VALIDATION AND DATABASE GENERATION OF 100 CANINE MICROSATELLITE PROFILES FOR CRIME AND PATERNITY TESTING

By:

Jonti Eyles

(33384684)

A thesis submitted in fulfilment of the requirements of the degree of Master of Forensic Science (Professional Practice)

In the School of Veterinary and Life Sciences

Murdoch University

Supervisors:

Brendan Chapman

Dr Claire Sharp

Dr Peter Spencer

Semester 2, 2019

DECLARATION

I declare that this manuscript does not contain any material submitted previously for the award of any other degree or diploma at any university or other tertiary institution. Furthermore, to the best of my knowledge, it does not contain any material previously published or written by another individual, expect where due references have been made in the text. Finally, I declare that all reported experimentations performed in this research were carried out by myself, except that any contribution by others, with whom I have worked is explicitly acknowledged.

Signed: *Jonti Eyles*

Date: *18/01/2020*

ACKNOWLEDGEMENTS

I would like to acknowledge Murdoch University for providing the funding, equipment and facilities required to complete this project and the Animal Hospital at Murdoch University for supplying the canine samples used in this study. I would like to thank my supervisors Brendan Chapman, Claire Sharp and Peter Spencer for providing support and assistance throughout this project.

Thanks, as always, to my family and friends.

TABLE OF CONTENTS

BLANK PAGE – NOT NUMBERED

PART ONE

LITERATURE REVIEW

VALIDATION AND DATABASE GENERATION OF 100 CANINE MICROSATELLITE PROFILES FOR CRIME AND PATERNITY TESTING

1. ABSTRACT

Almost two in five Australian households have dogs, subsequently, canine biological evidence is often part of the physical evidence found at crime scenes. The genetic analysis of canine biological material can provide valuable links between suspects, victims and crime scenes, aiding investigations. However, as canine DNA evidence has been underutilised in forensic casework, a validated Western Australian canine microsatellite database is non-existent. Therefore, by employing a commercial canine STR kit, a Western Australian canine population database can be created containing allele frequencies, thus permitting the statistical weighting of evidence. The generation of the database will enable reliable interpretation of canine biological evidence for forensic casework.

Keywords: Forensic Science, Canine DNA, Canine quantitation, Canine genotyping, Canine database

TABLE OF CONTENTS

---

2. LIST OF TABLES

3. LIST OF ABBREVIATIONS

AKC	American Kennel Club
bp	Base pair
C_T	Cycle threshold
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
dsDNA	Double-stranded DNA
HV1	Hypervariable 1
HV2	Hypervariable 2
HWD	Hardy-Weinberg disequilibrium
HWE	Hardy-Weinberg equilibrium
ISAG	International Society for Animal Genetics
LE	Linkage equilibrium
LR	Likelihood ratio
MC1R	Melanocortin-1 receptor
mL	Millilitre
mtDNA	Mitochondrial DNA
ng	Nanogram
nm	Nanometre
pg	Picogram
qPCR	Quantitative real-time PCR
RFU	Relative fluorescence unit
RMP	Random match probability
SINE	Short interspersed element
SNP	Single nucleotide polymorphism
STR	Short tandem repeat
SWGDAM	Scientific Working Group for DNA Analysis Methods
U.S.	United States
µL	Microlitre
UV	ultraviolet
θ	theta
5′	Five prime

4. INTRODUCTION

As of 2016, an estimated 4.8 million dogs were owned as pets in Australia, averaging 1.3 dogs per household (1). With more than 100,000 people suffering dog-related injuries per year, the ability to identify the individual dog responsible based on biological evidence is critical for forensic investigations (2). Domestic canine material, in the form of hair, saliva, blood, urine and faeces, is often present on evidentiary exhibits and regularly encountered at crime scenes (3). These crime scenes are not limited to cases involving dogs as the perpetrators or victims of attack, but also include instances of property damage, animal cruelty, theft of animals and destruction to native fauna and livestock (4). The analysis of canine-related biological material has contributed to court proceedings that have ultimately resulted in the exoneration or destruction of the dog involved (4). Furthermore, the subsequent recovery and genetic analysis of dog-related material have proven to provide links between not only animal suspects and victims but also between human suspects, victims and crime scenes (5).

Routine human STR analysis is the mainstay of most government-funded and private forensic laboratories as it provides a reliable approach for weighting evidence (6). The techniques employed to recover and analyse samples containing human biological material are considered well established, sensitive and highly discriminative (6). Following the collection of biological material, the sample is exposed to DNA extraction and then DNA quantitation methods to determine the concentration of DNA recovered (6). Next, utilising a commercial kit, STR regions of the extracted sample are amplified via PCR, allowing the generation and interpretation of an STR profile (6). Similarly to human DNA analysis, the examination of canine DNA involves many well studied STR markers that have been

validated through prior research and casework (4). Hence the working method implemented by forensic laboratories to analyse human DNA is applicable to the examination of canine biological material (4).

Despite the potential for providing intelligence to police investigations, the analysis of canine DNA evidence is not common in forensic casework (7). Subsequently, there is a lack of standardisation concerning methodology, interpretation and reporting. Recent court challenges regarding canine DNA evidence necessitated the need for a standard and validated canine STR panel (7). The Canine Genotypes™ Panel 2.1 was manufactured specifically for forensic casework and optimised for standard PCR conditions allowing easy implementation into forensic laboratories (7). Through developmental studies, the Canine Genotypes™ Panel 2.1 demonstrated high allelic diversity at the 18 STR markers allowing reliable identification of individual dogs (7). As the commercial kit requires 1–2 ng of a high quality DNA sample in a 20 µL reaction volume to deliver optimal results, accurate detection and quantitation of canine DNA is critical for STR genotyping (8). Unlike human DNA quantitation, commercial canine-specific DNA quantitation kits or instruments are not available. Therefore, current canine DNA quantitation methods rely on total DNA quantitation or laboratory-developed qPCR assays. As forensic examinations depend on validity and standardisation, a viable working method for canine DNA analysis is critical to obtaining reliable results.

Comparable to forensic cases involving human DNA, cases involving canine biological material frequently require the examination and comparison of reference samples to crime scene samples. Therefore, a relevant dog database is necessary to

effectively interpret the information obtained from the resultant canine reference and crime scene STR profiles (9). Studies pertaining to canine microsatellite databases are generally limited to the U.S., predominated by pedigreed dogs and are private in-house databases (9). Currently, in Western Australia, there are no databases containing population allele frequencies for the 18 STR markers examined by the Canine Genotypes™ Panel 2.1 kit. Consequently, results obtained from the analysis of dog biological material found at crime scenes or investigations regarding paternity verification cannot be dependably interpreted due to the lack of a relevant Australian database. The generation and validation of a Western Australian database will enable the statistical significance of a canine DNA match to be reliably calculated for use in forensic casework. Furthermore, the creation of the canine microsatellite database has the potential to be a valuable national resource and further the application and development of canine DNA analysis in the forensic science community.

5. CANINE DNA QUANTITATION

Quantitation of target DNA extracted from forensic samples is crucial to obtaining reliable DNA profiles and minimising sample consumption (6). As commercial kits available for STR genotyping function best with a narrow concentration range of template DNA, accurate quantitation of template DNA is required for obtaining optimal results (6). When not enough target DNA is added to PCR, stochastic amplification and failure to equally sample the STR alleles may result in allelic dropout or imbalanced DNA profiles (6). On the other hand, when too much target DNA is added to PCR, the fluorescent signal can be overwhelming and off-scale fluorescent peaks may be evident, making interpretation of results more difficult (6). Hence based on quantitation data, the amount of template DNA

in the PCR can be diluted or concentrated to fit the optimal window for PCR amplification to produce robust and reliable results (6). Furthermore, various quantitation methods have been implemented in laboratories to determine the quantity of total DNA and human DNA, but little research has focused on quantifying species-specific DNA, namely canine DNA (10). As canine DNA is readily present at crime scenes and may be mixed with human genetic evidence, the ability to detect and quantify canine genetic material in a mixture of non-target DNA is essential for forensic casework (11).

5.1 NON-CANINE SPECIFIC DNA QUANTITATION METHODS

Canine DNA quantitation has utilised non-specific DNA quantitation methods consistent with earlier techniques implemented by forensic laboratories that measure total DNA (10, 12). Early methods that were used to quantify DNA when species of origin was not of concern are UV and fluorescent spectroscopy (12). UV spectroscopy is considered a simple method that is most applicable to dsDNA by assessing optical density at 260 nm and 280 nm (12, 13). As UV spectroscopy methods require large amounts of sample, yields DNA in microgram quantities and suffers from absorption interference from RNA and contaminants, it is not suited for reliable quantitation of forensic samples (12-14). Dissimilar to UV spectroscopy, fluorescence spectroscopy is an alternate method for quantitation of total DNA that offers increased sensitivity of 25 pg/mL, permitting measurement of low-level DNA (12, 15). In order to quantitate DNA via fluorescence, the addition of an intercalating dye is necessary (12). The quantitation of DNA is measured by the mechanism of the intercalating dye binding to DNA to produce a fluorescence enhancement (15). Many fluorescent dyes have been utilised for DNA quantitation purposes, such as PicoGreen™ and ethidium bromide used for quantitation involving

solution-based or gel analysis (12). However, fluorescence spectroscopy measurements are relative and require comparison with quantified DNA standards or calculation via standard curve (12). Therefore, canine forensic samples can be quantified using UV or fluorescence spectroscopy; however, the quantitation measurement reflects the total DNA present in the sample, including possible human, bacteria, plant and animal DNA. Subsequently, a sensitive and canine-specific quantitation method is essential for optimising downstream processes.

5.2 CANINE-SPECIFIC DNA QUANTITATION METHODS

Studies relating to human DNA quantitation have demonstrated that qPCR methods provide more sensitive and accurate results compared to UV and fluorescent spectroscopy techniques (12, 13, 16, 17). qPCR assays utilise fluorescence to monitor the accumulation of PCR product with each cycle (18). Hence, by monitoring the data readout, the concentration of the sample can be determined by comparing the C_T value, cycle number correlating to exponential growth, to a series of standards (18). One of the most common types of qPCR utilises TaqMan®, a fluorogenic 5'nuclease assay, that during polymerisation releases the reporter dye from the probe and as the reporter dye and quencher dye are no longer in close proximity, fluorescence is detected (18). Additionally, the use of an intercalating dye, such as SYBR® Green I dye that detects the formation of PCR products by fluorescing when bound to dsDNA, offers an alternate approach to qPCR (19). Both TaqMan® and SYBR® Green I dye approaches have been applied to detect and quantify canine DNA to optimise downstream processes (10).

The TaqMan® assay developed for canine DNA quantitation is genus-specific and has been routinely utilised for in-house animal casework (20). The MC1R gene was chosen as the target for the TaqMan® qPCR assay as the MC1R sequence was conserved within canines and possessed enough variation to exclude non-target DNA (21). Additionally, as the MC1R gene is involved in hair colouration, the sequence data had been characterised and readily available (21). Alternatively, the SYBR® Green I dye was employed for SINE-based quantitation (22). As SINE-based qPCR assays have been designed for human DNA analysis, the high copy number of SINEs in the canine genome was considered an ideal target for species-specific qPCR assays (22). Both MC1R TaqMan® and SYBR® Green I dye qPCR assays can be used to dependably detect and quantitate canine DNA with sensitivity levels down to 5 pg of DNA with no effect from non-target species (10). As DNA is quantitated based on comparison to a standard, a calibrated reference standard is recommended for analysis that requires quantitation using an alternate technique, such as PicoGreen® (10). Furthermore, based on a study comparing MC1R TaqMan® and SYBR® Green I dye qPCR assays, it was found that the MC1R TaqMan® assay was more suited for forensic casework due to the shorter run time and potential toxicity risk of SYBR® Green I dye (10). The employment of qPCR as a technique for quantifying human DNA, has reduced laboratory costs, decreased sample consumption and increased turnaround time (16). Hence implementing qPCR, specifically the MC1R TaqMan® assay, for routine canine DNA quantitation can improve the efficacy and reliability of canine STR genotyping results.

6. CANINE FORENSIC GENETIC MARKERS

Genetic information obtained from canine biological material has a wide range of applications to forensic investigations (23, 24). Forensic genetic markers, including STR and

SNP genotyping and mtDNA sequence analysis, have been targeted to analyse canine-related evidence (24-28). Due to the similarity in genomes and inheritance patterns between humans and canines, the interpretation and analysis approach to canine genetic examinations are comparable to those methods involved in human DNA analysis (29). Consequently, the similarity has permitted an increase in application and development of canine DNA analysis in forensic cases (29).

