J Forensic Sci, Mar. 2005, Vol. 50, No. 2
Paper ID JFS2004207
Available online at: www.astm.org

Joy Halverson,[1] *D.V.M. and Christopher Basten,*[2] *Ph.D.*

A PCR Multiplex and Database for Forensic DNA Identification of Dogs

ABSTRACT: Animal-derived trace evidence is a common finding at crime scenes and may provide an important link between victim(s) and suspect(s). A database of 558 dogs of pure and mixed breeds is described and analyzed with two PCR multiplexes of 17 microsatellites. Summary statistics (number of alleles, expected and observed heterozygosity and power of exclusion) are compared between breeds. Marked population substructure in dog breeds indicates significant inbreeding, and the use of a conservative θ value is recommended in likelihood calculations for determining the significance of a DNA match. Evidence is presented that the informativeness of the canine microsatellites, despite inbreeding, is comparable to the human CODIS loci. Two cases utilizing canine DNA typing, *State of Washington v. Kenneth Leuluaialii and George Tuilefano* and *Crown v. Daniel McGowan*, illustrate the potential of canine microsatellite markers for forensic investigations.

KEYWORDS: forensic science, canine, DNA typing, polymerase chain reaction, microsatellites, STR, database, likelihood ratio, CODIS

Animals live in close contact with humans and may witness or be victims of human crimes. If pets are killed or injured at a crime scene, they may leave significant biological evidence. Even as passive witnesses, pets can provide important information to the observant investigator by the transfer of trace evidence. Dog and cat hairs are ubiquitous in the homes, cars, and on the clothing of pet owners and are readily transferred to and from crime scenes. Animal hairs can be analyzed with techniques similar to those used for human hairs. Microsatellite loci, successfully amplified from dog hairs and other sample types, have contributed to over 20 criminal investigations since 1996. Evidence from these investigations has been admitted in nine criminal trials in California, Florida, Illinois, Oklahoma, Washington, Iowa, Pennsylvania, and Great Britain. Depending on the trial dates, the significance of the DNA matches in these cases was based on population allele frequency data from 438–558 dogs. We report herein on frequency data from 558 dogs of pure and mixed breed origin. The data are compared to a large, privately owned data set of 9548 dogs analyzed with the same microsatellite loci as well as two human forensic databases.

Background

Animal Microsatellites in the Scientific Literature

In 1989, microsatellites were first described in the human genome as a string of dinucleotide repeats with flanking unique DNA sequences amenable to amplification as a single locus (1). The utility of these markers for gene mapping was readily apparent, and they were soon described in many other species. Horses and cattle were the focus of early development; the first report of canine microsatellite loci appeared in 1993 (2). Early investigators were concerned that, in contrast to humans, the inbreeding common in domestic animals would reduce the variation found in microsatellites. However, as more loci were developed, the polymorphism and ease of use demonstrated both their utility and their high discrimination. The animal identification community quickly phased out blood typing technologies in favor of DNA typing just as the human parentage and forensic communities had done.

Microsatellites with tetranucleotide repeat structure were first reported in dogs in 1996 (3). Genetic research on dogs using microsatellites has continued with investigations into the genetic variation found in dog breeds (4–9) and the publication of several genetic maps (10–13). Microsatellites have been used for linkage mapping in the investigation of genetic diseases in dogs (14). In 1997 the American Kennel Club (AKC) and the United Kennel Club began pilot programs using microsatellites for pedigree verification (15). Zoogen, a commercial laboratory, had identified 20 tri- and tetranucleotide repeat STR (short tandem repeat) loci having characteristics suitable for robust, cost-effective DNA typing across many dog breeds (16). At that time, tri- and tetranucleotide microsatellites were uncommon in animal STR markers; bovine and equine verification programs were based on dinucleotide repeat markers. Zoogen followed the example set by the human forensic community and developed tri- and tetranucleotide microsatellites to reduce the stutter artifacts characteristic of dinucleotide repeats and to enhance standardization of allele scoring. After purchasing Zoogen, Applied Biosystems (ABI) developed two canine STR kits (Stockmarks® Canine I and II) to be used with their fluorescent analysis platforms. The kit loci consisted of one published tetranucleotide locus ($CATA_1$) (17), one trinucleotide and twelve tetranucleotide Zoogen loci, and three public domain tetranucleotide loci developed at the Fred Hutchinson Cancer Center (3).

Canine Microsatellites and Forensics

Despite their utility for the general scientific community, there are relatively few publications in the peer-reviewed literature on forensic identification with canine microsatellites. The use of microsatellites for the detection of wolf-dog hybrids has been reported,

[1] Joy Halverson, D.V.M., QuestGen Forensics, 29280 Mace Bld., Davis, CA 95616.
[2] Christopher Basten, Ph.D., Bioinformatics Research Center, North Carolina State University, Raleigh, NC 27695-7566.

Received 18 May 2004; and in revised form 2 Oct. 2004; accepted 2 Oct. 2004; published 2 Feb. 2005.

an application of note to law enforcement (18,19). Shutler et al. (20) reported the first use of canine microsatellites in a homicide investigation in Canada in 1991 in which an elderly man and his dog were killed by blunt trauma. The suspect had mixed bloodstains on his clothing, but the amount was insufficient for the RFLP-VNTR testing of the time. In 1996, the case was reopened, and the stains were tested with both human and ABI canine microsatellites. The human stains matched the victim, and the canine stains matched the victim's dog. There are two reports of canine STR typing to identify the canine perpetrator(s) in attacks on people (one fatal) and another on zoo animals (21–23). Another report describes the use of DNA typing to identify a dog that caused a traffic accident (24). Two of these reports originated in Hungary; the ABI Stockmarks® Canine I kit was used for the analyses. At the European Academy of Forensic Sciences meeting in Istanbul, Turkey in 2003, Dr. Andreas Hellman of the Kriminaltechnisches Institut in Germany reported ongoing investigation into the use of the Canine I kit for forensic analysis of animal evidence (25).

