J Forensic Sci, January 2008, Vol. 53, No. 1
doi: 10.1111/j.1556-4029.2007.00615.x
Available online at: www.blackwell-synergy.com

Andrea L. Himmelberger,[1] *M.S.; Theresa F. Spear,*[1] *M.A.; Jessica A. Satkoski,*[2] *Ph.D.; Debra A. George,*[2] *M.A.; Wendy T. Garnica,*[2] *B.S.; Venkat S. Malladi,*[2] *B.S.; David G. Smith,*[2,3] *Ph.D.; Kristen M. Webb,*[4] *B.S.; Marc W. Allard,*[4] *Ph.D.; and Sreetharan Kanthaswamy,*[1,2,3] *Ph.D.*

Forensic Utility of the Mitochondrial Hypervariable Region 1 of Domestic Dogs, in Conjunction with Breed and Geographic Information*

ABSTRACT: The 608-bp hypervariable region 1 (HV1) sequences from 36 local dogs were analyzed to characterize the population genetic structure of canid mitochondrial DNA (mtDNA). Sixteen haplotypes were identified. A 417-bp segment of this sequence was compared with GenBank sequences from a geographically representative sample of 201 dogs, two coyotes, and two wolves. Sixty-six haplotypes were identified including 62 found only in domestic dogs. Fourteen of these correspond to the 16 local haplotypes and were among the most frequent haplotypes. The local sample was judged to be representative of the much broader geographic sample. No correlation was observed between local haplotypes and the owner's characterization of dog breed. A 60-bp variation "hotspot" within the canid HV1 was identified as a potentially valuable molecular tool, particularly for assaying limited or degraded DNA samples.

KEYWORDS: forensic science, trace evidence, domestic dog, mixed and pure breed studies, geographic origin, mitochondrial DNA, sequence variation, hypervariable region 1

According to the American Kennel Club (AKC), households in the United States include on average 1.7 dogs as companion animals. In these homes, canine evidence, such as shed hair, can be plentiful and is frequently collected with other trace evidence at crime scenes. Despite its abundance in the environment, canine evidence is underutilized in forensic investigations. In the United States, evidence from domestic dog forensic DNA testing has contributed to only about 20 criminal investigations since 1996 (1). Reasons for the infrequent use of canine evidence in investigations include the lack of awareness of its value by law enforcement personnel, the absence of forensically validated panels of robust microsatellite markers or short tandem repeats (STRs) and associated canine databases. Often such evidence, particularly single, fallen hairs (with little or no root or follicle), yield very small quantities of DNA and consequently cannot provide complete STR profiles. However, the high success rate of DNA typing of trace and transfer samples using canine mtDNA can dramatically increase the value of this evidence as this approach to genetic analysis will often yield informative results even when STR-based systems fail (1,2). Lacking a standard reference population it is of interest to know whether or not a local sample leads to a bias in estimating the probability of genetic identity between two randomly selected samples whose geographic origin cannot be known with certainty.

The canine mtDNA genome has been completely sequenced and the length of the sequence is *c*. 16727 bp (3). Most of the sequence variation among individuals is found in two segments of the control region: hypervariable region 1 (HV1) and hypervariable region 2 (HV2), i.e., between nucleotide position (np) 15458 and 16727 (4). Just like the human HV1, the canine HV1 is highly polymorphic and is of keen forensic interest because it can also be successfully amplified and typed from limited or severely degraded DNA (5–7). Sufficient mtDNA can be reliably typed from a single dog hair shaft (1,5,8–11).

Worldwide *c*. 400 dog breeds and varieties of dogs have been described and most of these originated relatively recently in canid evolution. The AKC, whose databases register just under one million dogs annually and contain an average of 13 million living dogs at any time, recognizes over 150 breeds of dogs in the United States alone (J. L. Halverson, personal communication). Most mtDNA studies have been based on purebred dogs. However, most dogs in the world are of mixed ancestry ranging from crosses of at least two distinct breeds to multi-breed combinations. Recent surveys suggest that there are over 70 million dogs in the United States and over half are mixed breed dogs (AKC and the American Pet Products Manufacturers Association [APPMA]). Because of the prevalence of mixed breed dogs in households in the United States, biomaterial from such dogs is expected to predominate at crime scenes. An appropriate forensic canine database must reflect the population that includes the suspected animal for that database to be relevant to crime scene investigation.

In the absence of a standard reference canine mtDNA database, forensic analysis is likely to be based on samples collected

[1]Graduate Group in Forensic Science, University of California, Davis, CA 95616.
[2]Department of Anthropology, University of California, Davis, CA 95616.
[3]California National Primate Research Center, University of California, Davis, CA 95616.
[4]Department of Biological Sciences, The George Washington University, Washington, DC 20052.

*Part of the research described in this article has been reported in Ms. Andrea Himmelberger's Masters of Science (Forensic Science) thesis which was submitted to the UC Davis Office of Graduate Studies on November 16, 2006, and orally presented in the UC Davis Graduate Group in Forensic Science exit seminar as part of the program's requirement.

Received 28 Jan 2007; and in revised form 20 June 2007; accepted 27 June 2007.

opportunistically or on an ad hoc basis. We describe a study that aimed to (1) characterize this local mixed breed variation and compare it with a much larger geographic sample of pure and mixed breed dogs and (2) to establish the local genetic structure of mtDNA variation. This study is part of a larger endeavor to organize an exhaustive canine mtDNA population database.

Methods

Buccal cells were taken from 58 domestic dogs by using BuccalAmp™ DNA Extraction Kit (EPICENTRE Biotechnologies, Madison, WI). Thirty-one samples were obtained from the University of California, Davis (UC Davis), Veterinary Genetics Laboratory (VGL), with the remaining 27 samples obtained from volunteers' dogs in northern and southern California. These samples consist mostly of mixed breed animals representing 19 distinct breeds, consistent with the predominance of mixed breed dogs in the United States. The animals' owners reported the breeds or the ancestry of dogs sampled for this study. A dog was characterized as purebred or of single breed ancestry based on either the owner's account or the dogs' size and appearance. Most of these dogs have formal pedigree records as those available for AKC or UKC registered animals. The mtDNA dataset generated here might be more representative of forensically important mtDNA variation among household dogs in this geographic region of the United States than samples consisting only of purebred dogs if mtDNA of dogs in the United States is sufficiently geographically structured.