6.1 CANINE mtDNA

The analysis of canine biological material has utilised mtDNA analysis in numerous criminal investigations, generally involving dog hair evidence (30). In the reported cases, canine mtDNA analysis has been implemented when STR typing fails due to inadequate quantities of DNA or degraded evidentiary samples (30). MtDNA has been widely used in animal typing and has been examined through sequencing, SNP assays or restriction digestion typing (11). Prior research has predominately focused on studying the noncoding HV1 and HV2 regions of canines and comparing the sequence to a consensus sequence for the examination of SNPs and insertions or deletions for association purposes (27). Similarly to humans, the canine HV1 and HV2 regions have demonstrated to be highly polymorphic and of forensic value in relation to degraded and limited DNA samples (27). However, the significance of canine mtDNA analysis is limited due to the domestication of dogs, resulting in fewer haplotypes and common haplotypes occurring at increased frequencies in comparison to human mtDNA examinations (30). Consequently, using data from the control region, the probability of a random match between two unrelated humans for approximately 720 bp was 2.7×10^{-8} in a Switzerland population (31) and the probability of a random match between two unrelated canines for approximately 595 bp was 7.0×10^{-2}

in a UK dog population (32). Therefore, the RMP is significantly higher in dogs than in humans, proving that canine mtDNA analysis is less discriminatory. Hence the forensic value of canine mtDNA may serve as a valuable tool for inclusionary and exclusionary purposes but is not suitable for the identification of an individual from canine biological samples (27).

6.2 CANINE STRs

The analysis of STRs has become the gold standard for modern DNA profiling (33). Microsatellites with dinucleotide repeats were first reported in domestic dogs in 1993 (34) and following the progression of human DNA analysis involving tetranucleotide markers, canine trinucleotide and tetranucleotide repeats were published in 1996 (35). Ideally, tetranucleotide repeat markers are preferential for forensic analysis due to the decrease in stutter percentages as well as increasing the resolution of closely spaced alleles, subsequently, easing the interpretation of genotyping results, consistent with human DNA profiling approaches (6). In contrast to humans, the effect of inbreeding in domestic animals generated concern regarding the degree of variation found in canine microsatellites (26). However, as further research was conducted, the abundant, highly polymorphic nature and ease of use of microsatellites demonstrated the sufficient application for identification and parentage verification, offering a high level of discrimination (36).

Numerous independent studies have investigated genotype and sequence data for various canine STR markers via singleplex amplification (35, 37, 38). Prior studies have validated the use of multiplex amplification by PCR for human forensic samples by

establishing dependable sensitivity and high discrimination whilst limiting sample consumption and being cost and time effective (39). Hence subsequent research has investigated the efficacy and advantages of the co-amplification of canine STR markers in a single PCR multiplex or multiple PCR multiplex reactions (3, 26, 38, 40-42). Additionally, combining various STR markers into multiplexes has enhanced the utility of canine microsatellites for forensic casework, leading to the development of canine STR panels and commercial canine STR kits (3).

6.3 CANINE STR PANELS

Prior to the availability of commercially available canine STR multiplex kits designed for forensic purposes, in-house assembled STR panels for canine genetic analyses were implemented for research and casework (9). The in-house panels that were developed and used for casework included 10 to 15 canine-specific STR markers that have been previously validated for sensitivity, species specificity and considered efficient for use in casework (4, 5, 40, 43). The selected STR markers within the in-house STR panels were chosen based on research conducted by Primmer et al. (44) and Francisco et al. (35), generally in conjunction with in-house validation testing. As a result, the STR markers included in the in-house panels consisted of four markers that were present in all panels, FH2010, FH2054, PEZ06, and PEZ12 (4, 5, 40, 43), as evident in Table 1. Hence the in-house panels differed based on the inclusion or exclusion of a combination of 19 varying markers (4, 5, 40, 43). As the in-house panels have been developed by the associated laboratory to analyse canine biological material for individual identification and parentage verification (4, 5, 40, 43), the lack of a uniform panel disallows the exchange and comparison of inter-laboratory data relating to analysed canine samples.

Table 1: Selected canine STR markers examined using in-house STR panels and DogFiler.

STR Marker	Shutler et al. (5)	Padar et al. (43)	Eichmann et al. (40)	Clarke et al. (4)	DogFiler (3, 45)
FH2010					
FH2054					
FH2079					
FH2087Ua					
FH2087Ub					
FH2132					
FH2611					
PEZ01					
PEZ02					
PEZ03					
PEZ05					
PEZ06					
PEZ08					
PEZ11					
PEZ12					
PEZ15					
PEZ16					
PEZ18					
PEZ20					
VWF.X					
WILMS-TF					
ZUBECA4					
ZUBECA6					
VGL0760					
VGL0910					
VGL1063					
VGL1165					
VGL1541					
VGL1606					
VGL1828					
VGL2009					
VGL2136					
VGL2409					
VGL2918					
VGL3008					
VGL3112					
VGL3235					
VGL3438					
SRY					

Recently, a study was conducted regarding the developmental validation of DogFiler, a STR panel assembled into a single multiplex containing 15 novel markers and a sex identification marker, SRY (3). As outline in Table 1, the STR markers contained in DogFiler are not present in other in-house panels, nor in previous published literature concerning canine DNA profiling. The newly selected STR markers were chosen based on data provided by the published 7.6X dog genome that met established in-house criteria, hence exclusive to DogFiler (3). The study demonstrated that the DogFiler STR panel is a useful tool for canine forensic identification and parentage verification (3). The DogFiler panel claims to produce optimum results with 0.5 – 1.5 ng of template DNA and based on in-house pooled data, the average RMP yielded was 7.81×10^{-16} with a conservative θ value of 0.142 (3). Additionally, Dog-Filer has been redesigned into mini-DogFiler containing mini-STRs that have undergone developmental validation for degraded canine biological samples(45). Although DogFiler and mini-DogFiler panels were validated in accordance with SWGDAM guidelines for developmental validation (45), the efficacy of the panels for forensic applications has not been published in relation to casework. Additionally, the lack of reliable evaluations regarding the selection of the novel STR markers in place of the STR markers included in the in-house panels, hinders the progression of inter-laboratory comparisons of STR data. Subsequently, the inclusion of incompatible STR markers in the in-house panels and DogFiler and mini-DogFiler panels impedes the standardisation of canine DNA analysis thus the demand for commercial canine STR kits (24).

6.4 COMMERCIAL CANINE STR KITS

Based on previous research and case findings, STR markers have successfully been employed for match comparisons and kinship testing in cases involving dog biological

material (4, 5, 36, 42, 43, 46). Having commercially available kits accessible to forensic laboratories has facilitated the ease and efficiency of forensic DNA analysis by ensuring that the primer design and optimised PCR multiplexes are suitable for the amplification and genotyping of the selected markers (47). As demonstrated by, but not limited to, the STR loci included in the in-house panels, many STR markers have been validated for individual identification and parentage verification of canines (4, 5, 40, 43). Similarly, the canine STR kits that have been manufactured for commercial purposes consist of varying combinations of numerous STR markers, as displayed in Table 2. As the available commercial kits vary based on which STR markers are included, the primer sequences utilised for PCR amplification will also vary due to the amplification of STR markers in a single PCR multiplex. Hence the commercial STR kits available for canine DNA analysis, dictate which STR markers will be used by the testing laboratory.

Table 2: Selected canine STR markers examined using commercial canine STR kits.

STR Marker	Repeat motif	Canine Genotypes™ Panel 1.1 (48)	Canine ISAG STR Parentage Kit (49)	StockMarks® (51)	Canine Genotypes™ Panel 2.1 (8)
AHT121	di				
AHT137	di				
AHTh130	di				
AHTh171	di				
AHTh260	di				
AHTk211	di				
AHTk253	di				
CXX279	di				
FH2848	di				
INRA21	di				
INU005	di				
INU030	di				
INU055	di				
REN105L03	di				
REN162C04	di				
REN169D01	di				
REN169O18	di				
REN247M23	di				
REN54P11	di				
REN64E19	di				
PEZ03	tri				
FH2001	tetra				
FH2004	tetra				
FH2010	tetra				
FH2017	tetra				
FH2054	tetra				
FH2079	tetra				
FH2088	tetra				
FH2107	tetra				
FH2309	tetra				
FH2328	tetra				
FH2361	tetra				
FH3313	tetra				
ZFX/Y	tetra				
PEZ01	tetra				
PEZ02	tetra				
PEZ05	tetra				

PEZ06	tetra				
PEZ08	tetra				
PEZ10	tetra				
PEZ11	tetra				
PEZ12	tetra				
PEZ13	tetra				
PEZ15	tetra				
PEZ16	tetra				
PEZ17	tetra				
PEZ20	tetra				
PEZ21	tetra				
FH3377	penta				
vWF.X	hexa				
Amelogenin	-				
ZFX/Y	-				

Note: The STR markers highlighted in dark green represent the STR markers examined by StockMarks® for Dogs Canine Genotyping Kit (StockMarks® for Dogs Canine 1 Kit) and light green represent the STR markers examined by StockMarks® for Dogs Canine 2 Kit. The Amelogenin and ZFX/Y STR markers are included in the corresponding kit for sex-typing purposes.

The Thermo Fisher Canine Genotypes™ Panel 1.1 and Thermo Fisher Canine ISAG Parentage Kit are primarily designed for routine parentage analysis and individual identification (48, 49). The Canine Genotypes™ Panel 1.1 contains reagents necessary to coamplify 18 STR markers, that include the core panel of loci recommended by the ISAG for canine parentage testing, in addition to amelogenin for sex-typing in a single multiplex PCR reaction (48). All STR markers in the Canine Genotypes™ Panel 1.1 contain dinucleotide repeat motifs with the exception of FH2054, a tetranucleotide repeat locus (48). Following the production of the Canine Genotypes™ Panel 1.1, the Canine ISAG

Parentage Kit was manufactured with an additional three markers, containing dinucleotide repeat motifs, hence coamplifies 21 STR loci and the amelogenin locus (49). As both the Canine Genotypes™ Panel 1.1 and Canine ISAG Parentage Kit includes a Master Mix containing optimized buffer, dNTPs and Phusion™ Hot Start DNA Polymerase, Primer Mix for the loci examined containing forward and reverse primers with a fluorescent dye end-labelled to one primer from each pair, a positive control DNA sample and requires 1–2 ng of a high quality DNA sample in a 20 µL reaction volume to deliver optimal results (48, 49). Therefore, the key difference between the two parentage kits is the additional three loci examined in the Canine ISAG Parentage Kit, as outlined in Table 2. Based on a study conducted by Kanthaswamy et al. (42), Canine Genotypes™ Panel 1.1 resulted in an average conservative RMP of 1.45×10^{-21} with a validated θ value of 0.09 whereas there is no published data concerning the average RMP yielded from the Thermo Fisher Canine ISAG Parentage Kit. The addition of AHTh130, REN105L03 and REN64E19 may decrease RMP values compared to those values associated with the Canine Genotypes™ Panel 1.1, subsequently increasing the appeal of the Canine ISAG Parentage Kit for forensic purposes (50). However, neither kit has been validated nor published literature relating to such use in forensic casework.

Similar to the Canine Genotypes™ Panel 1.1 and Canine ISAG Parentage Kit, the StockMarks® for Dogs Canine Genotyping Kit was also designed to genotype dogs for parentage verification (51). Formerly marketed as the StockMarks® for Dogs Canine 1 Kit, the StockMarks® for Dogs Canine Genotyping Kit consists of 10 STR markers recommended by the American Kennel Club (AKC), consisting of tetranucleotide repeat motifs with the exception of PEZ03, a trinucleotide repeat locus, outlined in Table 2 (26). The StockMarks®

for Dogs Canine 2 contained an additional 7 loci, comprised of tetranucleotide repeat motifs, hence allowing the amplification of 17 loci with two PCR multiplexes (26). The current StockMarks® for Dogs Canine Genotyping kit includes Amplification primer mix containing dye-labeled forward primers and reverse primers, AmpliTaq Gold® Polymerase, dNTP mix, StockMarks® PCR Buffer, $MgCL_2$ and control DNA. Unlike other commercial kits available for parentage canine STR analysis, the StockMarks® for Dogs Canine Genotyping Kit does not include a locus for sex determination and does not utilise Phusion™ Hot Start DNA Polymerase. However, the StockMarks® for Dogs Canine 1 Kit has been used in legal cases regarding the analysis of canine biological material (26). Owing to the State of Washington cautioning the use of canine DNA analysis as evidence in court, studies have been published that have sufficiently validated the microsatellites within the StockMarks® for Dogs Canine 1 and 2 Kits by demonstrating that the markers are reproducible, informative and robust for parentage and DNA identification testing (26, 30, 36). Based on a study concerning the verification of the panels, the StockMarks® for Dogs Canine 1 Kit yielded an average RMP of 3.6×10^{-5}, and 3.2×10^{-8} including the additional 7 loci in the Canine 2 panel with no evidence of an incorporated θ adjustment, significantly lower compared to the Fisher Canine Genotypes™ Panel 1.1 (36). The literature claims that the addition of the Canine 2 panel provides further power for identification appropriate for forensic investigations (26). However, as the Canine 2 panel is no longer available, the current StockMarks® for Dogs Canine Genotyping Kit with a required DNA sample concentration of 1 –10 ng/µL (51) and decreased power of exclusion with 10 STR markers analysed is more suitable for routine parentage verification opposed to samples involved in forensic casework.