Canine Microsatellites and Population Substructure

To use canine microsatellites effectively for forensic identification, it is imperative to assess the population substructure of dogs. Purebred dogs, estimated to comprise approximately half of the American dog population, do not mate randomly and could be expected to exhibit significant substructure. Even mixed breed dogs may show some degree of population substructure from recent purebred ancestry. If population substructure exists, then matching genotypes between two different dogs is not entirely random. In order to estimate the significance of a DNA match, the association of alleles within loci and between loci must be addressed.

The Hardy Weinberg proportion describes the relationship between the allele frequencies and genotype frequencies at a single locus. Suppose alleles A_i and A_j have frequencies p_i and p_j, respectively. If an individual has two copies of the same allele at a locus of interest, then that individual is said to be a homozygote. The population frequency of an A_i A_i homozygote is written P_{ii} and will have a frequency of p_i^2 predicted from the Hardy Weinberg proportions. An individual with different alleles at the locus would have genotype A_iA_j. The frequency of the A_iA_j genotype would be $P_{ij} = 2p_ip_j$. This relationship between the allele frequencies and the genotype frequency depends upon the assumption that the alleles come together independently. If population substructure is present, then an analysis of the degree of disequilibria provides an adjustment value θ that can be applied in calculating statistics for DNA tests such as identity and parentage. The correction value θ takes into account the history of the populations. In humans, the value of $\theta = 0.01$ is conservative and reflects the minimal degree of population substructure in most human populations.

Linkage disequilibrium refers to associations of alleles or genotypes between loci. The loci do not necessarily have to be physically linked to show some association (and loci on the same chromosome may show no association if far enough apart). If there is equilibrium across loci, then the genotypic frequencies can be multiplied to get a profile frequency. The reciprocal of this number ($X = 1/P$) is often used and stated as the inverse cumulative match probability of "1 in X."

Rather than a simple inverse match probability, the 1996 NRC report suggests that the weight of a DNA match between an evidence sample and a reference sample can be expressed as a likelihood ratio. The likelihood ratio requires the formulation of two contrasting hypotheses for the DNA match. The first hypothesis forms the numerator of the ratio, and it usually argues that the DNA profiles match because they came from the same source. This would generally be the prosecution's explanation of the match. An example of a second hypothesis (the defense explanation) might be that the DNAs did not come from the same source but matched by random chance. The defense may have other explanations of a DNA match; a likelihood ratio can be devised for any alternative explanation.

The likelihood ratio makes the assumptions that DNA samples from the same individual will match and that two different individuals have independent probabilities of having the profile (individuals are unrelated). If these assumptions are true, then the inverse match probability and the likelihood calculation are equal. However, population substructure impacts the second assumption and must be accounted for in canine DNA matches. The 1996 NRC report recommends that the genotype frequency calculations and the resulting likelihood calculation be modified to address substructure. For a heterozygote with alleles A_i and A_j with frequencies p_i and p_j

$$P(A_iA_j \mid A_iA_j) = 2[\theta + (1-\theta)p_i][\theta + (1-\theta)p_j]/[(1+\theta)(1+2\theta)] \quad (1)$$

And for an A_iA_i homozygote with the A_i population frequency p_i

$$P(A_iA_i \mid A_iA_i) = [2\theta + (1-\theta)p_i][3\theta + (1-\theta)p_i]/[(1+\theta)(1+2\theta)] \quad (2)$$

We account for the occurrence of allelic dropout by summing Eqs 1 and 2 over the dropped allele. If the known dog is A_iA_j, and the evidence is A_i with possible allelic dropout, then Eq 1 becomes

$$P(A_i - \mid A_iA_j) = [\theta + (1-\theta)p_i][4\theta + (1-\theta)(2-p_i)]/[(1+\theta)(1+2\theta)]. \quad (3)$$

If the known dog is a homozygote A_iA_i, then Eq 2 becomes

$$P(A_i - \mid A_iA_i) = [2\theta + (1-\theta)p_i][3\theta + (1-\theta)(2-p_i)]/[(1+\theta)(1+2\theta)] \quad (4)$$

If the evidence sample has a limited amount of DNA, the profile obtained may exhibit allelic dropout. Currently there is no quantitation assay for small amounts of canine DNA. In casework with telogen hairs, amplification is attempted without the opportunity to optimize template input. Research with human forensic kits has shown that allele dropout is common with template input in the range of 50–250 pg. (26,27) While further optimization of the Stockmarks® kit may minimize allele dropout, the impact of potential allelic dropout in forensic cases is currently handled by appropriate probability calculations. Loci that are clearly heterozygous and match the known dog allow the use of Eq 1. Single peaks may indicate that allelic dropout is a concern, and it is conservative to use Eqs 3 and 4 depending on the genotype of the known individual.