The mtDNA was extracted using conventional methods of extraction (12). Briefly, each buccal swab was immersed in 200 μL of 200 mM NaOH and denatured for 5 min to release the DNA from the buccal cells. The swab was then removed and 200 μL of 200 mM HCl and 100 mM Tris–HCl (pH 8.5) were added to the tube and mixed so as to prepare the DNA for analysis.

Following isolation, DNA amplification of an approximately 800 bp fragment was carried out using primers H15360 (5′-ATTACCTTGGTCTTGTAAACC-3′) and L16106 (5′-AAACTATATGTCCTGAAACC-3′) (13). These primers target parts of the HV1 segment np 15458 to np 16727 (3) and the tRNA genes for the amino acids proline and threonine (*tRNA-Pro;* np 15392 to 15457 and *tRNA-Thr*; np 15323 to 15392). The polymerase chain reaction (PCR) amplification consisted of 3 μL of mtDNA extract, 7 μL of each primer (0.15 μM), 2.5 μL of MgCl (50 mM), 2.5 μL of each dNTP (100 mM), and 0.2 μL of Taq Gold (5 U/μL) in 25 μL. The samples were amplified on a PTC-100 thermalcycler (Bio-Rad [formerly MJ Research], Hercules, CA) following protocols in the manufacturer's manual. Cycle parameters were 5 min at 95°C followed by 5 min at 85°C (loading temperature) then 36 cycles of 95°C for 20 sec, 51°C for 30 sec, 72°C for 40 sec, and then a 4°C hold. After amplification, the PCR products were purified using the Millipore purification system. Approximately 4 μL of the product was run on a 1.4% agarose gel for 1 h at 120 V to evaluate the quantity and specificity of the desired product. The gels were evaluated for the presence and quantity of the PCR product using UV light after SYBR® Green I (Cambrex, Rockland, ME) staining for 2 min and destaining in ddH_2O for 2 min.

Approximately 40 ng of product was used for sequencing with BigDye Terminators using primers H15422 (5′-CTCTTGCTCCACCATCAGC-3′) and L16102 (5′-AACTATATGTCCTGAAACCATTG-3′) (13). The 10 μL sequencing reaction consisted of 1 μL of PCR product, 2.5 μL of primer, 0.5 μL of BigDye Terminator v3.1 Ready Reaction Premix (Applied Biosystems, Foster City, CA), 2 μL of 5X BigDye Sequencing Buffer (Applied Biosystems), and ddH_2O. The cycle parameters were 40 cycles of 95°C for 20 sec, 50°C for 10 sec, 60°C for 4 min, followed by a 4°C hold. Sequences of approximately 640 bp were generated.

After cycle sequencing, the unincorporated dyes were removed using the Millipore purification system (Millipore, Billerica, MA). Electrophoresis was performed on the ABI 3730 Genetic Analyzer (Applied Biosystems) according to the manufacturer's directions. After analyzing the forward and reverse sequences, a 608-bp consensus sequence was obtained for each sample. Samples that did not yield a matching forward and reverse sequence were resequenced. If manual editing still did not resolve the problem, the sample was not used in analysis. All sequences were aligned using ClustalW (14). The sequences were then studied for the presence of polymorphic nucleotide positions using the Bioedit (15) program.

Using GenBank, we expanded the local dataset to include sequences of the HV1 region from the reference dog, a Sapsaree breed, which is native to Korea (GenBank Accession Number U96639) reported by Kim et al. (3), and corresponding sequences from 200 dogs of various breed and geographic affiliations, two coyotes (GenBank Accession Numbers AY172674 and AY240094), and two wolves (GenBank Accession Numbers AY240073 and AY240155). The GenBank domestic dog sequences also included some of those reported in Gundry et al. (16; GenBank Accession Numbers AY240030–AY240157) that represent 43 mostly western dog breeds of which 37 are AKC recognized and those submitted by Kim et al. (17), which represent several breeds indigenous to Asia (Asian Pug, Akita, Chejudo, Chin, Jindo, Pekingese, Sapsaree, Shiba Inu, and Tosa Inu), and some western breeds (Collie, German Shephard and Yorkshire Terrier; $n = 17$; GenBank Accession Numbers AFO64570–AFO64586).

Each of the 16 local dog haplotypes (GenBank Accession Numbers EF122413–EF122428), which represented the 36 sequences that were generated in our laboratory, as well as the 205 canid sequences obtained from GenBank (accession numbers and breed names or other identification for each of these sequences are given in Appendix 1) were analyzed. Lab-generated 608-bp sequences were trimmed to 417-bp to maximize overlap with the GenBank sequences. Information on the specific geographic location that some of these samples represent is limited (R. L. Gundry, personal communication).

For both groups of data, the mean number of uncorrected pairwise differences was calculated with ARLEQUIN version 2.001 (18). Genetic heterogeneity was expressed as nucleotide diversity, π, i.e., the ratio of the number of pairwise differences to the total number of nucleotides studied, and this estimate is analogous to the gene diversity of nuclear loci (19–21). The AMOVA program from ARLEQUIN was used to estimate the degree of variation within the local and global populations and the degree of differentiation between both samples.

Results

From the original sample set of 58 animals, a 608-bp sequence spanning the HV1 region was successfully sequenced in 36 local domestic dogs representing 19 different breeds. Suboptimal storage and transfer of buccal swab samples may have caused the PCR inhibition and subsequent failed DNA amplification. Nonetheless, it is likely that this limited sampling more realistically reflects opportunistic collections of nonhuman material for casework analysis when standardized reference databases are nonexistent.