The Canine Genotypes™ Panel 1.1, Canine ISAG Parentage kit and StockMarks® for Dogs Canine Genotyping kit have been designed to analyse canine DNA samples that contain high quantity and high quality DNA (26, 42). The efficacy of the parentage kits is considered dependable for routine DNA sample types for canine parentage testing, such as blood samples and buccal swabs (26, 42). However, in forensic casework, DNA evidentiary material is often degraded and limited in quantity. Therefore, a kit with applicable sensitivity and power of discrimination to reliably genotype canine DNA evidentiary material is necessary for forensic-related examinations. Additionally, developmental information regarding the validation of samples and databases employed by Canine Genotypes™ Panel 1.1, Canine ISAG Parentage Kit and StockMarks® for Dogs Canine Genotyping kit is absent or limited in published literature. Currently, the Canine Genotypes™ Panel 2.1 kit is the only commercially available dog STR kit specifically designed and validated for forensic analysis of canine material (42).

6.4.1 CANINE GENOTYPES™ PANEL 2.1 KIT

Comparable to human DNA analysis, where the development of PowerPlex 21 system has increased the efficiency, reliability and standardisation of human DNA profiling, the development of Canine Genotypes™ Panel 2.1 kit aims to standardise canine DNA analysis for forensic casework (7, 47). The generation of Canine Genotypes™ Panel 2.1 kit and accompanying validated U.S. population database was the first effort to assemble a canine STR typing kit in accordance with the SWGDAM guidelines (7). Stemming from developmental validation studies, the Canine Genotypes™ Panel 2.1 kit contains 18 polymorphic STR loci, comprised of tetranucleotide repeat motifs with the exception of FH3377 a pentameric marker and vWF.X a hexameric marker as displayed in Table 2, and

the ZFX/Y locus for sex-typing (8). The kit includes Canine Genotypes™ Panel 2.1 kit Master Mix containing dNTPs, Phusion™ Hot Start DNA Polymerase in an optimised buffer, Canine Genotypes™ Panel 2.1 Primer Mix consisting of forward and reverse primers with a fluorescent dye end-labeled to one primer from each pair for the loci examined and Canine Genotypes ™Panel 2.1 Control DNA for PCR and electrophoresis verification (8). Similar to Canine Genotypes™ Panel 1.1 and StockMarks® for Dogs Canine Genotyping kit, Canine Genotypes™ Panel 2.1 kit requires 1–2 ng of a high-quality DNA sample in a 20 µL reaction volume to deliver optimal results (8). Unlike commercial human DNA forensic kits, the Canine Genotypes™ Panel 2.1 kit does not include an allelic ladder. The developmental validation study recommends using the canine control DNA included in the kit for calibration of observed allele sizes (7). Based on reproducibility studies concerning the analysis of the canine control DNA, the level of precision and accuracy obtained suggests that there is no need for an allelic ladder (42). Nevertheless, an allelic ladder would reduce the effect of migration variation between examinations and enhance the comparison of canine STR data between laboratories (42). As the Canine Genotypes™ Panel 2.1 kit is the only commercially available kit specifically designed for forensic analysis of canine biological material, it has undergone an independent developmental validation study, ensuring the kit's efficacy as a forensic canine genotyping system (42).

The development of Canine Genotypes™ Panel 2.1 kit utilised markers from previously established canine STR panels and canine STR kits. As a result, six of the loci included in the StockMarks® for Dogs Canine Genotyping Kit were included in the development of the Canine Genotypes™ Panel 2.1 kit, evident in Table 2 (7). The combination of the 18 STR proved to be of high genetic diversity regardless of breed or

geographic region (7). This was established from the population studies on the accompanying database, constructed from allele diversity and frequency distribution data of U.S. domestic dog populations regarding the examined 18 STR markers (7). Therefore, the allele frequencies obtained from the 18 polymorphic STR loci can be used to reliably generate RMPs due to the high level of genetic diversity among the selected markers. The conservative RMP of a hypothetical canine STR profile was estimated as 2.26×10^{-40} with a θ value of 0.09 based on the national U.S. population database (7). Hence the STR markers included in the Canine Genotypes™ Panel 2.1 kit can reliably quantify the evidential value of a match between a reference sample and crime scene sample with a high level of confidence.

The developmental validation study also investigated the efficiency of Canine Genotypes™ Panel 2.1 kit in regards to the production and interpretation of canine STR profiles (24). A study concerning the comparison of Canine Genotypes™ Panel 1.1 kit and Canine Genotypes™ Panel 2.1 kit established that the Canine Genotypes™ Panel 2.1 kit containing STR markers that contain predominantly tetranucleotide repeat motifs, increased processing and interpretation efficiency, hence is more capable of identifying individual dogs (42). Moreover, the amplification study established that an annealing temperature of 60°C and 30 cycles for amplifying canine DNA produced the best intercolour balance, called all peaks and appropriate signal intensity was observed for all peaks (24). This data coincides with the manufacturer's recommendations for amplification of canine template DNA (8). Furthermore, due to the 3' to 5' exonucleases activity of the Phusion™ Hot Start DNA Polymerase, no plus-A or shouldering from non-templated nucleotide addition was present in any of the amplification results, hence increasing the

validity of interpretation outcomes (8, 24). Regarding STR profile interpretation, several dye blobs were evident in samples containing peaks with low RFU values and reagent blanks at the following locations:

- ~123 bp in the blue dye channel (7).
- ~117 bp in the green dye channel (7).
- ~96 bp in the yellow dye channel (7).
- ~112 and ~118 bp in the red dye channel (7).

In addition to the identified artifacts, for the animals typed in the study, microvariant alleles were seen at loci FH2361, FH3313, FH2107, FH2309 and FH3377 (7). Furthermore, three peaks were observed at loci FH2328 for the positive canine control DNA and was determined to be a tri-allelic pattern linked with the cell line of the control DNA (7). Hence for interpretation purposes, the identification of known artifacts, microvariants and tri-allelic pattern can limit complication in allele calling and facilitate the interpretation process.

The development of the canine STR nomenclature relating to the 18 STR markers included in Canine Genotypes^{TM} Panel 2.1 kit permits inter-laboratory data comparisons by increasing fragment sizing accuracy (7). The nomenclature method used to characterise and sequence the STR markers has been internationally accepted and was designed to dedicate STR alleles based on the number of full repeat motifs, followed by separation with a decimal, the number of nucleotides in the incomplete repeat motifs or microvariants (38, 40). As a result of the nomenclature development, the core repeat units of the STR loci were described and STR markers were classified as simple, compound or complex repeat structures (7). The proposed canine allelic nomenclature promotes inter-laboratory sharing

of STR typing results and increases the opportunity for combining population data hence furthering the progression of canine DNA analysis. In summary, the Canine Genotypes™ Panel 2.1 kit serves as a valuable tool for forensic casework involving identity testing and parentage verification based on the reliability and informativeness of the STR panel (7).

7. CANINE POPULATION DATABASES

To use canine STR genetic analysis effectively for forensic casework purposes, access to a relevant canine population database is imperative (26). The frequency for a canine STR profile can be calculated based on information concerning the alleles present in the profile, population frequencies of the alleles and appropriate genetic formulas incorporating correction factors for population substructure and degree of relatedness (52). It is crucial for reliable results that the database used for canine STR analysis is large enough and representative of the population from which the sample originated (52). Therefore, by calculating the frequency of the observed STR genotype in a population database, the canine DNA evidence can be statistically weighted, comparable to human DNA profiling evidence admitted in courts of law (52).

7.1 CANINE ALLELE FREQUENCY DATA

Data sets including STR allele frequencies have been published for domestic dogs and have been used to calculate RMP and LR for canine DNA evidence (3, 4, 9, 26). Prior to the establishment of the canine database developed for the validation of Canine Genotypes™ Panel 2.1 kit, private databases were assembled based on in-house panels or kits that were designed for routine parentage verification (3, 4, 9, 26). A study authorised by the AKC investigated the efficacy of the StockMarks® for Dogs Canine 1 Kit for routine

parentage testing by the generation of the AKC data set (36). 9561 samples were collected at Parent Breed Club National Specialty dog shows, representing 108 dog breeds (36). The Zoogen database includes STR allele frequencies obtained from the analysis 489 pure breed and 69 mixed breed canines using the StockMarks® for Dogs Canine 1 Kit, similar to the AKC data set (26). As the AKC data set and Zoogen database are predominated by pure breed dogs, it is likely that some genetic lines may be overrepresented, hence not adequately representing the actual allele frequencies of the U.S. canine population. Therefore, the AKC data set and Zoogen database are not suitable for the statistical weighting of canine DNA evidence. Ideally, a canine population database should include allele frequencies from a range of domestic dogs, including pure breed, mixed breed and hybrids of two or more breeds (7). Based on this data, a U.S. canine database was created using the Canine Genotypes™ Panel 2.1 kit consisting of 236 pure breed dogs representing nine breeds and 431 mixed breed dogs (7). The Canine Genotypes™ 2.1 U.S. canine database contains population data concerning the analysis of 18 STR markers that can be typed using the kit (9). The database includes data relating to locus informativeness, allele frequencies and inbreeding coefficients (9). This database is the only published canine database utilising a commercial kit specifically designed for forensic purposes.

Generally, the match of two DNA profiles, such as the DNA profile obtained from a reference sample matching the DNA profile obtained from a crime scene sample, involves the use of the product rule by analysing relative allele frequencies in a database to obtain profile match probabilities (53). Therefore, in order to generate a valid population database, the frequency of alleles present in the population must be established (52). The Canine Genotypes™ Panel 2.1 U.S. database achieved this by sampling 676 U.S. domestic dogs containing individual genotypic information for all 18 STR loci examined (7). The

national database is also geographically representative of four regions across the U.S. to assess genetic diversity concerning geographic location (7). The allelic information contained in the published genotype database includes the number, range and frequencies for regional and national sample sets (7). Therefore, obtaining allele frequency information permits the storage of data and enables the application of genetic formulas to ensure the validity of the database for forensic purposes (52).

7.2 VALIDATION OF CANINE DATABASES

As dependable estimations of allele frequencies in relevant populations are required to reliably determine the probability of a random match between two DNA profiles, the database utilised is required to be examined for Hardy-Weinberg Equilibrium (HWE) and Lineage Equilibrium (LE) (52). The primary purpose of HWE and LE tests is to assess the genetic independence of alleles within loci and between loci, respectively, to permit the use of the product rule for RMP calculations (52). As the product rule should only be applied when there is no significant deviation from HWE and LE, if the relevant population is not in HWE and LE, the departure must be accounted for when calculating the RMP statistic (52). Hence, as domestic dogs have known selection and inbreeding histories, a correction factor, θ, can be applied to estimates regarding the frequency of a canine DNA profile to adjust for population substructure (52).

The Canine Genotypes^{TM} Panel 2.1 U.S. database found that mixed breed samples produced an increased number of pairs out of LE compared to pure breed dogs, supporting the results of the Zoogen database study (9, 26). Nevertheless, the data obtained from the Canine Genotypes^{TM} Panel 2.1 U.S. database exemplifies that the allele frequencies across

the 18 STR loci can be multiplied together to calculate RMPs (7). On the other hand, the Canine Genotypes™ Panel 2.1 U.S. database analyses for HWE indicated that the population was not in HWE (7). Hence, the indication of population substructure required appropriate adjustments for the degree of HWD to be applied to statistical calculations (52). Therefore, a θ correction value of 0.09 was incorporated into RMP calculations to compensate for potentially underestimating the allele frequencies in the U.S. domestic dog population (7, 52). Overall, the validation study concerning the Canine Genotypes™ Panel 2.1 U.S. database yielded a high level of gene diversity regardless of pure breed, mixed breed or geographical region for the combined 18 STR markers analysed and strengthened the validity of canine DNA evidence (9).

8. EXPERIMENTAL DESIGN

Currently, in Western Australia, there has been no concerted effort to generate a forensically validated canine microsatellite database. Therefore, the creation of the population database will contain accessible allele frequency data from pure breed and mixed breed dogs from Western Australia. Furthermore, the generation of the database will strengthen the validity of canine biological evidence for use in forensic casework. The database will be constructed using the proposed methodology listed below:

I. Sample collection

A total of 100 DNA samples from purebred and mixed bred dogs will be collected from historically stored canid EDTA blood and labelled with barcodes.

II. DNA extraction

DNA will be extracted using the PrepFiler™ kit (manual) according to the manufacturer's specifications (54).

III. DNA quantitation

DNA will be quantitated using an MC1R TaqMan® assay in accordance with the TaqMan® Fast Advances Master Mix manufacturer's guidelines (55) with the QuantStudio™ 6 Flex Real-Time PCR System as per the user guide (56).

