Tom. B. K., Mattila. A. M., Johnston. E.,

Halverson. J., Fantin. D., DeNise.S., Budowle. B., Smith. D. G., and Kanthaswamy. S. 2009. Developmental validation of short tandem repeat reagent kit for forensic DNA profiling of canine biological material. *Croat. Med. J.* 50: 268-285.

D'Andrea, F. et al. 1998. Preliminary experiments on the transfer of animal hair during simulated criminal behavior. *J. Forensic Sci.* 43: 1257–1258.

DeNise. S., Johnston. E., Halverson. J., Marshall. K., Rosenfeld. D., McKenna. S., Sharp. T., and Edwards. J. 2004. Power of exclusion for parentage verification and probability of match for identity in American Kennel Club breeds using 17 canine microsatellite markers. *Anim. Genet*. 35: 14-17.

Desmyter. S., and Gijsbers. L. 2012. Belgian canine population and purebred study for forensics by improved mitochondrial DNA sequencing. *Forensic Sci. Int. Genet.* 6: 113-120.

Desmyter. S., and Comblez. S. 2009. Belgian dog mitochondrial DNA

database for forensics, *Forensic Sci. Int. Genet. Suppl. Ser. 2*: 286-287.

Dimsoski, P. 2003. Development of a 17-plex microsatellite polymerase chain reaction kit for genotyping horses. *Croat. Med. J.* 44: 332–335.

Eichmann. C., and Parson. W. 2007. Molecular characterization of the canine mitochondrial DNA control region for forensic applications. *Int. J. Legal Med.* 121: 411-416.

Eichmann. C., Berger. B., Steinlechner. M., and Parson. W. 2005. Estimating the probability of identity in a random dog population using 15 highly polymorphic canine STR markers. *Forensic Sci. Int.* 51: 37-44.

Ellegren. H. 2004. Microsatellites: simple sequences with complex evolution. *Nat. Rev.Genet.* 5: 435–445.

Ganco. L., Carvalho. M., Serra. A., Balsa. F., Bento. A. M., and Anjos. M. J. 2009. Genetic diversity analysis of 10 STR's loci used for forensic

identification in canine hair samples. *Forensic Sci. Int. Genet.* 2: 288-289.

Gundry. R. L., Allard. M. W., Moretti. T. R., Honeycutt. R. L., Wilson. M. R., Monson. K. L., and Foran. D. R. 2007. Mitochondrial DNA analysis of the domestic dog: control region variation within and among breeds. *J. Forensic. Sci.* 52: 562-572.

Halverson. H., and Basten. C. 2005. Forensic DNA identification of animal-derived trace evidence: tools for linking and suspects. *Croat. Med. J.* 46: 598-605.

Halverson. J., and Basten. C. 2005. A PCR multiplex and database for forensic DNA identification of dogs. *J. Forensic Sci.* 50: 352-363.

Halverson. J. L., and Basten. C. 2005. Forensic DNA identification of animal-derived trace evidence: tools for linking victims and suspects. *Croat. Med. J.* 46: 598-605.

Heaton, M. P., et al. 2010. Ovine reference materials and assays for

prion genetic testing. *BMC Veter. Res.* 6: 23.

Imes. D. L., Wictum. E. J. Allard. M. W., and Sacks. B. N. 2012. Identification of single nucleotide polymorphisms within the mtDNA genome of the domestic dog to discriminate individuals with common HV1 haplotypes. *Forensic Sci. Int. Genet.*6: 630-639.

Kanthaswamy. S., Tom. B.K., Mattila. A. M., Johnston. E., Dayton. M., Kinaga. J., Erickson. B.J., Halverson. J., Fantin. D., DeNise. S., Kou. A., Malladi. V., Satkoski. J., Budowle. B., Smith. D. G., and Koskinen. M. T. Canine population data generated from a multiplex STR kit for use in forensic casework. *J. Forensic Sci.* 54: 829-840.

Kim. K. S., Lee. S. E., Jeong. H. W., and Ha. J. H. 1998. The complete nucleotide sequence of the domestic dog (*Canis familiaris*) mitochondrial genome. *Mol. Phylogenet. Evol*. 10: 210-220.