Sixteen domestic dogs' mtDNA HV1 haplotypes were identified among the 36 sequences and were numbered arbitrarily, while the reference dog and coyote haplotypes were not observed in any

TABLE 1—*Canine 608-bp mtDNA HV1 haplotypes of 36 local domestic dogs analyzed in this study.*

Haplotype	Breeds Represented (as Identified by Owner)	*n*	Frequency
1	Labrador Retriever, Dachshund, *Yellow Labrador Retriever, (African) Basenji mix*	4	0.111
2	Labrador Retriever, Labrador Retriever, Labrador Retriever, Yellow Labrador Retriever, *Airedale Terrier/Golden Retriever mix*	5	0.139
3	Cairn Terrier	1	0.028
4	Bull Mastiff, Mongrel, Yellow Labrador Retriever, German Shepherd mix, American Pit Bull Terrier	5	0.139
5	Boston Terrier	1	0.028
6	Labrador Retriever/Golden Retriever mix	1	0.028
7	Australian Shepherd/Rottweiler mix, German Shepherd/Border Collie mix, Mini Poodle, Mongrel, Black Labrador Retriever/Golden Retriever mix, *Australian Shepherd*	6	0.167
8	Border Collie	1	0.028
9	Rottweiler mix, Rottweiler mix, Chocolate Labrador Retriever mix, Rottweiler mix, *Labrador Retriever mix*	5	0.139
10	Standard Poodle mix	1	0.028
11	Bichon Frise	1	0.028
12	Queensland Heeler mix	1	0.028
13	Labrador Retriever/Border Collie mix	1	0.028
14	Australian Shepherd mix	1	0.028
15	Mixed Terrier	1	0.028
16	*Mongrel*	1	0.028

Italicized fonts indicate animals from southern California.

other animal. Table 1 reports the frequencies of the 16 different haplotypes identified among the local dogs and the associated breed types. The frequencies of these haplotypes ranged from 3% to 17%. The five most common haplotypes represented 69.4% of the local domestic dogs studied, 15 of which were sampled within a 50-mile radius of the university. Haplotype 7, which includes the greatest number of dogs (haplotypic frequency = 0.167), also represented most of the breeds including Australian Shepherd–Rottweiler mix, a German Shepherd–Border Collie mix, a Mini Poodle mix, a Labrador Retriever–Golden Retriever mix, and a purportedly purebred Australian Shepherd (Table 1). Haplotypes 2, 4, and 9 were the second most frequent haplotypes in this study, each with a frequency of 0.139. Haplotype 9 represents three Rottweiler mix dogs, a Chocolate Labrador Retriever mix, and a Labrador Retriever mix. Haplotype 2 represents Labrador Retrievers and an Airedale/Golden Retriever mix, while haplotype 4 represents a Bull Mastiff, a mongrel, a Yellow Labrador Retriever, a German Shepherd mix, and a Pit Bull Terrier. Haplotype 1 (frequency = 0.111) represents two Labrador Retrievers, a Dachshund, and Basenji mix. Thus, there was no correlation detected between the haplotype of a local dog and its breed make-up—as determined from its external features and/or its owners' description.

These haplotypes reflect 32 polymorphic nucleotide positions and one base deletion, which are identified in Table 2. Twelve of the 32 segregating sites in the 608-bp region were located within a *c.* 60-bp block between np 15595 and np 15653 in the HV1 region of the reference sequence, suggesting the presence of a mutational "hotspot". When the local 608-bp domestic dog sequences were compared with the reference dog's sequence, the number of observed transitions, transversions, and indels in the entire 608-bp region were 31, 1, and 1, respectively. All but haplotypes 1 and 5 exhibited a transversion at np $15639^{T/A/G}$, where the superscript denotes the mutation of T to either A or G following the polymorphism numbering format suggested by Pereira et al. (22). Haplotypes 7 and 10 are very similar and only differ at np 15931, while haplotype 10 has a deletion and haplotype 7 does not (six dogs exhibited haplotype 7 while only one animal exhibited haplotype 10). Haplotypes 8, 10, and 13 show a deletion, i.e., 15931^{del}.

The average number of pairwise differences among the local domestic dog haplotypes was 6.71 ± 3.24 reflecting an average nucleotide diversity (π) of 0.011 ± 0.01 (Table 3). The maximum π value between local haplotypes was 0.03 and the minimum was 0.002, representing one base pair difference between haplotypes. Haplotype 1 (representing purported purebred Labrador Retrievers, a Dachshund, and a Basenji mix) exhibited the highest identity to the reference dog sequence while haplotype 10 (a Standard Poodle) was the least similar to the reference sequence. Haplotypes 7 and 13 (representing Labrador Retriever–Border Collie mixes) exhibit a π value of 0.03, which represents 16 base pair differences and a deletion.

The exclusion capacity of the canid HV1 haplotype, or $1-\Sigma \mathbf{X}_i^2$ (where $\mathbf{X}_i$ is the frequency of the *i*th haplotype), is 0.894 and the random match probability, or $\Sigma \mathbf{X}_i^2$, is 0.106 for all 608-bp haplotypes (23). This implies that 89 out of 100 disputed individuals unrelated to a sample in question could be excluded from identity to that sample using this dataset.

When the 608-bp sequences were trimmed to 417 bp to maximize overlap with some of the GenBank sequences, the 60-bp variation hotspot was also removed and only 14 haplotypes were observed among the sequences generated from the 36 local dogs. When the trimmed local dog sequences were combined with the GenBank sequences, 62 haplotypes were observed (excluding the two coyote [*haplotypes 16* and *64*] and two wolf [*haplotypes 61* and *66*] sequences; Table 4). *Haplotypes 7, 2, 65, 3, 4,* and *1* occurred 37, 28, 23, 22, 18, and 15 times throughout the dataset. The 62 haplotypes from a total of 237 domestic dogs provided an exclusion capacity of 0.93 (a random match probability of 0.07). Most of the local dog haplotypes (57%) occurred only once in the local dataset (Table 4). This is concordant with the same estimates for the GenBank sequences (61%; Table 5) and that by Gundry et al. (16) who described 45 haplotypes in 43 breeds of dogs, a coyote and two wolves, 29 (64%) of which were observed only once in their dataset of solely purebred dogs.

In Table 5, a dog labeled as a "Chejudo" breed, which is indigenous to the Cheju Island in S. Korea (GenBank Accession Number AF064584), an Australian Shepherd/Rottweiler mix, a German Shepherd/Border Collie mix, a Mini Poodle, a Mongrel, a Black

TABLE 2—*The 32 different polymorphic sites in HV1 sequences of 36 local dogs analyzed in this study.*

Haplotype	Position																														
*	15483	15508	15526	15553	15557	15595	15611	15612	15620	15621	15627	15632	15639	15643	15650	15652	15653	15665	15710	15781	15800	15807	15814	15815	15912	15931	15932	15955	15959	16003	16025	16033
†	C	C	C	A	T	C	T	T	T	C	A	C	T	A	T	G	A	T	C	C	T	C	C	T	C	A	G	C	C	A	T	A
1																							T									
2									C		G		A										T									
3				G									A										T								C	
4									C		G		A										T					T				
5																	G						T									
6											G		A			A							T									
7			T			T		C				T	G	G		A					C		T	C	T			T		G		
8		T	T				C						G		C				T		C		T		T	d	A	T		G		
9													A										T								C	
10			T			T		C				T	G	G		A					C		T	C	T	d		T		G		
11					C								A										T									
12	T										G		A										T		T							
13											G		A										T			d			T		C	G
14													A										T									
15										T			A					C		T		T	T									
16											G	T	A										T								C	

d, deletion.
*Position of sequence variants based on reference sequence.
†Reference sequence.