IV. DNA profiling

The Canine Genotypes™ Panel 2.1 Kit will be used for genotyping the 100 dog samples in accordance with the manufacturer's specifications (8). PCR will be performed using the ProFLEX™ PCR System and electrophoresis will be performed using the ABI PRISM® 3130*x*l Genetic following the relevant user guides (57, 58).

V. Data analysis

The data obtained from the Canine Genotypes™ Panel 2.1 Kit will be interpreted using GeneMarker® as per the user manual (59).

VI. Database generation

The database will be generated from 100 canine microsatellite profiles using Microsoft Excel and the associated allele frequencies for each STR marker will be calculated and analysed. A computer program will then be used to perform statistical tests on the generated genetic data.

9. EXPERIMENTAL AIMS AND OBJECTIVES

This research project aims to sample 100 Western Australian domestic dogs to generate a canine microsatellite database for crime and paternity testing. The generation of the Western Australian canine database will permit statistical weighting to be applied to matches between dog reference samples and crime scene samples, hence allowing reliable interpretation of canine DNA profiles for forensic casework. The objectives of this research project are outlined below:

I. Develop a workable method to obtain a canine DNA profile.

II. Establish a viable quantification technique to optimise downstream processes.

III. Establish a validated database of 100 canine microsatellite profiles containing population allele frequencies.

Therefore, it is hypothesised that using PrepFiler™ kit for extraction, an MC1R TaqMan® qPCR assay for quantitation and the Canine Genotypes™ Panel 2.1 Kit for genotyping, a 100 dog microsatellite database can be generated from historically collected canid EDTA blood.

10. CONCLUSION

Dog-derived biological evidence is commonly present at crime scenes and subsequent analysis can aid forensic investigations. The utility of canine DNA evidence has demonstrated the demand for a viable working method and standardisation of analysis techniques. To improve the efficiency and reliability of canine evidentiary examinations, the MC1R TaqMan® assay is recommended for routine canine DNA quantitation. Furthermore, while there has been extensive research concerning the efficacy of STR markers for canine identification and paternity verification, the Canine Genotypes™ Panel 2.1 kit is the only commercial kit available designed specifically for forensic purposes. Hence establishing a standard protocol for the analysis of canine DNA, a population database can be reliably generated. As validated databases enable the statistical significance of canine DNA evidence to be reliably calculated, the STR genotyping results can be implemented into forensic casework. Hence the generation of a Western Australia canine microsatellite database has potential to be a valuable national resource and further the application and development of canine DNA analysis in the Australian forensic science community.

11. REFERENCE LIST

1. Animal Medicines Australia. Pet Ownership in Australia: Animal Medicines Australia Pty Ltd; 2016.

2. Rajshekar M, Blizzard L, Julian R, Williams A, Tennant M, Forrest A, et al. The incidence of public sector hospitalisations due to dog bites in Australia 2001–2013. Australian and New Zealand Journal of Public Health. 2014;41(4):377-80.

3. Wictum E, Kun T, Lindquist C, Malvick J, Vankan D, Sacks B. Developmental validation of DogFiler, a novel multiplex for canine DNA profiling in forensic casework. Forensic Sci Int Genet. 2013;7(1):82-91.

4. Clarke M, Vandenberg N. Dog attack: the application of canine DNA profiling in forensic casework. Forensic Sci Med Pathol. 2010;6(3):151-7.

5. Shutler G, Gagnon P, Verret G, Kalyn H, Korkosh S, Johnston E, Halverson J. Removal of a PCR inhibitor and resolution of DNA STR types in mixed human-canine stains from a five year old case. Journal of forensic sciences. 1999;44(3):623.

6. Butler J. Fundamentals of Forensic DNA Typing. Burlington: Elsevier Inc.; 2010.

7. Kanthaswamy S. Development and Validation of a Standardized Canine STR Panel for Use in Forensic Casework. National Institute of Justice (NIJ) National Criminal Justice Reference Service; 2009.

8. Thermo Scientific. Technical Manual: Canine Genotypes Panel 2.1: Thermo Fisher Scientific Inc.; 2014.

9. Kanthaswamy S, Tom BK, Mattila AM, Johnston E, Dayton M, Kinaga J, et al. Canine population data generated from a multiplex STR kit for use in forensic casework. J Forensic Sci. 2009;54(4):829-40.

10. Evans JJ, Wictum E, Penedo MCT, Kanthaswamy S. Real-Time Polymerase Chain Reaction Quantification of Canine DNA. Journal of Forensic Sciences. 2007;52(1):93-6.

11. Kanthaswamy S. Review: domestic animal forensic genetics - biological evidence, genetic markers, analytical approaches and challenges. Anim Genet. 2015;46(5):473-84.

12. Nicklas JA, Buel E. Quantification of DNA in forensic samples. Analytical and bioanalytical chemistry. 2003;376(8):1160-7.

13. Lee SB, McCord B, Buel E. Advances in forensic DNA quantification: A review. Electrophoresis. 2014;35(21-22):3044-52.

14. Georgiou CD, Papapostolou I. Assay for the quantification of intact/fragmented genomic DNA. Analytical Biochemistry. 2006;358(2):247-56.

15. Singer VL, Jones L, Yue ST, Haugland RP. Characterization of PicoGreen Reagent and Development of a Fluorescence-Based Solution Assay for Double-Stranded DNA Quantitation. Analytical Biochemistry. 1997;249(2):228-38.

16. Kline MC, Duewer D, Redman JW, Butler JM. Results from the NIST 2004 DNA Quantitation Study. Journal of forensic sciences. 2005;50(3):570.

17. Richard ML, Frappier R, Newman JC. Developmental validation of a real-time quantitative PCR assay for automated quantification of human DNA. Journal of forensic sciences. 2003;48(5):1041.

18. Schmittgen TD. Real-Time Quantitative PCR Methods. 2001;25(4):383-5.

19. Karsai A, Muller S, Platz S, Hauser M. Evaluation of a homemade SYBR green I reaction mixture for real-time PCR quantification of gene expression. BioTechniques. 2002;32(4):790-6.

20. Lindquist CD, Evans J, Wictum EJ. Developmental Validation of Feline, Bovine, Equine, and Cervid Quantitative PCR Assays. Journal of forensic sciences. 2011;56(s1):S29-35.

21. Newton JM, Wilkie A, He L, Jordan SA, Metallinos DL, Holmes NG, et al. Melanocortin 1 receptor variation in the domestic dog. Mammalian Genome. 2000;11(1):24-30.

22. Walker JA, Hughes D, Anders BA, Shewale J, Sinha SK, Batzer MA. Quantitative intra-short interspersed element PCR for species-specific DNA identification. Analytical Biochemistry. 2003;316(2):259-69.

23. Graham EAM. Nonhuman DNA. Forensic Science, Medicine, and Pathology. 2005;1(2):159-61.

24. Dayton M, Koskinen M, Tom BK, Mattila A, Johnston E, Halverson J, et al. Developmental validation of short tandem repeat reagent kit for forensic DNA profiling of canine biological material. Croatian Medical Journal. 2009;50(3):268-85.

25. Baute DT, Satkoski JA, Spear TF, Smith DG, Dayton MR, Malladi VS, et al. Analysis of forensic SNPs in the canine mtDNA HV1 mutational hotspot region. J Forensic Sci. 2008;53(6):1325-33.

26. Halverson J, Basten C. A PCR Multiplex and Database for Forensic DNA Identification of Dogs. J Forensic Sci. 2005;50(2):352-63.

27. Himmelberger AL, Spear TF, Satkoski JA, George DA, Garnica WT, Malladi VS, et al. Forensic utility of the mitochondrial hypervariable region 1 of domestic dogs, in conjunction with breed and geographic information. J Forensic Sci. 2008;53(1):81-9.

28. Webb KM, Allard MW. Mitochondrial genome DNA analysis of the domestic dog: identifying informative SNPs outside of the control region. J Forensic Sci. 2009;54(2):275-88.

29. Iyengar A, Hadi S. Use of non-human DNA analysis in forensic science: a mini review. Med Sci Law. 2014;54(1):41-50.

30. Halverson JL, Basten C. Forensic DNA Identification of Animal-derived Trace Evidence- Tools for Linking Victims and Suspects. Croatian Medical Journal. 2005;46(4):598-605.

31. Dimo-Simonin N, Grange F, Taroni F, Brandt-Casadevall C, Mangin P. Forensic evaluation of mtDNA in a population from south west Switzerland. International Journal of Legal Medicine. 2000;113(2):89-97.

32. Wetton JH, Higgs JE, Spriggs AC, Roney CA, Tsang CSF, Foster AP. Mitochondrial profiling of dog hairs. Forensic Science International. 2003;133(3):235-41.

33. Lynch M. God's signature: DNA profiling, the new gold standard in forensic science. Endeavour. 2003;27(2):93-7.

34. Ostrander EA, Sprague G, Rine J. Identification and characterisation of dinucleotide repeat (CA)n markers for genetic mapping in dog. Genomics. 1993;16(1):207-13.

35. Francisco LV, Langston A, Mellersh CS, Neal CL, Ostrander EA. A class of highly polymorphic tetranucleotide repeats for canine genetic mapping. Mammalian Genome. 1996;7(5):359-62.

36. DeNise S, Johnston E, Halverson J, Marshall K, Rosenfeld D, McKenna S, et al. Power of exclusion for parentage verification and probability of match for identity in

American kennel club breeds using 17 canine microsatellite markers. Animal Genetics. 2004;35(1):14-7.

37. Halverson J, Dvorak J, Stevenson T, inventor; The Perkin-Elmer Corporation, assignee. Microsatellite sequences for canine genotyping. US Patent patent 05874217. 1999.

38. Hellmann AP, Rohleder U, Eichmann C, Pfeiffer I, Parson W, Schleenbecker U. A proposal for standardization in forensic canine DNA typing: allele nomenclature of six canine-specific STR loci. J Forensic Sci. 2006;51(2):274-81.

39. Moretti TR, Baumstark A, Defenbaugh DA, Keys KM, Smerick JB, Budowle B. Validation of short tandem repeats (STRs) for forensic usage: Performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples. Journal of Forensic Sciences. 2001;46(3):647-60.

40. Eichmann C, Berger B, Steinlechner M, Parson W. Estimating the probability of identity in a random dog population using 15 highly polymorphic canine STR markers. Forensic Sci Int. 2005;151(1):37-44.

41. Berger B, Berger C, Hecht W, Hellmann A, Rohleder U, Schleenbecker U, et al. Validation of two canine STR multiplex-assays following the ISFG recommendations for non-human DNA analysis. Forensic Sci Int Genet. 2014;8(1):90-100.

42. Kanthaswamy S, Oldt RF, Montes M, Falak A. Comparing two commercial domestic dog (Canis familiaris) STR genotyping kits for forensic identity calculations in a mixed-breed dog population sample. Anim Genet. 2019;50(1):105-11.

43. Pádár Z, Angyal M, Egyed B, Füredi S, Woller J, Zöldág L, et al. Canine microsatellite polymorphisms as the resolution of an illegal animal death case in a

Hungarian Zoological Gardens. International Journal of Legal Medicine. 2001;115(2):79-81.

44. Primmer CR, Matthews M. Canine tetranucleotide repeat polymorphism at the VIAS-D10 locus. Animal Genetics. 1993;24(4):332.

45. Kun T, Lyons LA, Sacks BN, Ballard RE, Lindquist C, Wictum EJ. Developmental validation of Mini-DogFiler for degraded canine DNA. Forensic Science International: Genetics. 2013;7(1):151-8.

46. Zajc I, Mellersh C, Kelly EP, Sampson J. A new method of paternity testing for dogs, based on microsatellite sequences. The Veterinary record. 1994;135(23):545.

47. Ensenberger MG, Hill CR, McLaren RS, Sprecher CJ, Storts DR. Developmental validation of the PowerPlex 21 System. Forensic Sci Int Genet. 2014;9:169-78.

48. Thermo Scientific. User Guide: Canine Genotypes Panel 1.1 Thermo Fisher Scientific Inc.; 2012.

49. appliedbiosystems. User Guide: Canine ISAG STR Parentage Kit (2014): Thermo Fisher Scientific; 2017.

50. Ng J, Oldt R, Kanthaswamy S. Assessing the FBI's Native American STR database for random match probability calculations. Legal medicine (Tokyo, Japan). 2018;30:52-5.

51. appliedbiosystems. User Guide: StockMarks Kits for Horses, Cattle, and Dogs: Thermo Fisher Scientific Inc.; 2014.

52. Butler J. Advanced Topics in Forensic DNA Typing: Interpretation. Oxford: Elsevier Inc.; 2015.

53. Balding DJ, Nichols R. DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands. . Forensic Sci Int. 1994;64(2-3):125-40.