Forensic Casework

In 1996, PE Zoogen assisted the Royal Canadian Mounted Police in the case *Crown v. William Faulconer* as reported by Shutler et al. (20). In 1997, a similar case (*State of Washington v. Kenneth Leuluaialii and George Tuilefano*) occurred in Seattle, Washington. The suspects were accused in the shooting deaths of a young couple and their dog during a home invasion. Clothing articles from the suspects had been examined and were found to have non-human blood stains. PE Zoogen's analysis with Stockmarks® Canine 1

loci showed a match between the victims' dog and the stains. The statistical significance of the match was based on the initial Zoogen study of 438 dogs (henceforth called the "Zoogen database").

In 2003, the first author assisted the West Yorkshire Police of Leeds, Great Britain in a homicide investigation. Several men had invaded the home of Brian Keating and abducted him. His slain body was found posed in a church graveyard the next day. Investigators believed that dog hairs found on the victim's clothes had been transferred from a van used during the incident. The hairs matched a dog owned by Daniel McGowan, who was subsequently convicted along with three other suspects (*Crown v. Daniel McGowan*). The statistical significance of the match was based on the expanded Zoogen database of 558 dogs.

In October 2003, an appellate court ruled that admitting the evidence of a canine DNA match in the *State of Washington v. Kenneth Leuluaialii and George Tuilefano* trial was inappropriate due to the lack of a Frye hearing. (The conviction was upheld, however.) The court found that, at the time of the hearing, there was insufficient publication on the particular DNA markers used to consider canine DNA typing as accepted in the scientific community. The appellate authors also expressed concern over the degree of allele sharing (5 out of 18 possible alleles) between the dog in the Seattle case and the canine victim in the *Crown v. William Faulconer* case. *State of Washington v. Kenneth Leuluaialii and George Tuilefano* was the first case in the United States to use a canine DNA match as evidence. Since that time similar evidence has been admitted (several with Frye hearings) in nine assault and homicide trials nationwide and in Great Britain.

In 1996, the National Research Council had published its second report on standard procedures for the calculation of match probabilities and likelihood ratios (28). The Zoogen database (at the time of *State of Washington v. Kenneth Leuluaialii and George Tuilefano* consisting of 438 dogs) was evaluated for population substructure according to those recommendations. The analysis substantiated that the selective breeding of dogs affects canine population structure and hence the genotype frequencies. Testimony in the case included the use of the appropriate theta value and likelihood ratios. Herein we describe the current Zoogen database and compare it to a large canine database (n = 9548) and two human databases of sizes similar to the Zoogen database. We also describe canine population substructure and the implications for reporting the statistical significance of a DNA match of canine evidence in two criminal investigations.

Materials and Methods

Sample Collection

A total of 558 dogs were included in this study (Table 1). A total of 395 samples were provided by California Veterinary Diagnostic Laboratories (now IDEXX) in Sacramento, California. The samples had been submitted for routine diagnostic testing by veterinarians throughout northern California and southern Oregon. The breed, as indicated by the veterinarian, was the only information accompanying the sample. The samples belong to sixteen breeds (n = 17–32/breed) and mixed breed dogs (n = 40). DNA was extracted from these samples using the Qiagen BloodAmp Kit (Qiagen, Valencia, California) according to the manufacturer's instructions. In addition, Dr. Candy Gaiser (Zoogen, Inc., Davis, California) provided extracted DNA samples from unrelated dogs collected throughout the United States. These consisted of a panel of 96 purebred dogs, each from a different breed, panels of 19 unrelated Whippets and 19 unrelated Greyhounds, and 29 mixed breed dogs.

TABLE 1—*Sample sizes of breed groups in the Zoogen study.*

Breed	Sample Size
Akitas	18
Australian Shepherds	19
Chihuahuas	28
Dachshunds	28
Dalmatians	17
English Springer Spaniels	26
Golden Retrievers	20
Greyhounds	19
Lhasa Apsos	32
Chinese Shar Peis	18
Pomeranians	20
Poodles	30
Siberian Huskies	17
Shih Tzus	26
Whippets	19
Yorkshire Terriers	32
All Breed Panel*	96
Mixed Breed Dogs	69
Pit Bull Terriers	24
Total	558

* The All Breed Panel consists of one dog each from 96 different pure breeds.

PCR Amplification and Fragment Analysis

Initially, target DNAs were amplified in three PCR multiplexes of 7, 4, and 6 loci and analyzed in two gel lanes or capillaries. Later, Applied Biosystems kit reagents consisting of two PCR multiplexes of 10 and 7 loci [Stockmarks® for Dogs Canine I (Part # 4307481) and II, respectively] were used. Locus primers, repeat type, size range, and chromosomal location are listed in Table 2. Mixed breed dogs, the all-breed panel, and the Pit Bull Terriers were tested with PEZ 20; the original breed panels were tested with PEZ 18 instead. PEZ 18 was replaced with PEZ 20 because the former locus has a large base pair size range, making it difficult to multiplex. Amplification products were electrophoresed on ABI 377s (and on ABI 3100s later) and analyzed using GeneScan 3.0 software and Genotyper 2.0 software. An allelic ladder is not currently provided with the Stockmarks® Canine Kit. Instead, profiles from the positive control dog DNA included with the kit were used to offset allele bins. Alphabetic letters were used for standardized allele designations to make parentage results of commercial tests more accessible to dog owners.