Kim. R. N., Kim. D. S., Choi. S. H., Yoon. B. H., Kang. A., Nam. S. H., Kim. D.W., Kim. J. J., Ha. J. H., Toyoda. A., Fujiyama. A., Kim. A., Kim. M.Y., Park. K.H., Lee. K. S., Park. H. S. 2012. Genome

analysis of the domestic dog (Korean Jindo) by massively parallel sequencing. *DNA Res.* 19: 275-287.

Lygo. J. E., Johnson. P. E., Holdaway. D. J., Woodroffe. S., Whitaker. J.P., Clayton. T. M., Kimpton. C. P, and Gill. P. 1994. The validation of short tandem repeat (STR) loci for use in forensic casework. *Int. J. Legal. Med.* 107: 77-89.

Miller, K. W. & Budowle. B. 2001. A compendium of human mitochondrial DNA control region: Development of an international standard forensic database. *Croat. Med. J.* 42: 315–327.

Monson, K. L., et al. 2002. The mtDNA population database: An integrated software and database resource. *Forensic Sci. Comm.* 4(2).

Nass. M. 1969. Mitochondrial DNA. I. Intra mitochondrial distribution and structural relations of single and double length circular DNA. *J. Mol. Biol.* 42: 521-528.

Ogden, R. 2010. Forensic science, genetics and wildlife biology:

getting the right mix for a wildlife DNA forensics lab. *Forensic Sci. Med. and Path.* 6: 172–179.

Ogden. R., Mellanby. R. J., Clements. D., Gow. A. G., Powell. R., and McEwing. R. 2012. Genetic data from 15 STR loci for forensic individual identification and parentage analyses in UK domestic dogs (*Canis lupus familiaris*). *Forensic Sci. Int. Genet.* 6: e63-65.

Peakall. R., and Smoluse. P. E. 2006. GENEALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Exol. Notes*. 6: 288-295.

Pádár, Z., et al. 2001. Canine microsatellite polymorphisms as the resolution of an illegal animal death case in Hungarian zoological gardens. *Int. J. Legal Med.* 115: 79–81.

Pádár, Z., et al. 2002. Canine STR analyses in forensic practice: observation of a possible mutation in a dog hair. *Int. J. of Legal Med.* 116: 286–288.

Pereira, L., et al. 2004. Standardisation of nomenclature for dog mtDNA D-loop: a prerequisite for launching a *Canis familiaris* database. *Forensic Sci. Int.* 141: 99–108.

Piercy, R., et al. 1993. The application of mitochondrial DNA typing to the study of white Caucasian genetic identification. *Int. J. Legal. Med.* 106: 85–90.

Savolainen, P., & Lundeberg, J. 1999. Forensic evidence based on mtDNA from dog and wolf hairs. *J. Forensic Sci.* 44: 77–81.

Schneider, P. M., et al. 1999. Forensic mtDNA hair analysis excludes a dog from having caused a traffic accident. *Int. J. Legal. Med.* 112: 315–316.

Randi. E., Lucchini. V., Christensen. M. F., Mucci. N., Funk. S. M., Doilf. G., and Loeschcke. V. 2000. Mitochondrial DNA variability in Italian and East European Wolves: Detecting the consequences of small population size and hybridization. *Consev. Biol.* 14:463-473.

Satoh. M. & Kuroiwa. T. 1991. Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell. *Exp. Cell. Res.* 196: 137–140.

Reina H., Patrick H., Esther M. B. David. B., Cheryl H., Catherine T., Brian C., and Denise. S. C. 2008. Mitochondrial DNA analysis of domestic dogs in the UK. *Forensic Sci. Int. Genet. Suppl. Ser.* 1:598-599.

Sugiyama S., Chong. Y. H., Masayuki. S., Manami. K.,Tsuyoshi. K., Chihiro. U., Hiroshi. A., Makoto. B.,Shuichi. T., Atsushi. S., Hiroshi. O., Atsushi. N., and Toshinori. O. 2013. Analysis of mitochondrial DNA HVR1 haplotype of pure-bred domestic dogs in Japan. *Legal. Med.* 15: 303-309.

Savolainen.P., and Lundenberg. J. 1999. Forensic evidence based on mtDNA from dog and wolf hairs. *J. Forensic Sci.* 44: 77-81.