54. Thermo Scientific. PrepFiler Forensic DNA Extraction Kit: Applied Biosystems Inc.; 2008.

55. appliedbiosystems. User guide: TaqMan Fast Advanced Master Mix: Thermo Fisher Scientific Inc.; 2017.

56. appliedbiosystems. User guide: QuantStudio 6 and 7 Flex Real-Time PCR System Software: Life Technologies Corporation; 2013.

57. appliedbiosystems. User guide: ProFlex PCR System: Thermo Fisher Scientific Inc.; 2016.

58. appliedbiosystems. User guide: Applied Biosystems 3130/3130xl Genetic Analyzers: Life Technology Corporation; 2012.

59. SoftGenetics. User manual: GeneMarker: SoftGenetics LLC.; 2018.

BLANK PAGE – NOT NUMBERED

PART TWO

MANUSCRIPT

VALIDATION AND DATABASE GENERATION OF 100 CANINE MICROSATELLITE PROFILES FOR CRIME AND PATERNITY TESTING

1. ABSTRACT

Dogs are kept as pets in almost two in five Australian households. Subsequently, canine biological evidence often forms part of the physical evidence found at crime scenes. The genetic analysis of canine biological material has provided valuable intelligence to forensic investigations. However, as canine DNA analysis has been underutilised in Australian forensic casework, there is a lack of standardisation concerning methodology. Additionally, there has been no effort to develop an Australian canine microsatellite database using the Canine Genotypes™ Panel 2.1 Kit for the statistical weighting of evidence. Thus, this study aimed to generate a Western Australian domestic dog database and establish a viable working method that can be implemented into routine canine DNA analysis. By using the PrepFiler™ kit for extraction, the MC1R TaqMan® assay with validated standards for quantitation and the Canine Genotypes™ 2.1 Panel Kit for genotyping, a workable method was established to obtain a canine DNA profile. The resulting Western Australian domestic dog database contains allele frequency data for 14 out of the 18 STR loci, proving that further investigation is required regarding the four loci that did not produce interpretable data. However, the data presented in this study has the potential to form a valuable national resource and kick-start the application of canine DNA analysis in Australian forensic casework.

KEYWORDS: Forensic Science, Canine DNA, Canine quantitation, Canine genotyping, Canine database

TABLE OF CONTENTS

1. LIST OF TABLES

2. LIST OF FIGURES

3. LIST OF ABBREVIATIONS

bp	Base pair
C_T	Cycle threshold
DNA	Deoxyribonucleic acid
LR	Likelihood ratio
MC1R	Melanocortin-1 receptor
ng	Nanogram
pg	Picogram
qPCR	Quantitative real-time PCR
RFU	Relative fluorescence unit
RMP	Random match probability
STR	Short tandem repeat
µL	Microlitre
θ	theta (co-ancestry co-efficient)

5. INTRODUCTION

5.1 Forensic application of canine biological material

As of 2016, an estimated 4.8 million dogs were owned as pets in Australia (1). With more than 100 000 people suffering dog-related injuries per year, the ability to identify the individual dog responsible is critical for forensic investigations (2). The application of canine DNA analysis is not limited to cases involving dogs as the perpetrators, but also include instances of property damage, animal cruelty, theft of animals, paternity disputes and destruction to native fauna and livestock (3). As domestic canine material is often encountered at crime scenes, in the form of hair, saliva, blood, urine and faeces, the subsequent recovery and genetic analysis of dog-related exhibits can provide links between suspects, victims and crime scenes (4, 5).

The analysis of microsatellites or STR loci has become the gold standard for DNA profiling (6). As a result, the techniques employed by forensic laboratories to recover and analyse samples containing human biological material are considered well established, sensitive and highly discriminative (6). Similar to human DNA analysis, the examination of canine genetic material involves the examination of many well-studied STR markers that have been validated through prior research and casework (7). Subsequently, the working method and associated principles implemented for human STR profiling are applicable to the examination of canine biological material (3).

5.2 Canine-specific quantitation technique

The quantitation of target DNA extracted from forensic samples is crucial for reliable interpretation of DNA profiles and minimising sample consumption and accompanying costs (6). Due to commercial STR genotyping kits functioning best with a narrow concentration ranges of target DNA, an assay that can detect and quantify canine DNA in a mixture of non-target DNA is essential for forensic casework (8). As the Canine Genotypes™ Panel 2.1 Kit provides an optimal concentration window of 1 – 2 ng/µL of target DNA, accurate quantitation is essential for the effective analysis of canine STR profiles (9, 10). It has been demonstrated in human DNA analysis techniques, that qPCR offers sensitive, accurate and reproducible quantitation results with validated standards being critical for precise analysis of run data (11). Dissimilar to human DNA quantitation techniques, commercial canine-specific quantitation products are not available. As reported in Evans et al. (12), a canine-specific MC1R TaqMan® assay has been developed and validated for forensic samples by promoting successful and efficient canine-specific STR genotyping. Hence by implementing the primer and probe design of the MC1R TaqMan® assay, established by Evans et al. (12), in conjunction with the development of validated canine DNA standards, a viable quantitation technique can be implemented into a standard working method for routine canine DNA analysis.

5.3 Canine STR loci and population databases

Despite the potential for providing intelligence to police investigations, the analysis of canine biological material is not readily utilised in forensic casework, specifically in Australia (3, 13). Therefore, unlike the validated approaches to human DNA profiling, there is a lack of standardisation concerning methodology, interpretation and reporting with

regards to canine biological material. Furthermore, in the U.S., recent court challenges regarding canine DNA evidence dictated the need for a standardised and validated canine STR panel (13). As a response, the Canine Genotypes™ Panel 2.1 Kit was manufactured specifically for forensic analysis and demonstrated, through developmental studies, high allelic diversity at the 19 STR loci allowing reliable identification of individual dogs (13). The generation of the Canine Genotypes™ Panel 2.1 Kit and accompanying U.S. population database was the first effort to assemble a canine STR typing system validated explicitly for use in forensic casework (13).

To use canine STR data in forensic casework, access to a relevant canine population database is imperative for reliable statistical analysis (7). In addition, it is crucial to withstand legal scrutiny that the canine population database is large enough and representative of the population from which the sample originated (14). Currently, there has been no concerted effort to establish a canine population database using the Canine Genotypes™ Panel 2.1 Kit in Australia. The generation and validation of a Western Australian canine microsatellite database would enable the statistical weighting of a canine DNA match to be reliably calculated for use in Western Australian forensic casework. Furthermore, the creation of a Western Australian canine microsatellite database provides the framework necessary for promoting the sharing of national canine STR typing results and enhance the progression of canine DNA analysis in Australia. This study aims to develop a workable method to obtain a canine DNA profile and sample 100 Western Australian domestic dogs to generate a canine microsatellite database for crime and paternity testing.

6. MATERIALS AND METHODS

6.1 Sample collection

A total of 100 canine blood samples and four liver samples were provided by the Animal Hospital at Murdoch University with breed and sex information supplied by the Animal Hospital.

6.2 DNA Extraction

DNA extraction for the 100 canine blood samples was performed with the reagents from the PrepFiler™ automated extraction kit as per the PrepFiler™ Manual extraction kit protocol (15). The following adjustments to the prescribed method were made, 60 uL of magnetic particles were added in the binding step, and in the first wash 300 µL of Wash Buffer A was added.

DNA extraction for the liver samples was performed in accordance with the salting out procedure described in Miller et al., (16) with 200 µL of digestion buffer containing 20mM EDTA pH 8.0, 50mM Tris, 120mM NaCl and 1% SDS and 10 µL 20 mg/mL Proteinase K.

6.3 DNA Quantitation

6.3.1 Development of custom MC1R TaqMan® Assay

The MC1R TaqMan® Assay was designed according to the primer and probe design outlined in Evans et al., (12) including the recommendation to move the forward primer forward 8 bases and were supplied by Thermo Fisher Scientific.

6.3.2 Validation of canine qPCR standards

The canine liver DNA extracts were combined in equal parts and serially diluted in nuclease-free water *(ten-fold dilution series: neat, 10×, 100×, 1000×, 10000×, 100000× and 1000000×)* and quantitated using the Qubit® 3.0 Fluorometer with 198 µL Qubit® Working Solution and 2 µL of combined canine liver DNA extract, as per the user manual (17).

Real-time qPCR analysis was performed using the custom MC1R TaqMan® Assay in conjunction with TaqMan® Fast Advanced Master Mix using a Quant Studio™ 6 Flex Real-Time PCR System (12, 18, 19). The serially diluted canine liver extracts and Control DNA (0.25 ng/µL), included in the Canine Genotypes Panel 2.1 Kit, were subjected to qPCR using 1.0 µL of the custom MC1R TaqMan® Assay, 10.0 µL of the TaqMan® Fast Advanced Master Mix and 7.0 uL of Nuclease-Free Water, as described in the TaqMan® Fast Advanced Master Mix User Guide (19). The Qubit® data from the *neat* combined canine DNA extract was inputted as the concentration for Std. 1 (39.35 ng/µL). The serial dilutions were inputted as standards in duplicates with the neat, 10×, 100×, 1000×, 10000×, 100000× and 1000000× dilutions entered as Std. 1, Std. 2, Std. 3, Std. 4, Std. 5, Std. 6 and Std. 7, respectively, for the MC1R target, for standard curve generation.

6.3.3 qPCR of canine blood samples

The canine blood sample DNA extracts (2.0 µL) were subjected to qPCR with a Quant Studio™ 6 Flex Real-Time PCR System using 1.0 µL of the custom MC1R TaqMan® Assay, 10.0 uL of the TaqMan® Fast Advanced Master Mix and 7.0 µL of Nuclease-Free Water, as described in the TaqMan® Fast Advanced Master Mix User Guide (19). Canine DNA Std. 1 – 5 were inputted in duplicates for the MC1R target, for standard curve generation.

6.4 PCR Amplification

The 100 canine blood sample DNA extracts were amplified with the Canine Genotypes™ Panel 2.1 Kit, manufactured by Thermo Fisher Scientific. PCR was performed with 2 µL of template DNA (1.0 ng/uL) in a total PCR volume of 20 µL. Amplifications were performed as specified in the Canine Genotypes 2.1 Technical Manual with the ProFLEX™ PCR System, with thermal cycling parameters of 98ºC for 3 minutes; 30 cycles at 98ºC for 15 seconds; 60 ºC for 75 seconds; 72ºC for 30 seconds and 72ºC for 5 minutes (10).

6.5 Capillary electrophoresis and fragment analysis

For each sample, 1 uL of PCR product and 10 uL of reaction mix were used for capillary electrophoresis with the ABI PRISM® 3130*xl* genetic analyser according to the instructions published in the Canine Genotypes Panel 2.1 Technical Manual and the ABI PRISM® user instructions (10, 20). The run module was based on the FragmentAnalysis36_POP4_1 created by Populator using the following values for injection in combination with 36cm capillaries, Inj. Secs: 12, Inj. kV: 1.2, Run. kV: 15.0, Run ºC: 60 and Run Time: 1500 s.

6.6 Canine database generation

For the database generation, the electropherograms were analysed with Gene Marker Software v1.95. The panel file and bin file were supplied by the manufacturer and adjusted based on the fragment analysis data of the 100 canine samples. Possible microvariants which appeared to be one or two base pairs apart were binned with the immediate adjacent allele by increasing the size range around the respective allele. The genotyping data for each sample was inputted into Microsoft Excel and allele frequencies,

RMP and LR was calculated in accordance with Balding, et al., (1994) and recommendations 4.10a and 4.10b outlined by NRC II, (1996) (21, 22).

7. RESULTS AND DISCUSSION

7.1 Development of a viable canine qPCR technique

The canine liver samples supplied by the Animal Hospital at Murdoch University consisted of kelpie *(n=2)*, bull terrier *(n=1)* and mixed breed dog *(n=1)*. The combined liver samples were used for the development of the canine DNA standards to generate standard curve data for qPCR. The canine DNA standards were initially quantitated via fluorescence spectroscopy with the total concentration of DNA present in the standards displayed in Table 1. The concentration of Std. 4 – 7 were out of range, due to the sensitivity limitations of the Qubit® 3.0 Fluorometer, ranging from 10 pg/µL. – 100 ng/µL. Hence the concentration data for Std. 1 – 3 were used as a reference for subsequent qPCR work.