Data Analysis

Allele frequencies and observed heterozygosity (OH) were calculated with PopGene software (29). Expected heterozygosity (EH) (30), Polymorphism Information Content (PIC) (31), locus probability of exclusion (PE) (32), and Match Probability (28) were calculated according to the references noted. Cumulative PE was calculated as follows:

If e_i is the probability of exclusion (PE) for locus i, then the cumulative PE over the first l loci equals

$$1 - \prod_{i=1}^{l} (1 - e_i)$$

The cumulative inverse match probability (without a substructure adjustment) is the reciprocal of the product of the MPs over loci. Population substructure was analyzed with GDA software (33) by examining the associations within loci (Hardy-Weinberg disequilibria) and between loci (linkage disequilibria). Inbreeding

TABLE 2—*Primers for PCR amplification of microsatellites in Stockmarks Canine I & II.*

Stockmarks	Locus	Forward Primer	Reverse Primer	Repeat Motif	*Size Range (bp)		†Map Location	Reference
Canine I	$CATA_1$	GG CTG TCA CTT TTC CCT TTC	CAC CAC AAT CTC TCT CAT AAA TAC	CATA	95	136	‡ cfa 7	17
Canine I	PEZ03	CA CTT CTC ATA CCC AGA CTC	CAA TAT GTC AAC TAT ACT TC	AAG	95	154	cfa 19	16
Canine I	PEZ05	GC TAT CTT GTT TCC CAC AGC	GTC ACT GTA TAC AAC ATT GTC	AAAG	97	121	cfa 12	16
Canine I	PEZ06	AT GAG CAC TGG GTG TTA TAC	ACA CAA TTG CAT TGT CAA AC	AAAT	166	215	cfa 27	16
Canine I	PEZ08	TA TCG ACT TTA TCA CTG TGG	ATG GAG CCT CAT GTC TCA TC	AAAT	230	260	cfa 17	16
Canine I	PEZ12	GT AGA TTA GAT CTC AGG CAG	GTA GGT CCT GGT AGG GTG TGG	AAAG	250	317	cfa 3	16
Canine I	PEZ20	CC TAA ATT AGA GGT CTA ACC	GTA AGC GGG AAT GTG CTC CTC	AAAT	152	202	§Unmapped	16
Canine I	FHC2010	AA ATG GAA CAG TTG AGC ATG C	CCC CTT ACA GCT TCA TTT TCC	ATGA	220	248	cfa 24	3
Canine I	FHC2054	GC CTT ATT CAT TGC AGT TAG GG	ATG CTG AGT TTT GAA CTT TCC C	GATA	140	184	cfa 12	3
Canine I	FHC2079	CA GCC GAG CAC ATG GTT T	ATT GAT TCT GAT ATG CCC AGC	GGAT	263	299	cfa 24	3
Canine II	PEZ10	CT TCA TTG AAG TAT CTA TCC	CCT GCC TTT GTA AAT GTA AG	AAAG	230	330	cfa 14	16
Canine II	PEZ11	AT TCT CTG CCT CTC CCT TTG	GTG TGG ATA ATC TCT TCT GTC	AAAG	123	175	cfa 8	16
Canine II	PEZ13	AG TCT GGT GAT TTA ATT CGG	GTC TAG TCC CCA GTC TAG TTC ACT GCC C	AAAG	171	322	cfa 4	16
Canine II	PEZ15	CT GGG GCT TAA CTC CAA GTT C	CAG TAC AGA GTC TGC TTA TC	AAAG	193	284	||Unassigned	16
Canine II	PEZ16	GC TCT TTG TAA AAT GAC CTG	GTG GGA ATC GTC CTA AAA CCC	AAAG	263	334	cfa 27	16
Canine II	PEZ17	CT AAG GGA CTG AAC TTC TCC	GTG GAA CCT GCT TAA GAT TC	AAAG	196	245	cfa 4	16
Canine II	PEZ21	AA CCG GTT GTG ATT TCT GGG	GTC TGT GTC ATT AGT GAC ATC	AAAT	71	109	Unmapped	16

* Size ranges shown are from combined Zoogen and AKC study data.
† Ref (11) and website:http://www.fhcrc.org/science/dog_genome/breen2001/.
‡ cfa refers to Canis familiaris.
§ Unmapped loci were not included in mapping study referenced.
|| Unassigned loci were included in the mapping study referenced but could not be assigned.

coefficients were estimated for each locus. Data from the All Breed Panel were analyzed alone and as part of the entire group (n = 558); data from individuals of the All Breed Panel were not combined with breed groups. A population substructure value θ was estimated using standard methods (30) and three levels of θ recommended as lower, middle, and upper confidence limits for likelihood calculations. The application of the θ values in the likelihood calculations (34) for canine DNA matches in two homicide cases is illustrated.

Database Comparisons

The American Kennel Club collected 9548 samples from 108 breeds at Parent Club dog shows. These shows are national events held once or twice annually. Sample submission was on a voluntary basis; there was no effort to screen for close relatives in the sample groups. Summary statistics (number of alleles, expected heterozygosity) from breeds common to the Zoogen database were used for comparison (35).

We also compared the discrimination power of the Stockmarks® canine loci to differentiate canine individuals with the discrimination power of the human forensic CODIS microsatellites by determining the proportion of individuals with matching locus genotypes. The databases compared included the Zoogen Study (n = 558), a published FBI database (n = 622) (36), and a human database (n = 824) provided courtesy of Christine S. Tomsey, DNA Laboratory Manager, Pennsylvania State Police Bureau of Forensic Science and Criminal Identification, Forensics Division, Greensburg, PA 15601. The FBI database included the 13 CODIS loci, and the Penn database included the 13 CODIS loci plus Penta D and Penta E. Both databases are comprised of Caucasian, African American, and Hispanic populations. For each database, $n(n-1)/2$ pairwise comparisons were performed. For each comparison, the number of loci that have the same genotype between every pair of individuals was counted. Then the counts were divided by the number of comparisons to determine the frequency of individuals in each database sharing zero or more locus genotypes.