Labrador Retriever/Golden Retriever mix, and an Australian Shepherd from this study exhibited *haplotype 7*. *Haplotype 2* was the second most frequent haplotype and among the dogs that belonged to this haplotype is a dog classified as a Korean Sapsaree breed (GenBank Accession Number AF064571), as well as a Labrador Retriever, two Labrador Retrievers, a Yellow Labrador Retriever, and an Airedale Terrier/Golden Retriever mix from this study. *Haplotype 65* did not represent any of the local domestic dogs. A Pekingese (GenBank Accession Number AF064575) and a Cairn Terrier, three Rottweiler mix breed dogs and two Labrador Retriever mix breed dogs belonged to *haplotype 3*, while a dog that belongs to the Japanese Tosa Inu breed (GenBank Accession Number AF064578), a Bull Mastiff, a Mongrel, a Yellow Labrador Retriever, a German Shepherd mix, and an American Pit Bull Terrier belonged to *haplotype 4*. In addition to the other dogs in this study, *haplotype 1* represented a Labrador Retriever, a Dachshund, a Yellow Labrador Retriever, and a Basenji mix from the local domestic dog category. A dog labeled as "German Shepherd" breed (GenBank Accession Number AF064573) exhibited a 417-bp haplotype that was identical to the reference dog, i.e., *haplotype 15*.

The mean number of pairwise differences and value of π among the domestic dog 417-bp haplotypes were 5.70 ± 2.73 and 0.013 ± 0.01, respectively (Table 3). The local samples which were represented by 14 haplotypes (of these only one haplotype, *haplotype 10*, was restricted to the local animals), reflected a π value of 0.016 ± 0.01, while the geographically representative samples represented by 61 haplotypes (of which 48 were found only in the global population) had a π estimate of 0.023 ± 0.01 (Table 3). The pairwise differences between haplotypes and gene diversity estimates within the local samples were 4.95 ± 2.47 and 0.012 ± 0.01 and within the global samples were 5.84 ± 2.80 and 0.014 ± 0.01, respectively. The AMOVA test revealed that all of the genetic variance is explained by genetic diversity within local and global dog sequences (Table 6).

Discussion

When reference and evidentiary samples originate from the same geographic populations, as is likely to be the case, representativeness of the reference samples comprising the database is relevant in assessing the rarity of DNA profiles associated with evidence found at a crime scene. Therefore, a more precise understanding of the genetic structure of local and more global dog populations is needed to provide a better interpretation of the meaning of a "matching" DNA profile or haplotype.

The degree of haplotypic diversity we identified within a particular local geographic location is comparable to that reported by Gundry et al.'s (16) Massachusetts (MA) dog samples where 20 haplotypes were observed (Table 5). However, the MA diversity could be attributed to the much higher sample size from that state ($n = 51$) compared with our data from a local population which is based on only 36 local dogs. On the other hand, the presence of only three haplotypes among Gundry et al.'s (16) sample set of 62 dogs from Texas reflects strong founder effects on domestic dog matrilines. Five of the local dog haplotypes identified in the present study were found to represent 70% of the local dogs studied. In spite of the smaller sampling size compounded by the possible losses of significant local haplotypes resulting from the failed DNA tests, our study showed that the local haplotypic pairwise differences and gene diversity were only slightly lower than those estimated from the GenBank sequences which represented more geographically representative samples. This, together with the AMOVA, suggests that a local reference sample is sufficient to characterize most of the

TABLE 3—*Pairwise differences and nucleotide diversity (π) among all domestic dog haplotypes observed in this study.*

	Local 608-bp Haplotypes (n = 36)	Local and Global 417-bp Haplotypes (n = 237)	Local 417-bp Haplotypes (n = 36)	Global 417-bp Haplotypes (n = 201)
Pairwise differences	6.71 ± 3.24	5.70 ± 2.73	4.95 ± 2.47	5.84 ± 2.80
π	0.011 ± 0.01	0.014 ± 0.01	0.012 ± 0.01	0.014 ± 0.01

TABLE 4—*Haplotypes observed among all 241 animals based on the analysis of the 417-bp fragment.*

Haplotype*	Numbers†	Haplotype	n
1	15 (4)	34	2
2	28 (5)	35	1
3	22 (6)	36	1
4	18 (5)	37	1
5	2 (1)	38	1
6	3 (1)	39	1
7	37 (6)	40	1
8	3 (1)	41	1
9	2 (1)	42	1
10	1 (1)	43	2
11	6 (2)	44	1
12	2 (1)	45	2
13	4 (1)	46	1
14	3 (1)	47	1
15	3	48	1
16 (coyote)	1	49	1
17	4	50	1
18	1	51	1
19	1	52	1
20	1	53	1
21	1	54	1
22	4	55	1
23	1	56	1
24	3	57	1
25	1	58	2
26	2	59	1
27	1	60	1
28	1	62	1
29	1	61 (wolf)	1
30	1	63	2
31	4	64 (coyote)	1
32	1	65	23
33	3	66 (wolf)	1

**Haplotype 3* corresponds to haplotypes 3 and 9 in Table 1. Similarly, *haplotype 9* to 10, *haplotype 10* to 11, *haplotype 11* to 12 and 14, *haplotype 12* to 13, *haplotype 13* to 15, and *haplotype 14* to 16.

†California numbers are in parentheses.

variation within a global sample of dogs and to estimate parameters of probability of exclusion. The high level of sharing among the most common haplotypes, high frequency of singleton types, and high level of diversity in both the local and global samples support our conclusion that the two population samples are comparable and useful for forensic applications.