Table 1: *Quantitation of canine DNA standards using the Qubit® 3.0 Fluorometer.*

Standard	Dilution factor	Concentration (ng/µL)			Standard Deviation
		Reaction 1	Reaction 2	Mean	
Std. 1	Neat	39.30	39.40	39.35	0.0707
Std. 2	10×	3.49	3.30	3.395	0.1344
Std. 3	100×	0.18	0.203	0.1915	0.0163
Std. 4	1000×	Out of range	Out of range	-	-
Std. 5	10000×	Out of range	Out of range	-	-
Std. 6	100000×	Out of range	Out of range	-	-
Std. 7	1000000×	Out of range	Out of range	-	-

To validate the canine DNA standard concentration data evident in Table 1, the Control DNA (0.25 ng/µL), provided in the Canine Genotypes™ Panel 2.1 Kit, was subjected to qPCR using the MC1R TaqMan® Assay. As a result, the MC1R TaqMan® Assay in

conjunction with the inputted canine DNA standard concentrations successfully quantitated the concentration of the Control DNA with +0.013 ng/µL in variation. The MC1R TaqMan® Assay showed sensitivity down to 0.036 ng/µL with the assay unable to quantitate Std. 5 (~0.0034 ng/µL), consistent with published data concerning the quantitative limitations of the MC1R TaqMan® Assay showing sensitivity to 5 pg of DNA (12). The standard curve correlation value was desirable (R^2 >0.99), indicating a close fit between the regression line and data points, hence verifying the reliability of the developed canine DNA standards (18). Additionally, the C_T values fell within the desirable range (>8 and <35) and the efficiency of the qPCR, as measured by the slope of the standard curve, deviated from the theoretical maximum of -3.3 averaging -3.618 or 90.6% efficient for quantitation of the 100 canine samples (18). Although considered acceptable for genotyping purposes, the amplification efficiency can be improved by increasing the number of standard replicates (18). However, due to the unknown historical nature of the supplied canine samples, the possible presence of PCR inhibitors may have reduced amplification efficiency (18). Nevertheless, the MC1R TaqMan® Assay and the concentration of the canine DNA standards were considered fit for use for the quantitation of canine DNA, thus implemented into the standard working method.

7.2 Allele bin adjustments based on the analysis of the control DNA provided with the Canine Genotypes™ Panel 2.1 Kit

It is evident in Figure 1 (a) and (b), that the supplied control DNA profile (a) is consistent with the obtained control DNA profile (b). Furthermore, the results displayed in Table 2 compares the genotypes of the control DNA supplied by the manufacturer and the genotype obtained from fragment analysis performed in this study based on the

electropherograms in Figure 1 (a) and (b). It can be seen in Figure 1 (a) and (b) that three peaks were present at locus FH2328 in both DNA profiles of the control DNA. It has been determined that in a population of 667 United States dogs, only the control DNA exhibited three peaks at this locus, attributing the peak to an artifact or possible tri-allelic pattern present in the cell line of the control DNA (13, 23). As the possible third allele was not included in the supplied genotype, this peak was not included in the genotype information presented in Table 2 (a) and (b).

(a) **(b)**

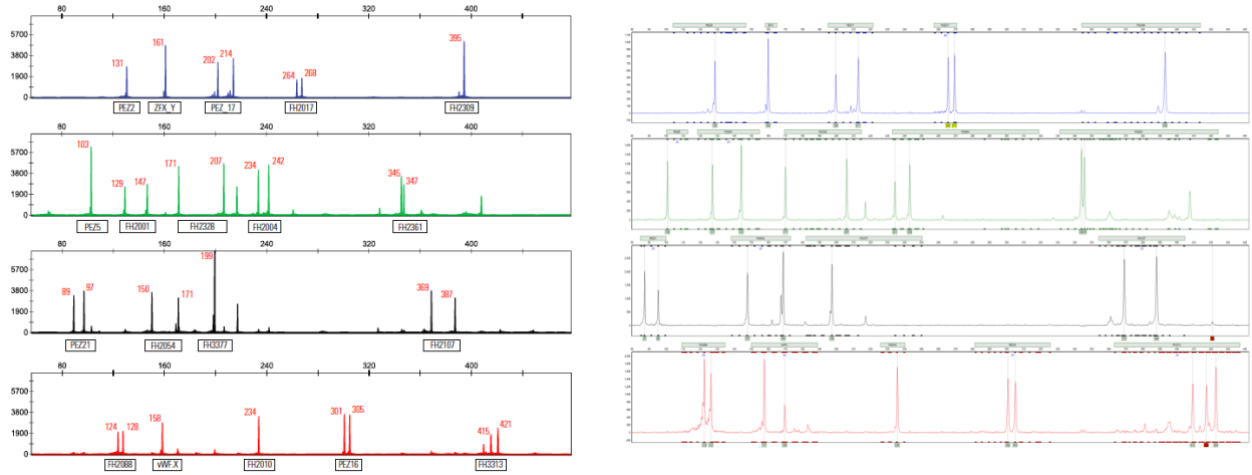


Figure 1: (a) *DNA profile of control DNA supplied by the manufacturer and* ***(b)*** *DNA profile of control DNA run with 100 canine DNA samples, both using LIZ 500 size standard with ABI PRISM® 3130xl genetic analyser.*

Table 2: (a) *Genotype of control DNA supplied by the manufacturer and* ***(b)*** *genotype of control DNA run with 100 canine DNA samples, both using LIZ 500 size standard with ABI PRISM® 3130xl genetic analyser, based on Figure 1 (a) and (b).*

(a)

PEZ02	ZFX/Y	PEZ17	FH2017	FH2309
131, 131	161, 161	202, 214	264, 268	395, 395
PEZ05	FH2001	FH2328	FH2004	FH2361
103, 103	129, 147	171, 207	234, 242	345, 347
PEZ21	FH2054	FH3377	FH2107	
89, 97	150, 171	199, 199	369, 387	
FH2088	vWF.X	FH2010	PEZ16	FH3313
124, 128	158, 158	234, 234	301, 305	415, 421

(b)

PEZ02	ZFX/Y	PEZ17	FH2017	FH2309
128, 128	160, 160	199, 211	266, 270	394, 394
PEZ05	FH2001	FH2328	FH2004	FH2361
100, 100	127, 143	171, 207	231, 239	343, 345
PEZ21	FH2054	FH3377	FH2107	
87, 95	147, 168	195, 195	370, 390	
FH2088	vWF.X	FH2010	PEZ16	FH3313
123, 127	157, 157	236, 236	300, 304	412, 420, 424

The Canine Genotypes™ Panel 2.1 Kit technical guide recommends using the control DNA to calibrate the allele sizes observed (10). As the bin file provided by the manufacturer was created based on evaluating genotyping data from 667 dogs from the United States, the resultant bins did not align with Western Australian dog population and required adjustment (24). Thus, the inconsistencies seen between the genotype data of the supplied control DNA and the run data obtained from analysing 100 Western Australian canines, calibration using the supplied control DNA genotype was not possible on the Western Australian domestic dog dataset. Hence the bins used to establish the allele calls displayed in Table 2 (b) is based on the Western Australian domestic dog data, differing from the allele calls established by the manufacturer in Table 2 (a). The development of an allelic ladder would mitigate the effects of migration variation between laboratories, possible resolving the differences seen between population genotypes associated with the control DNA.

7.3 Exclusion of samples FD0025, FD0030, FD0062, FD0031, FD0032 and FD0054

For the 100 canine samples genotyped, 94 of the samples produced interpretable DNA profiles with peaks detected for 10–15 loci when 0.25–1.0 ng of template DNA was amplified. Samples FD0025 (Boxer), FD0030 (mixed breed) and FD0062 (mixed breed) did not produce detectable peaks (RFU > 50) at all 19 loci. According to the quantitation data, sample FD0025 contained 0.120 ng/µL and sample FD0030 contained 0.115 ng/µL of DNA, significantly lower than the manufacturer's recommended input of 1–2 ng/µL and outside detection range of 10 – 0.125 ng/µL, hence resulting in allelic and locus dropout (10, 24). Similar to sample FD0025 and FD0030, sample FD0062 also failed to produce a DNA profile. However, sample FD0062 did contain sufficient concentrations of DNA (8.513 ng/µL).

Therefore, due to incorrect quantitation data or inaccurate pipetting when preparing PCR amplification and capillary electrophoresis plates, sample FD0062 did not produce detectable peaks. Consequently, samples FD0025, FD0030 and FD0062 were excluded from further analysis.

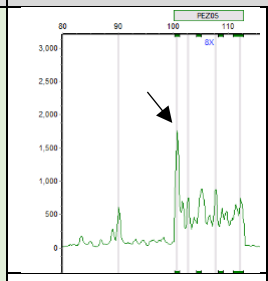
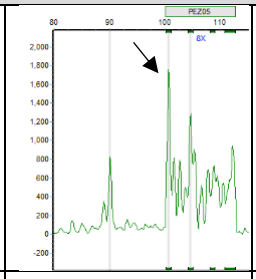
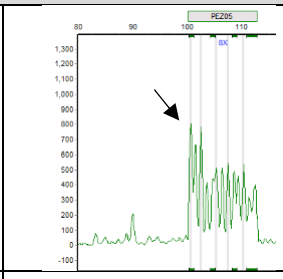
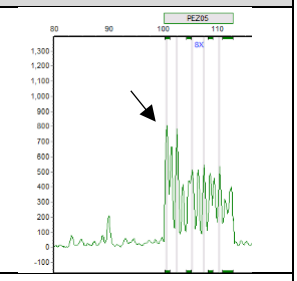
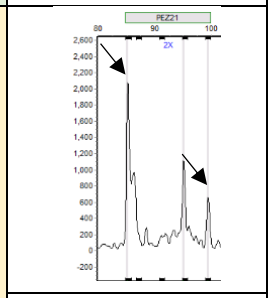
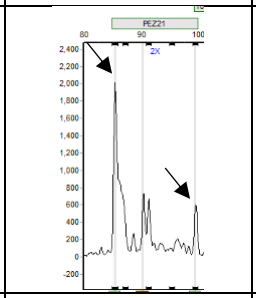
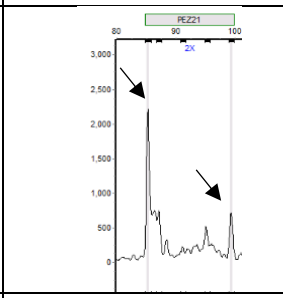
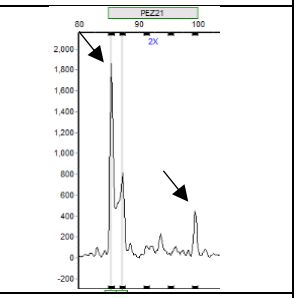
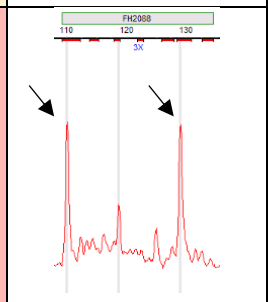
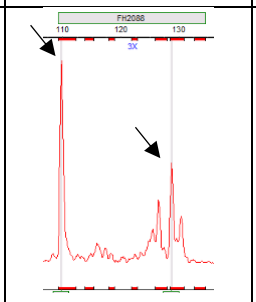
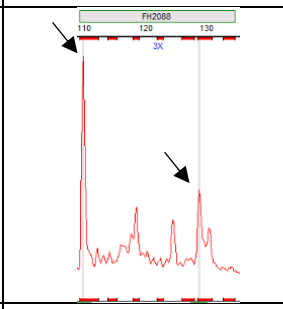
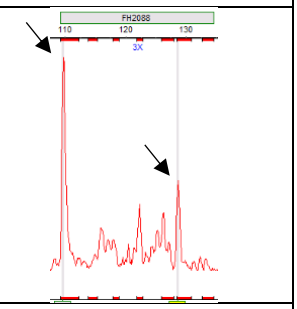
Samples FD0031(West Highland White Terrier), FD0032 (Staffordshire Bull Terrier) and FD0054 (Labrador Retreiver) produced profiles containing more than three peaks at single loci throughout the green and yellow colour channels, specifically at the PEZ05, FH2001, FH2361, FH2328, PEZ21, FH3377, FH2106 loci. The morphology of the multiple peaks detected in the green and yellow colour channel did not resemble that of the allelic peaks present in the blue and red colour channels; instead appears consistent with elevated baseline noise. Elevated baseline noise has been documented in the green and yellow channels for samples containing increased concentrations of DNA than the recommended 1–2 ng/µL (24). To rectify this issue, samples FD0031, FD0032 and FD0054 can be re-quantitated to ensure accurate concentration data and re-analysed via PCR amplification and capillary electrophoresis. Therefore, due to complications regarding interpretation, samples FD0031, FD0032 and FD0054 were also excluded from further analysis.