Results

Breed Diversity Estimates from the Zoogen Study

In the single breed panels, the mean number of alleles across loci ranged from 5.31 in Whippets and Greyhounds to 7.33 in Chihuahuas (Fig. 1). Other summary statistics (EH, OH, PIC, PE) based on allele frequencies follow similar trends with some breeds showing scores approaching mixed breed dogs (e.g., Australian Shepherds, Chihuahuas, Poodles, English Springer Spaniels, and Lhasa Apsos). Akitas and Whippets showed the lowest scores in summary statistics (Fig. 2). Table 3 (available at http://statgen.ncsu.edu/forensics/canine/Table3.xls) lists the

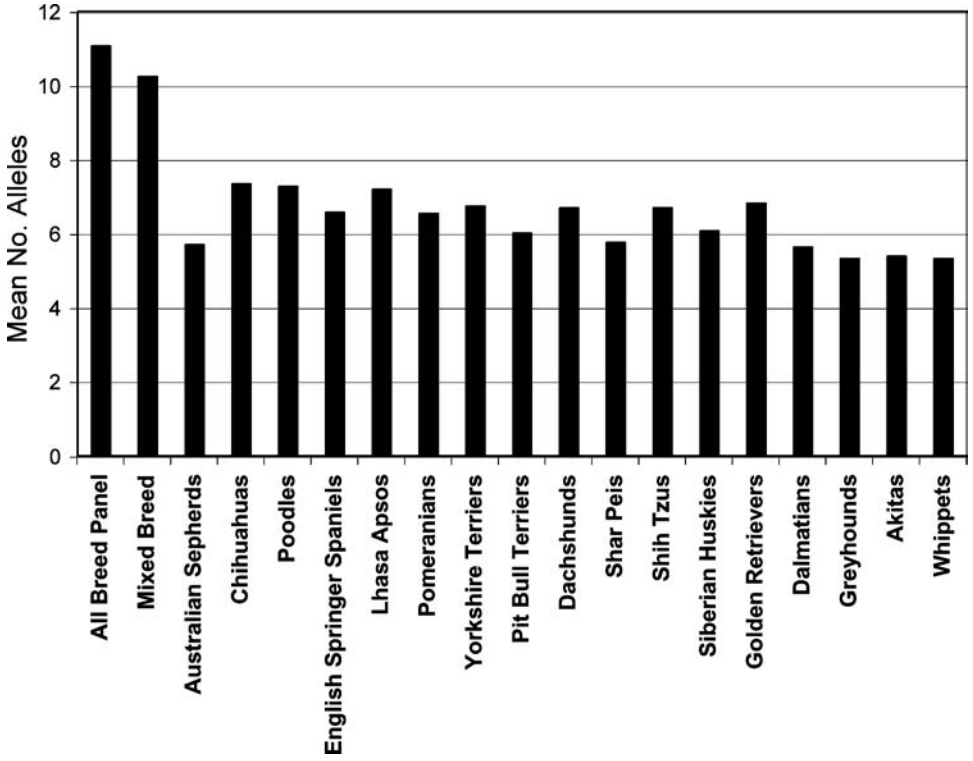

FIG. 1—*Mean number of alleles (averaged across loci) in the All Breed Panel, mixed breed dogs, and the 17 breed panels. Locus PEZ 18 was not included in mean calculation.*