Savolainen. P., Rosen. B., Holmberg. A., Leitner. T., Uhlen.M., and Lundeberg. J. 1997. Sequence analysis of domestic dog mitochondrial

DNA for forensic use. *J. Forensic Sci.* 42: 593-600.

Schneider. P. M., Seo. Y., and Rittner. C. 1999. Forensic mtDNA hair analysis excludes a dog from having caused a traffic accident. *Int. J. Legal Med*. 112: 315-316.

Tamura. K., Peterson. D., Peterson. N., Stecher. G., Nei. M., and Kumar. S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28: 2731-2739.

Tom, B. K., et al. 2010. Development of a nomenclature system for a canine STR multiplex reagent kit. *J. Forensic Sci.* 55: 597–604.

Tozaki, T., et al. 2001. Population study and validation of paternity testing for Thoroughbred horses by 15 microsatellite loci. *J. Vet. Med. and Sci.* 63: 1191–1197.

Robin. E. D. & Wong. R. 1988. Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. *J. Cell.*

Physio. 136: 507–513.

Robino, C., et al. (2008). Forensic application of a multiplex PCR system for the typing of pig STRs. *Forensic Sci. Int. Genet. Suppl. Ser.* 1: 614–615.

Röhl, A., et al. 2001. An annotated mtDNA database. *Int. J. Legal Med.* 115: 29–39.

van Asch, B., et al. 2009. A new autosomal STR nineplex for canine identification and parentage testing. *Electrophoresis*, 30: 417–423.

van de Goor, L. H. P., et al. 2010. A proposal for standardization in forensic equine DNA typing: allele nomenclature for 17 equine-specific STR loci. *Ani. Genet.* 41: 122–127.

Verscheure. S., Backeljau. T., and Desmyster. S. 2013. Reviewing population studies for forensic purposes: dog mitochondrial DNA. *Zookeys* 38: 1-411.

Verscheure. S., Backeljau. T., and Desmyter. S. 2014. Dog mitochondrial genome sequencing to enhance dog mtDNA discrimination power in forensic casework. *Forensic Sci. Int. Genet.* 12: 60-68.

Verscheure. S., Backeljau. T., and Desmyter. S. 2015. Coding region SNP analysis to enhance dog mtDNA discrimination power in forensic casework. *Forensic Sci. Int. Genet.* 14: 86-95.

Waits. L. P., Luikart. G., and Taberlet. P. 2001. Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Mol. Ecol*. 10: 249-256.

Walsh. P. S. et al. 1996. Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. *Nucleic Acids Res*. 24: 2807–2812.

Webb. K. M., and Allard. M. W. 2009. Identification of forensically informative SNPs in the domestic dog mitochondrial control region. *J. Forensic. Sci.* 54: 289-304.

Webb. K. M., and Allard. M. W. 2009. Mitochondrial genome DNA analysis of the domestic dog: identifying informative SNPs outside of the control region. *J. Forensic. Sci.* 54: 275-288

Webb. K. M., and Allard. M. W. 2010. Assessment of minimum samples sizes required to adequately represent diversity reveals inadequacies in datasets of domestic dog mitochondrial DNA. *Mitochondr. DNA 21*: 19-31.

Werrett, D. J., & Sparkes, R. 1998. 300 matches per week – the effectiveness and future development of DNA intelligence databases – parts 1 and 2. *Proceedings of the ninth international symposium on human identification* (pp. 55–62).

Wetton. J., Higgs. J., Spriggs. A., Roney. C., Tsang., and Foster. A. 2003. Mitochondrial profiling of dog hairs. *Forensic Sci. Int.* 133: 235-241.

Wictum. E., Kun. T., Lindquist. C., Malvick. J., Vankan. D., and Sacks. B. 2013. Developmental validation of DogFiler, a novel multiplex for

canine DNA profiling in forensic casework. *Forensic Sci. Int. Genet*. 7: 82-91.

Zenke. P., Egyed. B., Zoldag. L., and Padar. Z. 2011. Population genetic study in Hungarian canine populations using forensically informative STR loci. *Forensic Sci. Int. Genet.*5: e31-36.