No general correlation between breed and haplotype was found, although data from this study showed that four of the five animals represented by haplotype 2 were identified by their owners as Labrador Retrievers (Table 1). Besides the observation that some haplotypes were disproportionately represented in some breed mixes than other haplotypes, the intra-haplotypic breed representation is also quite notable. This was also observed in Gundry et al.'s purebred dog dataset (16), which has been summarized in our Table 5, as well as other dog mtDNA datasets (4,6,7,17,23). Confidence in the informal accounts of the dog owners concerning information on their pet's breed background is problematical because this determination is usually based on the animal's physical appearance and/or size.

Another possible reason why mtDNA markers fail to achieve breed and/or geographic resolution could be that 75% of all modern dogs are hypothesized to descend from one female Asian wolf (13,24). Dogs were first domesticated over 10 millennia ago (13,25–27). While some breeds such as retrievers, water spaniels, and terriers have existed for over 1000 years, most modern dog breeds were only developed as recently as the 17th and 18th centuries (25–27). Therefore, the divergence of canid mtDNA may be too recent for valid assessments of genetic relationships among breeds, and in some cases within breed (24), and any chance relationships between breed and haplotype are probably an artifact of random genetic drift. Because the maternally inherited mitochondrial genome is haploid, the effective population size at this locus is one-fourth that of biparentally inherited autosomal loci. Thus, genetic drift leads to the fixation of alleles four times more quickly in the mitochondrial genome than in the nuclear genome (28).

Moreover, compared with inferences on genetic distance and gene flow that are based on autosomal markers, inferences that are based on mtDNA are female biased, and Y-linked markers are male-biased. This is particularly noteworthy because the maintenance of breeds, types, or styles relies on the introduction of novel genes paternally (17) by means of inbreeding (mating involving first degree relatives such as the parents and siblings) and linebreeding (when there is an ancestor in the pedigree that is common to the sire and dam of the dog in question). As such, the problem of clarifying genetic relationships among mixed breed dogs with complicated breed history remains.

As more than one-third of the 32 point mutations discovered were found within a single 60-bp region, this region might be considered a "hotspot" for sequence polymorphism. The stark differences in estimates of pairwise differences and nucleotide diversity estimates between the 608-bp and 417-bp fragments from the local samples reflect the level of information within this variation hotspot of the canine HV1 region (Table 3). If primers were designed to amplify shorter templates of target DNA which comprise this highly polymorphic 60-bp block region rather than a much longer fragment, then canid HV1 assay could become more suitable for investigating severely environmentally challenged forensic samples, where the DNA is highly degraded or contaminated with PCR inhibiting agents.

It is interesting to note that the 14 417-bp long haplotypes generated in this study were among the most frequently found haplotypes within domestic dogs particularly *haplotypes 1*, *2, 3, 4,* and *7* (Table 4), again suggesting that they are representative of the more global set of haplotypes. The common haplotypes are not only represented by more individual dogs but also by more breeds of dogs. This was also observed by Angleby and Savolainen (23) and it holds true for dogs that also have disparate geographic provenances. In forensics, the rarer a haplotype the greater is its degree of forensic informativity for discriminating individuals and/or excluding the others; therefore, commonly found haplotypes are not as forensically valuable.

TABLE 5— *The 417-bp mtDNA HV1 haplotypes among the GenBank sequence.*

Haplotype	Individual dog/Species or breed information/ Geographic location
1	H195 - Massachusetts (MA)
	H196 – MA
	H230 - Michigan (MI)
	H287 – MA
	H318 – MA
	H985 – MA
	AB007392
	AF098147 Do22
	AF098146 Do21
	AF098144 Do19
	AF008147 D26a
2	H334 – MA
	H974 – MA
	H994 – MA
	H1697 - Texas (TX)
	H1698 – TX
	H1742 – TX
	H1743 – TX
	H2108
	H2120
	H2123
	H2124
	H2125
	H2126
	H2140
	H2142
	H2143
	H2150
	H2160
	H2165
	H2168
	AB007393
	AF098139 Do14
	AF064571 Sapsaree - Korea
3	H229 - MI
	H232 - MI
	H290 - MA
	H313 - MA
	H324 - MA
	H333 - MA
	H439 - MA
	H532 - MA
	H915 - MA
	H954 - MA
	H967 - MA
	H2141
	AB007385
	AF098128 Do3
	AF064575 Pekingese - Chinese
	AF008146 D3
4	H227 - MI
	H292 - MA
	H951 - MA
	H980 - MA
	H981 - MA
	H2110
	H2119
	H2132
	H2148
	AB007382
	AF098133 Do8
	AF064578 Tosa Inu - Japan
	AF008148 D4
5	AB007390
6	H984 - MA
	AB007383

TABLE 5— *(Continued)*

Haplotype	Individual dog/Species or breed information/ Geographic location
7	H233 – MI
	H234 - MI
	H289 - MA
	H293 - MA
	H325 - MA
	H326 - MA
	H913 - MA
	H959 - MA
	H1684 - TX
	H1686 - TX
	H1687 - TX
	H1688 - TX
	H1691- TX
	H1699 - TX
	H1700 - TX
	H1705 - TX
	H1729 - TX
	H1740 - TX
	H1741 - TX
	H1745 - TX
	H2114
	H2115
	H2116
	H2118
	H2122
	AB007391
	AB007387
	D83620
	AF064584 Chejudo - Korea
	AF008150 D6b
	AF008143 D24
8	H177B - MA
	H177C - MA
9	H956 - MA
11	H968 - MA
	H988 - MA
	AF098132 Do7
	AF098131 Do6
12	AF064581 Collie
13	H957 - MA
	H975 - MA
	AB007400
14	AB007394
	AF098130 Do5
15	H982 - MA
	Reference dog U96639 Sapsaree - Korea
	AF064573 German Shepherd
16	Coyote
17	H179B - MA
	H180 - MA
	H531 - MA
	AB007403
18	AB007402
19	AB007401
20	AB007395
21	AB007389
22	H167 - MA
	AB007388
	AF008157 D7c
	AF008152 D7a
23	AB007386
24	AB007384
	AF098141 Do16
	AF064572 Sapsaree - Korea
25	AB007381
26	H428
	AB007380