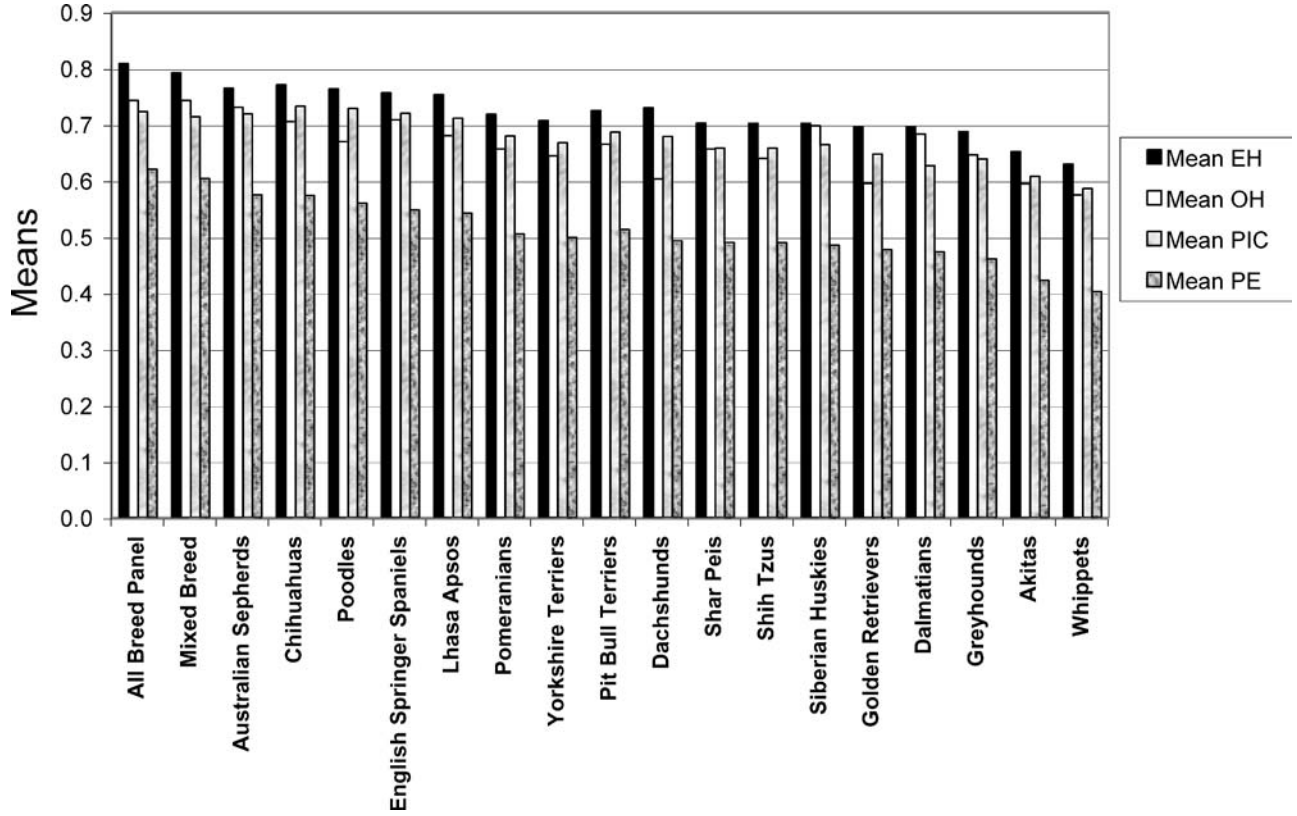

FIG. 2—*Mean Expected Heterozygosity (EH), mean Observed Heterozygosity (OH), mean Polymorphism Information Content (PIC), and mean Power of Exclusion (PE) (averaged across loci) in the All Breed Panel, mixed breed dogs, and the 17 breed panels. Locus PEZ 18 was not included in mean calculations.*

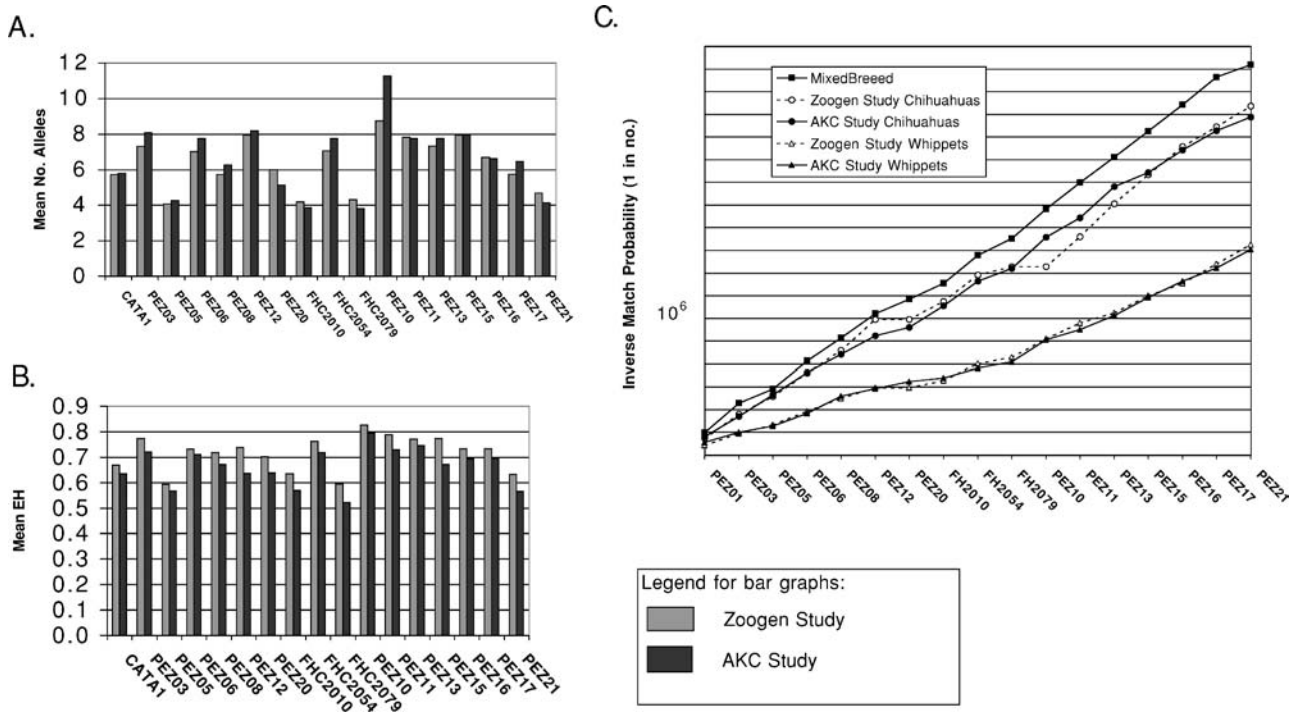

FIG. 3—*Comparison of the mean number of alleles (a), mean expected heterozygosity (both averaged across breeds) (b), and the inverse match probability (c) in breeds common to the Zoogen and AKC study. The Y-axis in (c) has a logarithmic scale. The two studies show similar trends in locus statistics.*

TABLE 4—*Comparison of mean expected heterozygosity (EH) averaged over loci in the Zoogen and AKC studies.*