7.4 Interpretation issues associated with PEZ05, PEZ21 and FH2088

As PEZ05, PEZ21 and FH2088 produce the smallest DNA fragments in their respective colour channels; it would be expected of these loci to produce the highest RFU signal as polymerase preferentially amplifies smaller fragments (25). However, PEZ05, PEZ21 and FH2088 were the loci with the lowest RFU values in the green, yellow and red

colour channel, respectively. A developmental study published by Dayton, et al., (2009) also found that the locus with the lowest RFU value in the green channel is PEZ05, differing 1000-3000 RFU from the highest normalised peak (24). Consequently, it can be stipulated that the primer pairs for PEZ05 were not performing optimally with the other primer pairs in the green channel (24). Furthermore, it is possible that the primer pairs for PEZ21 and FH2088 are also not performing optimally with the other pairs in the yellow and red channels. Subsequently, an elevation in baseline noise was present at PEZ05, PEZ21 and FH2088 loci. The low RFU values of possible allelic peaks in conjunction with the elevated baseline noise and artifacts producing high RFU values, significantly complicated interpretation at these loci, as displayed in Table 3. Increasing the concentration of input DNA may enhance interpretation at the PEZ05, PEZ21 and FH2088 loci, thus potentially increasing the signal in relation to the other markers in the colour channels, improving intracolour balance. As less than 10% of the canine samples analysed produced interpretable allelic data at PEZ05, PEZ21 and FH2088, for the purposes of this study, the PEZ05, PEZ21 and FH2088 loci were excluded from further population analysis.

Table 3: *Image of electropherogram at PEZ05, PEZ21 and FH2088 displaying peak data, elevated baseline noise and artifacts (indicated by the arrows) consistent across samples four randomly selected samples.*

Locus	Images displaying uninterpretable peaks from four randomly selected samples			
PEZ05				
	FD0011	FD0034	FD0061	FD0090
PEZ21				
	FD0021	FD0047	FD0070	FD0095
FH2088				
	FD0018	FD0044	FD0059	FD0084

7.5 Partial profiles obtained using the Canine Genotypes™ Panel2.1

Partial profiles were observed in 70 % of the canine samples analysed, with 90% of these samples exhibiting no data at the FH2309 locus, displayed in Table 4. It has been established that from degraded samples, there is an inverse relationship between the size of the locus and successful PCR amplification of that locus (6). A comparison of peak heights, within the blue colour channel, between the shortest and longest loci revealed a common "ski slope" effect occurring in the 62 samples that failed to produce data at the FH2309 locus, presented in Figure 2.

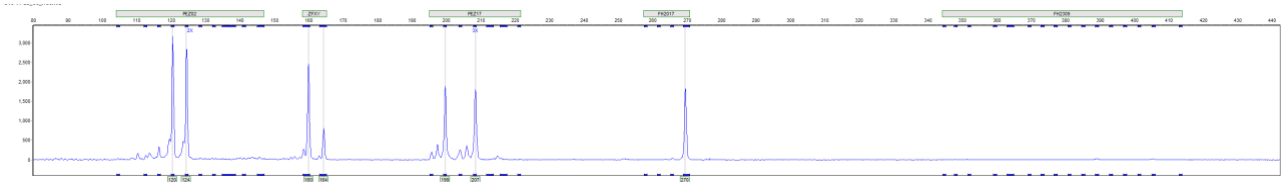


Figure 2: *Example of observed "ski slope" in the blue colour channel for sample FD0058 yielding a partial profile with no data detected for the FH2309 locus.*

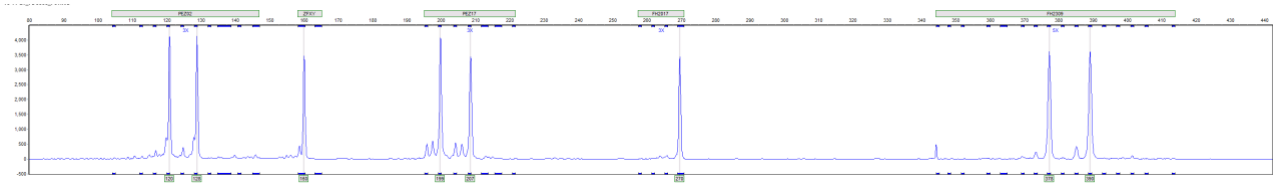


Figure 3: *Example of the blue colour channel for sample FD0006 yielding a full DNA profile with no sign of degradation or the presence of PCR inhibitors.*

It is evident in Figure 2 that the smaller loci (PEZ02 and ZFX/Y) produced significantly higher peak heights in comparison to the larger loci (PEZ17 and FH2017). However, for the 32 canine samples producing interpretable allelic information at the FH2309 locus, in addition to the control DNA, no obvious "ski slope" was apparent as evident in Figure 3. The observed "ski slope" suggests that PCR amplification was not performing optimally due to the lack of target DNA from DNA degradation or the presence of PCR inhibitors (6). The

PrepFiler™ Kit was used for the 100 canine samples analysed, and it has been established that the PrepFiler™ Wash Buffer maximises the removal of most PCR inhibitors (15). Hence it is possible that PCR inhibitors may remain present in the canine DNA extracts resulting in reduced amplification efficiency. Additionally, due to the unknown historic nature of the provide canine blood samples, the 62 samples may exhibit DNA degradation, also resulting in partial profiles with reduced sensitivity for larger DNA fragments (26). Furthermore, the degradation of DNA or presence of PCR inhibitors is also evident in the yellow and red colour channels with 13 combined samples failing to produce allelic and locus information at FH2107, largest locus in the yellow colour channel, and FH3313, largest locus in the red colour channel, displayed in Table 4. However, due to the interpretation issues associated with PEZ05, PEZ21 and FH2088, evidence of a "ski slope" cannot be observed. The development of a qPCR assay that can examine the degree of degradation by the analysis of a large and small amplicon, similar to those implemented in human qPCR techniques, could assess the level of DNA degradation and predict the effect on downstream processes (27).

Table 4: *Number of canine samples (n = 94) with data absent at a given locus indicated by colour channel and size range.*

Locus	Size range (bp)	No. of samples with absent peaks
PEZ02	104 - 145	0
ZFX/Y	159 - 164	0
PEZ17	190-225	0
FH2017	256-276	8
FH2309	339-428	62
PEZ05*	92-117	ND
FH2001	118-160	0
FH2328	171-213	4
FH2004	232-326	1
FH2361	322-439	0
PEZ21*	83-103	ND
FH2054	139-177	0
FH3377	183-305	0
FH2107	291-426	2
FH2088*	94-138	ND
vWF.X	151-187	0
FH2010	221-243	0
PEZ16	280-332	0
FH3313	340-446	11

**No interpretable data obtained for this locus (ND = no data).*

7.6 Common artifacts exhibited using the Canine Genoytpes™ Panel 2.1

Table 5 displays the location of common artifacts that have been published or observed in at least 50 % of canine samples producing allelic data at the locus exhibiting the artifact. Prior literature has established dye blobs occurring at PEZ02, PEZ05, FH2361, PEZ21, FH2088 and FH3313 (10, 24). Due to the interpretation issues associated with PEZ05, PEZ21 and FH2088, the peaks associated with dye blobs could not be characterised. Dye blobs were observed at PEZ02, FH2328, FH2004, FH2361, FH3377 and FH2107 in the STR profiles of the at least 50% of the Western Australian domestic dogs sampled, supplied

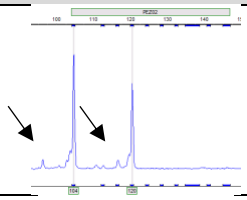
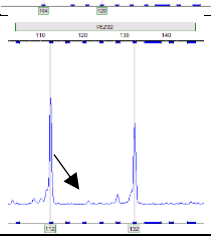
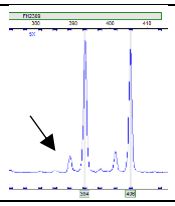
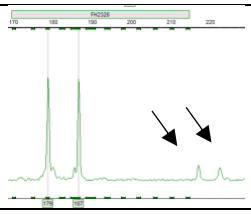
control DNA and negative control. The dye blobs generally appeared as broad low-level peaks occurring at characteristically consistent sizes in the relevant colour channel. Once characterised, the dye blobs did not cause interference with allele calls for the canine samples analysed, with the exception of nonspecific PCR products occurring at locus FH3313. However, caution should be taken when interpreting samples with peak heights of low RFU values such as potential evidentiary samples.

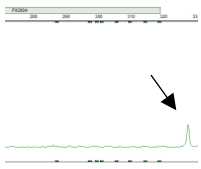
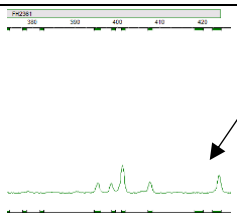
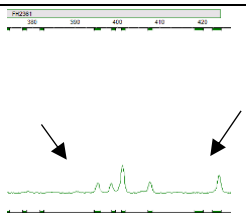
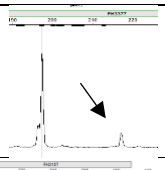
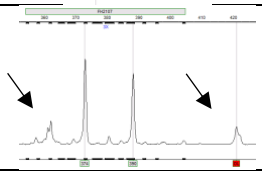
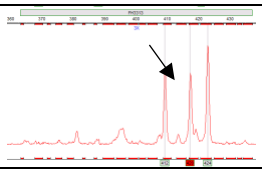
The artifact peak at the FH3313 locus caused interpretation complications, regarding the characterisation of the peak typically present at ~409 bp, presented in Table 5. Out of the 83 canine samples producing allelic data at the FH3313 locus, 47% exhibited the ~409 bp artifact peak, often called as allele 412 that proved difficult to determine if the allele call was a true peak. As the artifact produced RFU values consistent with that of allelic peaks, canines that were possible heterozygous at FH3313, exhibiting the artifact, were typically typed as tri-allelic at the FH3313 locus. Whereas, canines that were possible homozygous at FH3313, exhibiting the artifact, were typed as heterozygous at the FH3313 locus. However, due to the morphology and relative RFU values that the possible artifact displayed at the FH3313 locus for 39 canines, the peak could not be confidently characterised as a dye blob or allele 412. Further investigative studies are required regarding the characterisation of genotypes exhibiting a peak at ~409 bp, hence the FH3313 locus was excluded from population studies.

Examination of the canine DNA profiles revealed the presence of stutter commonly occurring at PEZ02 and FH2309. As evident in Table 5, PEZ02 contained both n-2 and n-4 stutter peaks, consistent with published literature concerning this locus (24). The canine samples processed also exhibited n-4 stutter peaks at locus FH2309, inconsistent with the

published developmental study (24). Dayton et al., (2009) detected n+4 stutter peaks when the height of the parent peak was above 200 RFU, however, only n-4 stutter at the FH2309 locus was observed with the current data set (24). Although stutter products did not impact allele calls for the 94 canine samples analysed, it may prove challenging to determine a stutter product of an adjacent allele from a true allele of a minor contributor in evidentiary samples containing more than one contributor (14).

Table 5: *Published and observed dye blobs and stutter (indicated by the arrows) seen in ≥50% of canine samples producing allelic data at the locus exhibiting the artifact.*

Locus	Location (bp)	Classification	Example	Published	Observed
PEZ02	n-2 n-4	Stutter		✓ Dayton et al., (2009)	✓
	~123	Dye blob		✓ Dayton et al., (2009)	✓
FH2309	n+4	Stutter	Not observed	✓ Dayton et al., (2009)	✗
	n-4	Stutter		✗	✓
PEZ05*	~117	Dye blob	ND	✓ Dayton et al., (2009)	✗
FH2328	~218 ~222	Dye blob		✗	✓

FH2004	~326	Dye blob		✕	✓
FH2361	~407	Dye blob		✓ Thermo Scientific (2014)	✓
	~391–425	Series of possible dye blobs		✕	✓
PEZ21*	~96	Dye blob	ND	✓ Dayton et al., (2009)	✕
FH3377	~218	Dye blob		✕	✓
FH2107	~362 ~420	Dye blob		✕	✓
FH2088*	~112 ~118	Dye blob	ND	✓ Dayton et al., (2009)	✕
FH3313	~409	Dye blob		✓ Thermo Scientific (2014)	✓

**No interpretable data obtained for this locus (ND = no data).*

In addition to dye blobs and stutter products, an examination of the electropherograms containing canine STR data displayed the presence of pull-up in less than 15% of canine DNA profiles. Pull-up was noticed in the blue colour channel as a result of peak signals that were carried over from the adjacent green colour channel, evident in Figure 4. Typically, the blue pull-up peaks were seen within the FH2361 locus (~340–350 bp) with an RFU value less than 100.