TABLE 5— *(Continued)*

Haplotype	Individual dog/Species or breed information/ Geographic location
27	AF098145 Do20
28	AF098143 Do18
29	AF098142 Do17
30	AF098140 Do15
31	AF098138 Do13
	AF098137 Do12
	AF008155 D8b
	AF008153 D8a
32	AF098136 Do11
33	H393 - Italy
	AF098135 Do10
	AF098134 Do9
34	H328 - MA
	AF098129 Do4
35	AF098127 Do2
36	AF098126 Do1
37	AF064586 Yorkshire Terrier
38	AF064585 Akita
39	AF064582 Chejudo -Korea
40	AF064580 Akita
41	AF064579 Asian Pug
42	AF064577 Chin - Japan
43	H533 - Italy
	AF064576 Shiba Inu - Japan
44	AF064574 Jindo -Korea
45	H443 - MA
	AF064570 Sapsaree -Korea
46	AF064569 Jindo - Korea
47	AF008156 D7b
48	AF008154 D18b
49	AF008151 D21
50	AF008149 D26b
51	AF008145 D18a
52	AF008144 D6a
53	H168
54	H169
55	H231 - MI
56	H296 - MA
57	H426
58	H440 - MA
	H972 - MA
59	H441 - MA
60	H442 - MA
61	H919 Wolf - North West Territories, Canada
62	H958 - MA
63	H970 - MA
	H993 - MA
64	H1649 Coyote - TX
65	H1685 - TX
	H1689 - TX
	H1690 - TX
	H1692 - TX
	H1693 - TX
	H1694 - TX
	H1696A - TX
	H1696B
	H1701 - TX
	H1702 - TX
	H1703 - TX
	H1704 - TX
	H1728 - TX
	H1730 - TX
	H1731- TX
	H1732 - TX
	H1733 - TX
	H1734 - TX
	H2113
	H2127
	H2157
	H2178
	H2151
66	H173 Wolf - Minnesota (MN)

TABLE 6—*AMOVA design and results of the local and global 417-bp haplotypes based on 100,172 permutations ($p < 0.001$).*

Source of Variation	Degrees of Freedom	% Variation
Among sample sets (local and global)	1	−0.88
Within sample sets (local and global)	235	100.88
Total	236	100

The results developed from this canine database demonstrate that there is sufficient variation in the domestic dog HV1 region among individual dogs to make sequence analysis of this region a valuable canine forensic genetics tool. For instance, in a dog attack case, based on the exclusion capacity of the HV1 region used in this study, the probability of excluding a wrongfully accused dog by sequence analysis of this region is 89%. However, in the database of 64 haplotypes, a 4% increase in exclusion capacity, i.e., from 0.89 to 0.93, was obtained despite the increase in sample size from only 36 dogs to 237—almost a sevenfold increase in sampling size. A study by Savolainen et al. (7,9) yielded comparable results from their analysis of 52 of the most common dog breeds in Sweden. With a sampling of 102 Swedish dogs, they were able to calculate an exclusion capacity of 0.88 by analyzing a 257-bp segment of the control region containing HV1. The fact that they analyzed DNA from pure bred dogs could have also contributed to an unrealistically high exclusionary capacity of a region that is less than half the length of sequences analyzed in this study. Another study (23) which used a 573-bp segment within the mtDNA region that overlapped with the region analyzed in this study, and involved dog populations from China, Germany, Japan, Sweden, and the United Kingdom suggested that average exclusion capacities of *c.* 0.90 (range: 0.86–0.95) could be expected for most geographic populations especially where data from various parts of each country are combined. Our estimates of 0.89 based on the local dogs, albeit with sequences 35 bp longer, are within their predicted range. Our estimates of 0.93 with even shorter sequences of 417 bp may have inflated this estimate because of the pooling of dogs of diverse origins, a reflection of Wahlund's principle (19). Therefore, the estimates we have obtained from only the local mixed breed dogs are probably not only more realistic but also more relevant because bioevidence from outbred dogs are more likely encountered in the majority of forensic cases.

The limited discrimination of the mtDNA testing may have been an important reason why canine hair evidence typing results obtained from mtDNA typing techniques have not always been used during criminal trials. In one closed homicide case (State of California v. David Westerfeld, 2002) where canine mitochondrial typing was admitted as evidence because of a lack of nuclear DNA, the matching haplotype occurred at least once in every 11 dogs, i.e., the approximate haplotype frequency is 8.9% (J. L. Halverson, personal communication, http://www.questgen.biz/mitodogs.htm). In contrast to the limited mtDNA exclusion ratios and relatively high haplotype frequencies presented above, in two other closed homicide investigations (2), likelihood ratios of 4.82×10^9 (or one in 4,820,000,000 individuals) and 5.27×10^{14}, respectively, were estimated using canine STRs. With random match probabilities of 2.07×10^{-10} and 1.90×10^{-15}, respectively, there was virtually no opportunity for a false inclusion.

Because mtDNA can be considered as only a single locus and this "locus" does not segregate, the exclusion capacity of mtDNA will never surpass that of nuclear DNA markers, which include the hypervariable STRs. Halverson and Basten (1) showed that the power of discrimination based on 17 canine STRs is *c.* 1 in 10^{12}

among mixed breed dogs. Because of its significantly lower discriminatory power (i.e., the probability of exclusion is 8.6–9.5 per 10 dogs [23]), mtDNA analysis should always be performed in combination with more polymorphic systems where possible.

Nonetheless, as each somatic cell contains up to 1000 mitochondria compared with only two copies of nuclear DNA per cell, typing of shed canine hair using mtDNA analysis is more likely than the analysis of nuclear DNA to produce positive results for degraded samples. The ability to obtain a profile from degraded or trace amounts of forensic samples using mtDNA analysis tends to offset drawbacks of the marker's limited power of exclusion. The discrimination power of this technique will likely improve when combined with sequence analysis of HV2.

Conclusions

This study has found that outbred dogs cannot be reliably assigned to specific breed groups (as described by their owners) using mtDNA typing but that the HV1 region in domestic dog mtDNA can provide discriminating typing information that is likely to be useful for exclusionary purposes. Thus, it is recommended that forensic analysts rely primarily on information from the HV1 region in domestic dog mtDNA analysis for including or excluding an individual animal as a possible source of evidence in a criminal investigation.