Breed	*Sample n	AKC Mean EH	Zoogen Mean EH	†Lower Limit	†Upper Limit	‡AKC Included?
Akitas	117	0.649	0.677	0.599	0.756	yes
Australian Shepherds	151	0.682	0.788	0.725	0.851	no
Chihuahuas	105	0.747	0.743	0.674	0.811	yes
Shar Peis	104	0.721	0.734	0.654	0.813	yes
Dachshunds	130	0.720	0.681	0.595	0.768	yes
Dalmatians	96	0.672	0.721	0.636	0.807	yes
Golden Retrievers	116	0.611	0.675	0.593	0.757	yes
Greyhounds	13	0.544	0.707	0.637	0.776	no
Lhasa Apsos	117	0.689	0.768	0.719	0.817	no
Pomeranians	124	0.692	0.701	0.622	0.779	yes
Poodles	186	0.717	0.735	0.667	0.802	yes
Shih Tzus	41	0.612	0.724	0.659	0.789	no
Siberian Huskies	149	0.635	0.735	0.66	0.811	no
Whippets	147	0.596	0.653	0.573	0.732	yes
Yorkshire Terriers	56	0.654	0.682	0.598	0.765	yes

* Sample number in AKC study.
† Limits of 95% confidence interval.
‡ Is the AKC Study mean EH included in the 95% confidence interval calculated from the mean EH of the Zoogen Study for a given breed?

summary statistics for 17 breed panels, the All Breed panel (consisting of one dog each per 96 breeds), and mixed breed dogs. In non-forensic applications, such as routine parentage testing, the 10plex loci are analyzed first; 7plex loci are occasionally needed when putative parents are closely related. The cumulative PE and MP calculations shown in Table 3 are based entirely on allele frequencies and do not include adjustments for population substructure.

Locus Performance in the Zoogen Study and the AKC Parent Club Breed Study

In order to compare data from the Zoogen study and the AKC Parent Club Study, only the fifteen breeds in common between the two studies were used, narrowing the Zoogen study group to 355 dogs and the AKC study group to 1655 dogs. In the Zoogen study, the mean number of alleles, averaged across the breed panels, ranged from 4.06 at PEZ5 to 8.73 at PEZ10 (Fig. 3*a*). In the AKC study, the mean number of alleles ranged from 4.27 at PEZ5 to 11.27 at PEZ10. Loci showing low mean numbers of alleles are FHC2010 (4.18 in the Zoogen study, 3.87 in the AKC study) and FHC2079 (4.29 in the Zoogen study, 3.80 in the AKC study). Overall, the mean number of alleles seen in the narrowed AKC study was only 6% greater than in the Zoogen study group, despite including over fourfold as many dogs.

Unlike the mean number of alleles, the Zoogen study had overall higher values for mean EH (Fig. 3*b*, Table 4), PIC, and PE (data not shown) than the AKC Parent Club study. While higher, these

TABLE 5—*The frequency of individuals sharing the same locus genotype at the indicated number of loci in the Zoogen canine database and two human databases (FBI and Pennsylvania State Bureau of Forensic Science).*

Matching Loci	No. (Canine)	Freq. (Canine)	No. (FBI)	Freq. (FBI)	No. (Penn)	Freq. (Penn)
0	51 465	0.331	74 049	0.383	122 944	0.363
1	58 351	0.375	72 872	0.377	129 606	0.382
2	31 381	0.202	34 107	0.177	63 435	0.187
3	10 865	0.070	9 800	0.051	18 701	0.055
4	2 745	0.018	1 968	0.010	3 757	0.011
5	481	0.003	301	0.002	546	0.002
6	91	0.001	30	>0.001	77	>0.001
7	15	>0.001	4	>0.001	7	>0.001
8	5	>0.001	0	0	2	>0.001
9	2	>0.001	0	0	1	>0.001
10	1	>0.001	0	0	0	0
11	1	>0.001	0	0	0	0
12	0	0	0	0	0	0
13	0	0	0	0	0	0
14	0	0			0	0
15	0	0			0	0
16	0	0				
17	0	0.000				
18	0	0.000				
Total Number	155 403		193 131		339 076	

TABLE 6—*The numbers of loci showing Hardy-Weinberg disequilibrium (HW) and linkage disequilibrium (LD) in dog breeds.*

Breed	HW	LD
Akitas	4	7
Australian Shepherds	2	4
Chihuahuas	4	4
Chinese Shar Peis	3	4
Dachshunds	3	4
Dalmatians	2	2
English Springer Spaniels	4	7
Golden Retrievers	3	6
Greyhounds	2	6
Lhasa Apsos	3	7
Mixed Breed Dogs	3	7
Pit Bull Terriers	3	7
Pomeranians	3	4
Poodles	5	3
Shih Tzus	2	6
Siberian Huskies	2	5
Whippets	3	7
Yorkshire Terriers	6	13
Total	57	103
Expected at 95% confidence	15.3	122.4
Number of tests performed	306	2448

measures of locus informativeness show the same trends between loci, with PEZ 5, PEZ 21, FHC 2010, and FHC 2079 having relatively low scores; PEZ 3, 6, 10, 11, 12, 13, 15, and FHC 2054 having high scores; and PEZ 8, 16,17, and 20 having intermediate scores (Fig. 3*b*). Figure 3*c* compares the cumulative inverse match probability of two breeds common to both studies and to the mixed breed dogs in the Zoogen study. The match probabilities shown were calculated using the most common alleles for each locus in the given breeds so as to represent a "worst case scenario" rather than an actual case based on an individual. As the inverse match probabilities climb with each additional locus tested, the lines for each breed from the two studies show similar slopes. Reaching a given benchmark for match probability, such as 1 in 10^6, requires 7 loci in Chihuahuas and 13 loci in Whippets.