국문 논문 초록

한국 내 개의 STRs (Short Tandem Repeats) 및 미토콘드리아 고변이 지역 분석을 통한 집단 유전학적 데이터베이스 구축에 대한 연구

(지도교수: 조 성 범)

서울대학교 대학원
수의학과 수의공중보건학 전공
문 서 현

삶의 질 향상, 사회구성원의 고령화, 핵가족화로 반려견에 대한 관심과 수요가 증가하여, 반려견을 기르는 대한민국 인구는 약 천만 명, 반려견의 수는 5백만 마리, 반려견을 키우는 가정의 비율은 17.4% 이다. 사람과 가장 밀접한 생활을 하는 동물인 개의 생물학적 시료(혈흔, 타액, 모발, 소변 등 생체 분비물)는 범죄현장에서 빈번하게 발견되고, 이는 범죄를 해결하는 유용한 증거가 되기도 한다. 즉, 개의 DNA는 살인, 강도, 성폭행, 동물 학대, 절도 등의 범죄현장과 범인을 성공적으로 연결할 수 있는 도구가 될 수 있다. 범죄현장에서 개의 DNA 타이핑을 통한 현장 내 불상의 시료와 대조 시료와

의 일치함을 확인할 경우, 일치에 대한 평가는 매우 중요한 법과학적 의미를 가진다. 이때 인간 개체식별 분야와 마찬가지로 개의 표준집단의 유전학적 데이터베이스는 확률상의 일치율 측정을 가능하게 한다. 미국, 영국, 호주, 벨기에, 터키, 헝가리, 오스트리아, 이탈리아 등의 나라들은 STR(Short Tandem Repeat)을 통한 개의 개체식별 결과를 효과적으로 적용시키기 위하여 각 나라의 순종과 잡종으로 구성된 시료집단의 유전학적 데이터베이스 구축을 최근 완료하였다. 또한 개의 미토콘드리아 DNA 분석을 통한 canine mtDNA haplotype reference database 구축 필요성이 요구되고 있는 실정이다. 세계 여러 나라에서 극미량의 검체에서도 분석 가능한 개의 미토콘드리아 DNA의 증거력으로써 유용함을 연구하였고, 실제로 개의 미토콘드리아 분석 및 이를 발전시킨 개의 미토콘드리아 DNA haplotype 데이터베이스는 개 특이적이고, 범죄현장과 연관이 되지 않은 개체의 개를 배제시키는 신속한 방법이 될 수 있다. 이 연구는 이미 세계 각 나라별로 구축된 개의 개체 식별 분석 및 모계 유전자형 분석 두 가지 방법을 한국 내 개 집단에 적용하여 진행되었다. 한국 내 개에 대한 통계학적 분석을 통하여 개의 집단 유전학적 분석 및 법과학적 분석을 수행하였다. 개의 개체식별 분석법(600마리)의 결과는 모든 좌위에서 하디-와인버그 평형 상태에 있는 결과를 보였으며, 동일 염색체 상에 존재하는 모든

좌위들에서 나타난 독립적인 상태는 법과학적 적용을 위한 독립적인 확률 계산 시에 실제 적용이 가능함을 보였다. 또한, 핵 DNA 손상 시 개체 식별 분석이 불가능한 자연 탈락된 개의 털, 유골 등과 같이 손상되거나 극미량의 DNA를 가진 시료 분석에 사용되는 개의 미토콘드리아 고변이 지역 유전자 분석법으로, 한국 내 개(180마리) 집단에서 35개의 단일 염기 다형성을 발견하여 mitochondria DNA haplotype 23종을 분류하여 한국 내 개의 mtDNA haplotype database 구축하였다. 이는 사건현장에서 개와 관련된 증거물의 배제 여부를 신속하게 판단함으로써, 개체 식별 (STR typing) 분석을 보완하는 역할을 할 것이다.
이 연구는 한국 내 개의 개체 식별 분석법 및 미토콘드리아 고변이 지역 유전자 분석법을 수행하여 국내 개의 집단 유전학적 데이터베이스를 처음으로 구축함으로써 개와 관련된 사건에서 유전자 분석의 법과학적 증거 능력을 향상시킬 수 있을 것이다.

주요어: 개의 생물학적 시료, 범죄 현장, 개체 식별 분석법, 미토콘드리아 고변이 지역 유전자 분석법, 집단 유전학적 데이터베이스 구축, 유전자 분석의 법과학적 증거능력
학번: 2012-30451