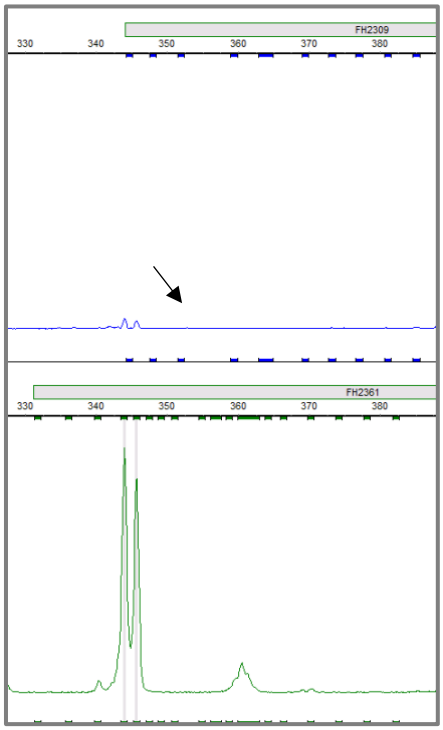


Figure 4: *FH2309 (blue colour channel) exhibiting pull-up (indicated by the arrow) due to carry over of peak signal from FH2361 (green colour channel).*

7.7 Development of the Western Australian domestic dog microsatellite database

The dog samples used in this study were assumed to be unrelated and consisted of 69 pure breed canines representing 41 dog breeds and 31 mixed breed canines. The canine samples that contained sex information were correctly typed by the ZFX/Y locus (n = 92), with the exception of sample FD0055 (Dalmatian). The ZFX/Y locus exhibited low peak height ratios with the Y peak (164) amplifying shorter than the X peak (160). In five samples, conformational review of the Y peak was required due to the low peak height ratios associated with the ZFX/Y locus. Regarding sample FD0055, the sex information provided specified the dog as female, however, the genotype at the ZFX/Y locus was consistent with a male genotype. Verification of the origin of the sample and re-testing is required to verify the male genotype observed in sample FD0055.

The STR allele frequencies obtained for the Western Australian domestic dog population are displayed in Table 6. For the Western Australian domestic dogs that were examined, the number of alleles typed per locus was less than the number of alleles typed in the United States dog population at all loci. Due to the adjustment of allele bins and lack of a nomenclature system generated from the number of repeat units, the frequency data obtained from Western Australian dogs cannot be reliably compared to that of United States dogs. Hence the impact of genetic variation based on geographical location cannot be determined.

To calculate match probabilities based on canine STR data, the inbreeding history of domestic dogs needs to be assessed. The developmental study concerning 667 United States dogs determined that a θ correction value of 0.09 was suitable to account for population substructure among different breeds, much higher than that compared to recommended θ values for humans (13). To evaluate the gene diversity for the combined 14 loci, the STR allele frequencies for the Western Australian domestic dog population, displayed in Table 6 can be used to determine the RMP and LR for the control DNA sample. Incorporating a θ correction value of 0.09, the RMP and LR for the control DNA for the combined 14 loci using the Western Australia domestic dog database was 4.24×10^{-14} and 2.36×10^{13}, respectively. The Unites States dog database calculated the RMP and LR for the control DNA at 3.47×10^{-23} and 2.89×10^{22}, respectively, using the combined 18 STR loci included in the kit (13). Subsequently, the loci excluded in the calculation of match probabilities using the Western Australian domestic dog database, PEZ05, PEZ21, FH2088 and FH3313, significantly decreased the significance of a match. Therefore, to effectively

compare the gene diversity for the combined 18 loci using the Western Australia domestic dog database to the United States dog database the issues associated with PEZ05, PEZ21, FH2088 and FH3313 would need to be resolved.

Locus	Allele	Frequency	Locus	Allele	Frequency	Locus	Allele	Frequency	Locus	Allele	Frequency
PEZ02 (n=10)	104	0.010582011	PEZ05*	ND	ND	PEZ21*	ND	ND	FH2088*	ND	ND
	112	0.042328042	FH2001 (n=9)	119	0.015957447	FH2054 (n=9)	139	0.010638298	vWF.X (n=6)	157	0.484042553
	116	0.052910053		123	0.015957447		143	0.04787234		163	0.335106383
	120	0.26984127		127	0.303191489		147	0.170212766		169	0.122340426
	124	0.296296296		131	0.106382979		152	0.223404255		175	0.015957447
	128	0.248677249		135	0.021276596		156	0.10106383		181	0.031914894
	132	0.031746032		139	0.367021277		160	0.111702128		187	0.010638298
	136	0.037037037		143	0.095744681		164	0.106382979	FH2010 (n=4)	228	0.074468085
	140	0.005291005		147	0.031914894		168	0.180851064		232	0.377659574
	128	0.005291005		151	0.042553191		172	0.04787234		236	0.212765957
ZFX/Y (n=2)	160	0.734042553	FH2328 (n=10)	175	0.077777778	FH3377 (n=12)	183	0.074468085		240	0.335106383
	164	0.265957447		179	0.016666667		187	0.042553191	PEZ16 (n=11)	280	0.005291005
PEZ17 (n=7)	195	0.015957447		183	0.15		191	0.175531915		284	0.042328042
	199	0.218085106		187	0.161111111		195	0.164893617		288	0.142857143
	203	0.218085106		191	0.183333333		199	0.015957447		292	0.227513228
	207	0.265957447		195	0.127777778		203	0.063829787		296	0.169312169
	211	0.180851064		199	0.127777778		207	0.069148936		300	0.206349206
	215	0.079787234		203	0.122222222		211	0.021276596		304	0.105820106
	219	0.021276596		207	0.027777778		235	0.085106383		308	0.068783069
FH2017 (n=4)	258	0.069767442		215	0.005555556		239	0.143617021		312	0.010582011
	262	0.087209302	FH2004 (n=11)	231	0.142076503		243	0.090425532		316	0.015873016
	266	0.529069767		235	0.459016393		247	0.053191489		324	0.005291005
	270	0.313953488		239	0.142076503	FH2107 (n=13)	354	0.011111111	FH3313*	ND	ND
FH2309 (n=15)	344	0.045454545		243	0.120218579		358	0.033333333			
	348	0.090909091		247	0.021857923		362	0.072222222			
	352	0.045454545		286	0.005464481		364	0.005555556			
	360	0.03030303		296	0.005464481		366	0.077777778			
	364	0.015151515		300	0.038251366		370	0.216666667			
	370	0.060606061		304	0.038251366		374	0.144444444			
	378	0.075757576		313	0.021857923		378	0.172222222			
	382	0.075757576		317	0.005464481		382	0.127777778			
	386	0.060606061	FH2361 (n=19)	331	0.005208333		386	0.077777778			
	390	0.106060606		335	0.020833333		390	0.05			
	394	0.151515152		339	0.161458333		396	0.005555556			
	398	0.121212121		343	0.234375		404	0.005555556			
	402	0.045454545		345	0.005208333						
	406	0.060606061		347	0.234375						
	414	0.015151515		349	0.015625						
				351	0.166666667						
				355	0.067708333						
				357	0.010416667						
				359	0.010416667						
				361	0.020833333						
				363	0.010416667						
				367	0.005208333						
				383	0.010416667						
				399	0.005208333						
				401	0.005208333						
				409	0.005208333						
				423	0.005208333						

Table 6: *Observed Western Australian domestic dogs STR allele frequencies (n = number of different allele types)* **No interpretable data obtained for this locus (ND = no data)*

8. CONCLUSION

As dog derived evidence is commonly encountered at crime scenes, the analysis of canine material can aid forensic investigations. The utility of canine DNA evidence has demonstrated the demand for a viable working method and an Australian microsatellite database. By validating the use of the canine-specific MC1R TaqMan® assay and associated standards, the efficacy of the assay is suitable for routine canine DNA quantitation. Subsequently, using the PrepFiler™ kit for extraction, the MC1R TaqMan® assay with validated standards for quantitation and the Canine Genotypes™ 2.1 Panel Kit for genotyping, a workable method was established to obtain a canine DNA profile. Based on the Western Australian domestic dog data set, common interpretation issues associated with the PEZ05, PEZ21, FH2088 and FH3313 loci, complicating the identification of true allele peaks. Further studies relating to the effect of input DNA concentration and adjustments of injection time and voltage is recommended to increase the signal intensity, thus aiding interpretation. Additionally, complications arising from artifact peaks were reduced by establishing the location and morphology of common artifacts exhibited using the Canine Genotypes™ Panel 2.1. As a result, the observed Western Australian domestic dog STR allele frequencies were obtained for 14 loci examined, forming the foundation of the canine microsatellite database. The data presented in this study has the potential to be a valuable national resource and further the application and development of canine DNA analysis in the Australian forensic science community. It is recommended to expand the database geographically to include other Australian states and territories for the database to be representative of the Australian domestic dog population.

9. REFERENCE LIST

1. Animal Medicines Australia. Pet Ownership in Australia: Animal Medicines Australia Pty Ltd; 2016.
2. Rajshekar M, Blizzard L, Julian R, Williams A, Tennant M, Forrest A, et al. The incidence of public sector hospitalisations due to dog bites in Australia 2001–2013. Australian and New Zealand Journal of Public Health. 2014;41(4):377-80.
3. Clarke M, Vandenberg N. Dog attack: the application of canine DNA profiling in forensic casework. Forensic Sci Med Pathol. 2010;6(3):151-7.
4. Shutler G, Gagnon P, Verret G, Kalyn H, Korkosh S, Johnston E, Halverson J. Removal of a PCR inhibitor and resolution of DNA STR types in mixed human-canine stains from a five year old case. Journal of forensic sciences. 1999;44(3):623.
5. Wictum E, Kun T, Lindquist C, Malvick J, Vankan D, Sacks B. Developmental validation of DogFiler, a novel multiplex for canine DNA profiling in forensic casework. Forensic Sci Int Genet. 2013;7(1):82-91.
6. Butler J. Fundamentals of Forensic DNA Typing. Burlington: Elsevier Inc.; 2010.
7. Halverson J, Basten C. A PCR Multiplex and Database for Forensic DNA Identification of Dogs. J Forensic Sci. 2005;50(2):352-63.
8. Kanthaswamy S. Review: domestic animal forensic genetics - biological evidence, genetic markers, analytical approaches and challenges. Anim Genet. 2015;46(5):473-84.
9. Ostrander EA, Sprague G, Rine J. Identification and characterisation of dinucleotide repeat (CA)n markers for genetic mapping in dog. Genomics. 1993;16(1):207-13.
10. Thermo Scientific. Technical Manual: Canine Genotypes Panel 2.1: Thermo Fisher Scientific Inc.; 2014.
11. Richard ML, Frappier R, Newman JC. Developmental validation of a real-time quantitative PCR assay for automated quantification of human DNA. Journal of forensic sciences. 2003;48(5):1041.
12. Evans JJ, Witcum E, Penedo MCT, Kanthaswamy S. Real-Time Polymerase Chain Reaction Quantification of Canine DNA. Journal of Forensic Sciences. 2007;52(1):93-6.

13. Kanthaswamy S. Development and Validation of a Standardized Canine STR Panel for Use in Forensic Casework. National Institute of Justice (NIJ) National Criminal Justice Reference Service; 2009.
14. Butler J. Advanced Topics in Forensic DNA Typing: Interpretation. Oxford: Elsevier Inc.; 2015.
15. Thermo Scientific. PrepFiler Forensic DNA Extraction Kit: Applied Biosystems Inc.; 2008.
16. Miller SA, Dykes D, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Research. 1988;16(3):1215.
17. Thermo Scientific. User Guide: Qubit 3.0 Fluorometer: Life Technologies; 2014.
18. appliedbiosystems. User guide: QuantStudio 6 and 7 Flex Real-Time PCR System Software: Life Technologies Corporation; 2013.
19. appliedbiosystems. User guide: TaqMan Fast Advanced Master Mix: Thermo Fisher Scientific Inc.; 2017.
20. appliedbiosystems. User guide: Applied Biosystems 3130/3130xl Genetic Analyzers: Life Technology Corporation; 2012.
21. Balding DJ, Nichols R. DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands. . Forensic Sci Int. 1994;64(2-3):125-40.
22. National Research Council. The evaluation of forensic DNA evidence. Washignton: United States: National Academy Press; 1996.
23. Baute DT, Satkoski JA, Spear TF, Smith DG, Dayton MR, Malladi VS, et al. Analysis of forensic SNPs in the canine mtDNA HV1 mutational hotspot region. J Forensic Sci. 2008;53(6):1325-33.
24. Dayton M, Koskinen M, Tom BK, Mattila A, Johnston E, Halverson J, et al. Developmental validation of short tandem repeat reagent kit for forensic DNA profiling of canine biological material. Croatian Medical Journal. 2009;50(3):268-85.
25. Leibelt C, Budlowe B, Collins P, Daoudi Y, Moretti T, Nunn G, et al. Identification of a D8S1179 primer binding site mutation and the validation of a primer designed to recover null alleles. Identification of a D8S1179 primer binding site mutation and the validation of a primer designed to recover null alleles. 2003;133(3):220-7.

26. Niederstätter H, Kochl S, Grubwieser P, Pavlic M, Steinlechner M, Parson W. A modular real-time PCR concept for determining the quantity and quality of human nuclear and mitochondrial DNA. Forensic Science International: Genetics. 2007;1(1):29-34.

27. Hudlow WR, Chong M, Swango KL, Timken MD, Buoncristiani MR. A quadruplex real-time qPCR assay for the simultaneous assessment of total human DNA, human male DNA, DNA degradation and the presence of PCR inhibitors in forensic samples: A diagnostic tool for STR typing. Forensic Science International: Genetics. 2008;2(2):108-25.