Despite the lack of genetic subdivision between the local and global populations, each population contained highly variable haplotypes. Estimates of haplotypic diversity based on local dog populations provide a more accurate picture of local mtDNA structure than samples of dogs combined from different and diverse origins. Therefore, for casework, analysts ought to rely more on databases using local samples as they would be sufficient and perhaps a more relevant dataset to assist in assessing the power of typing results in most forensic investigations. The 60-bp mutational "hotspot" segment in the HV1 region identified in this study may become an important molecular tool for future canine forensic typing applications and consideration should be given to its further evaluation for assaying degraded canid DNA and as a rapid primary screening assay for excluding noncontributors. While longer sequences may be more informative, these regions are only forensically useful if they can be PCR amplified.

Acknowledgments

We are grateful to Joy L. Halverson, D.V.M. (QuestGen Forensics) for her review of a draft of this article. We thank the volunteers who provided samples from their dogs. We express our sincere gratitude to the UC Davis Veterinary Genetics Laboratory (VGL) Wildlife and the Companion Animal Research Units for their contribution of samples used in this study. We also wish to thank the journal's reviewers for their constructive comments.

This research was partially funded by the National Institute of Justice (Grant No. 2004-DN-BX-K007), the Genetic Resource Conservation Program Grant, University Extension, UC Davis, and the Alpha Phi Foundation.

References

1. Halverson J, Basten C. A PCR multiplex and database for forensic DNA identification of dogs. J Forensic Sci 2005;50(2):1–12.
2. D'Andrea F, Fridez F, Coquoz R. Preliminary experiments on the transfer of animal hair during simulated criminal behavior. J Forensic Sci 1998;43(6):1257–8.
3. Kim KS, Lee SE, Jeong HW, Ha JH. The complete nucleotide sequence of the domestic dog (*Canis familiaris*) mitochondrial genome. Mol Phylogenet Evol 1998;10(2):210–20.
4. Okumura N, Ishiguro N, Nakano M, Matsui A, Sahara M. Intra- and interbreed genetic variations of mitochondrial DNA major non-coding regions in Japanese native dog breeds (*Canis familiaris*). Anim Genet 1996;27(6):397–405.
5. Wilson MR, Polanskey D, Butler J, DiZinno JA, Replogle J, Budowle B. Extraction, PCR amplification, sequencing of mitochondrial DNA from human hair shafts. Biotechniques 1995;18:662–9.
6. Vilà C, Savolainen P, Maldonado JE, Amorim IR, Rice JE, Honeycutt RL, et al. Multiple, ancient origins of the domestic dog. Science 1997;276(5319):1687–9.
7. Savolainen P, Rosen B, Holmberg A, Leitner T, Uhlen M, Lundeberg J. Sequence analysis of domestic dog mitochondrial DNA for forensic use. J Forensic Sci 1997;42(4):593–600.
8. Randi E, Lucchini V, Christensen MF, Mucci N, Funk SM, Dolf G, et al. Mitochondrial DNA variability in Italian and East European wolves: detecting the consequences of small population size and hybridization. Conserv Biol 2000;14(2):464–73.
9. Savolainen P, Lundeberg J. Forensic evidence based on mtDNA from dog and wolf hairs. J Forensic Sci 1999;44(1):77–81.
10. Schneider PM, Seo Y, Rittner C. Forensic mtDNA hair analysis excludes a dog from having caused a traffic accident. Int J Legal Med 1999;112(5):315–6.
11. Pfeiffer H, Huhne J, Ortmann C, Waterkamp K, Brinkmann B. Mitochondrial DNA typing from human auxiliary, pubic and head hair shafts—success rates and sequence comparisons. Int J Leg Med 1999;112:287–90.
12. Allen M, Engstrom AS, Meyers S, Handt O, Saldeen T, von Haeseler A, et al. Mitochondrial DNA sequencing of shed hairs and saliva on robbery caps: sensitivity and matching probabilities. J Forensic Sci 1998;43:453–64.
13. Savolainen P, Zhang Y-P, Luo J, Lundeberg J, Leitner T. Genetic evidence for an East Asian origin of domestic dogs. Science 2002;298:1610–3.
14. Higgins D, Thompson T, Gibson T. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl Acids Res 1994;22:4673–80.
15. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 1999;41:95–8.
16. Gundry RL, Allard MW, Moretti TR, Honeycutt RL, Wilson MR, Monson KL, et al. Mitochondrial DNA analysis of the domestic dog: control region variation within and among breeds. J Forensic Sci 2007;52(3):562–72.
17. Kim KS, Jeong HW, Park CK, Ha JH. Suitability of AFLP markers for the study of genetic relationships among Korean native dogs. Genes Genet Sys 2001;76:243–50.
18. Schneider S, Roessli D. ARLEQUIN 2.001: a software for population genetic analysis. Geneva, Switzerland: Laboratory of Genetics and Biometry, University of Geneva, 2000.
19. Nei M. Molecular evolutionary genetics. New York: Columbia University Press, 1987.
20. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of the mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 1993;10:512–26.
21. Smith DG, McDonough J. Mitochondrial DNA variation in Chinese and Indian rhesus macaques (*Macaca mulatta)*. Amer J Primatol 2005;65:1–25.
22. Pereira L, Asch BV, Amorim A. Standardization of nomenclature for dog mtDNA D-loop: a prerequisite for launching a *Canis familiaris* database. Forensic Sci Intl 2004;141:99–108.
23. Angleby H, Savolainen P. Forensic informativity of domestic dog mtDNA control region sequences. Forensic Sci Intl 2005;154:99–110.
24. Parker HG, Kim LV, Sutter NB, Carlson S, Lorentzen TD, Malek TB, et al. Genetic structure of the purebred domestic dog. Science 2001;291:1160–4.
25. Fogel B. The encyclopedia of the dog. New York: Dorling Kindersley Publishing, 1995.
26. Fogel B. Eyewitness companions—dogs. New York: Dorling Kindersley Publishing, 2006.
27. Crowley J, Adelman B, editors. The complete dog book: official publication of the American Kennel Club. New York: Howell House, 1998.

28. Neigel JE, Avise JC. Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. In: Karlin S, Nevo E, editors. Evolutionary processes and theory. New York: Academic Press, 1986;513–34.

Appendix 1—GenBank sequences used in this study

The GenBank Accession Numbers for the lab-generated haplotypes range from EF122413 to EF122428. The domestic dog *(Canis lupus familiaris)* reference sequence (GenBank Accession Number U96639 [Sapsaree—a Korean breed]), coyote (*Canis latrans*) sequences (AY172674 and AY240094.1 [H 1649]), wolf (*Canis lupus*) sequences (AY240073.1 [H 919], and AY240155.1 [H 173]) and domestic dog sequences (AB007392, AF098147 (Do22), AF098146 (Do21), AF098144 (Do19), AF008147 (D26a), AB007393, AF098139 (Do14), AF064571 (Sapsaree), AB007385, AF098128 (Do3), AF064575 (Pekingese—a Chinese breed), AF008146 (D3), AB007382, AF098133 (Do8), AF064578 (Tosa Inu—a Japanese breed), AF008148 (D4), AB007390, AB007383, AB007391, AB007387, D83620, AF064584 (Chejudo—a Korean breed), AF008150 (D6b), AF008143 (D24), AF098132 (Do7), AF098131 (Do6), AF064581 (Collie), AB007400, AB007394, AF098130 (Do5), AF064573 (German Shepherd), AB007403, AB007402, AB007401, AB007395, AB007389, AB007388, AF008157 (D7c), AF008152 D7a, AB007386, AB007384, AF098141 Do16, AF064572 (Sapsaree), AB007381, AB007380, AF098145 (Do20), AF098143 (Do18), AF098142 (Do17), AF098140 (Do15), AF098138 (Do13), AF098137 (Do12), AF008155 (D8b), AF008153 (D8a), AF098136 (Do11), AF098135 (Do10), AF098134 (Do9), AF098129 (Do4), AF098127 (Do2), AF098126 (Do1), AF064586 (Yorkshire Terrier), AF064585 (Akita—Japanese breed), AF064582 (Chejudo), AF064580 (Akita), AF064579 (Asian Pug), AF064577 (Chin—a Japanese breed), AF064576 (Shiba Inu—a Japanese breed), AF064574 (Jindo—a Korean breed), AF064570 (Sapsaree), AF064569 (Jindo), AF008156 (D7b), AF008154 (D18b), AF008151 (D21), AF008149 (D26b), AF008145 (D18a), AF008144 (D6a), AY240030.1 (H 167), AY240031.1 (H 168), AY240032.1 (H 169), AY240033.1 (H 177B), AY240034.1 (H 177C), AY240035.1 (H 179B), AY240036.1 (H 180), AY240037.1 (H 195), AY240038.1 (H 196), AY240039.1 (H 227), AY240040.1 (H 229), AY240041.1 (H 230), AY240042.1 (H 231), AY240043.1 (H 232), AY240044.1 (H 233), AY240045.1 (H 234), AY240046.1 (H 287), AY240047.1 (H 289), AY240048.1 (H 290), AY240049.1 ((H 292), AY240050.1 (H 293), AY240051.1 (H 296), AY240052.1 (H 313), AY240053.1 (H 318), AY240054.1 (H 324), AY240055.1 (H 325), AY240056.1 (H 326), AY240057.1 (H 328), AY240058.1 (H 333), AY240059.1 (H 334), AY240060.1 (H 393), AY240061.1 (H 426), AY240062.1 (H 428), AY240063.1 (H 439), AY240064.1 (H 440), AY240065.1 (H 441), AY240066.1 (H 442), AY240067.1 (H 443), AY240068.1 (H 531), AY240069.1 (H 532), AY240070.1 (H 913), AY240071.1 (H 533), AY240072.1 (H 915), AY240074.1 (H 951), AY240075.1 (H 954), AY240076.1 (H 956), AY240077.1 (H 957), AY240078.1 (H 958), AY240079.1 (H 959), AY240080.1 (H 967), AY240081.1 (H 968), AY240082.1 (H 970), AY240083.1 (H 972), AY240084.1 (H 975), AY240085.1 (H 974), AY240086.1 (H 980), AY240087.1 (H 981), AY240088.1 (H 982), AY240089.1 (H 984), AY240090.1 (H 985), AY240091.1 (H 988), AY240092.1 (H 993), AY240093.1 (H 994), AY240095.1 (H 1684), AY240096.1 (H 1685), AY240097.1 (H 1686), AY240098.1 (H 1687), AY240099.1 (H 1688), AY240100.1 (H 1689), AY240101.1 (H 1690), AY240102.1 (H 1691), AY240103.1 (H 1692), AY240104.1 (H 1693), AY240105.1 (H 1694), AY240106.1 (H 1696A), AY240107.1 (H 1696B), AY240108.1 (H 1697), AY240109.1 (H 1698), AY240110.1 (H 1699), AY240111.1 (H 1700), AY240112.1 (H 1701), AY240113.1 (H 1702), AY240114.1 (H 1703), AY240115.1 (H 1704), AY240116.1 (H 1705), AY240117.1 (H 1728), AY240118.1 (H 1729), AY240119.1 (H 1730), AY240120.1 (H 1731), AY240121.1 (H 1732), AY240122.1 (H 1733), AY240123.1 (H 1734), AY240124.1 (H 1740), AY240125.1 (H 1741), AY240126.1 (H 1742), AY240127.1 (H 1743), AY240128.1 (H 1745), AY240129.1 (H 2108), AY240130.1 (H 2110), AY240131.1 (H 2113), AY240132.1 (H 2114), AY240133.1 (H 2115), AY240134.1 (H 2116), AY240135.1 (H 2118), AY240136.1 (H 2119), AY240137.1 (H 2120), AY240138.1 (H 2122), AY240139.1 (H 2123), AY240140.1 (H 2124), AY240141.1 (H 2125), AY240142.1 (H 2126), AY240143.1 (H 2127), AY240144.1 (H 2132), AY240145.1 (H 2140), AY240146.1 (H 2141), AY240147.1 (H 2142), AY240148.1 (H 2143), AY240149.1 (H 2150), AY240150.1 (H 2157), AY240151.1 (H 2160), AY240152.1 (H 2165), AY240153.1 (H 2168), AY240154.1 (H 2178), AY240156.1 (H 2148), AY240157.1 (H 2151)).

Additional information and reprint requests:
Sree Kanthaswamy, Ph.D.
Department of Anthropology
University of California
One Shields Avenue
Davis, CA 95616
Email: skanthaswamy@ucdavis.edu