To evaluate the statistical significance of the differences in the two studies, a bootstrap analysis was performed using the mean expected heterozygosity (averaged over loci) as the test statistic. Table 4 presents the mean expected heterozygosities for each breed in the AKC and Zoogen databases with 95% confidence intervals for the Zoogen means. In ten out of the 15 breeds, the Zoogen confidence intervals capture the AKC means. There are five breeds in which the AKC mean falls outside the 95% confidence intervals, and in all cases the AKC means are lower.

Comparison of Canine Microsatellites to Human CODIS Microsatellites

The individual discrimination of the Stockmarks® canine microsatellites are compared to CODIS loci in Table 5. Population data from the FBI (n = 622), from the Pennsylvania State Police Bureau of Forensics Sciences and Criminal Identification (n = 824), and from the Zoogen database (n = 558) were evaluated by comparing every possible pair of individuals at each locus, counting the number of pairs with genotypes in common, and dividing by the total number of comparisons to determine a frequency. Table 5 shows the counts of individuals and frequencies in each of the three databases sharing matching genotypes at 0–9 loci. In both human databases, the proportion of individuals sharing genotypes at four loci or less was 0.998. In the Zoogen database, the proportion of individuals sharing genotypes at four loci or less was 0.996.

Population Substructure in the Zoogen Study

Table 6 summarizes the measures of Hardy-Weinberg (HW) and linkage (LD) disequilibria in the Zoogen study, showing the numbers of tests significantly deviating from HW expectations at the 5% probability level. To test for HW equilibrium for 17 loci in 18 breed panels, 306 tests were performed. At 5% significance level, it is expected that approximately 5% of these 306 tests (15 tests) would show departures from equilibrium due to chance alone. Instead, there are 57 tests deviating significantly from HW expectations, indicating that the study population is not in HW equilibrium. Testing genotypes across loci takes HW disequilibria into account and provides an estimate of linkage disequilibria (LD). 122 tests out of 2448 total are expected to be significant at the 5% significance level due to chance alone. The number of tests showing linkage disequilibrium at the 5% probability level is 103, which is close to the 122 tests expected by chance, indicating that the study population is in linkage equilibrium.

Fisher (37) addressed the problem of multiple testing. His method is applied to HW and LD tests (Tables 7*a* and 7*b*). Combined p-values obtained from Fisher exact tests over loci within breeds (Table 7*a*) indicate HW disequilibrium within loci in all breeds except Greyhound and Siberian Husky. Tests for linkage using genotypes rather than alleles failed to indicate LD in any breed except for Yorkshire Terriers. Combining p-values over breeds (each breed can be considered an independent test of Hardy-Weinberg equilibrium) for each locus is again significant (Table 7*b*), leading to rejection of random association of alleles within loci.

Inbreeding Coefficients

Estimates of the inbreeding (correlation between alleles) within individuals in a population (f), within individuals over all

TABLE 7*a*—*Fisher's combined P-values for Hardy-Weinberg and linkage disequilibria in dog breeds.*

Breed	HW Statistic	d.f.*	Comb. P	LD Statistic	d.f.	Comb. P
Akitas	68.473	34	0.000	251.784	272	0.805
All Breed Panel	185.182	28	0.000	296.281	304	0.614
Australian Shepherds	54.242	34	0.015	113.618	272	1.000
Chihuahuas	63.408	34	0.002	153.820	272	1.000
Dachshunds	51.510	34	0.028	187.849	272	1.000
Dalmatians	45.966	34	0.083	175.228	272	1.000
English Springer Spaniels	82.700	34	0.000	237.358	272	0.936
Golden Retrievers	58.710	34	0.005	224.759	272	0.983
Greyhounds	38.913	34	0.258	196.441	272	1.000
Siberian Huskies	42.286	34	0.156	203.782	272	0.999
Lhasa Apsos	59.823	34	0.004	243.869	272	0.889
Mixed Breed Dogs	61.272	34	0.003	269.437	272	0.533
Pit Bull Terriers	57.452	34	0.007	191.323	272	1.000
Pomeranians	59.428	34	0.004	149.419	272	1.000
Poodles	82.625	34	0.000	162.707	272	1.000
Shar Peis	48.953	34	0.047	169.666	272	1.000
Shih Tzus	58.629	34	0.005	212.779	272	0.997
Whippets	60.824	34	0.003	242.802	272	0.898
Yorkshire Terriers	68.522	34	0.000	317.521	272	0.030
Total	1248.917	640	0.000	4000.446	5200	1.000

* Degrees of freedom.

TABLE 7*b*—*Fisher's combined P-values for each locus.*

Locus	Statistic	d.f.*	Comb. P Value
PEZ1	87.551	38	0.000
PEZ3	69.258	38	0.001
PEZ5	83.266	38	0.000
PEZ6	81.816	38	0.000
PEZ8	83.193	38	0.000
PEZ10	62.726	38	0.007
PEZ11	64.796	38	0.004
PEZ12	81.398	38	0.000
PEZ13	66.606	38	0.003
PEZ15	69.888	38	0.001
PEZ16	69.529	36	0.001
PEZ17	60.684	38	0.011
PEZ18	73.771	32	0.000
PEZ20	28.26	6	0.000
PEZ21	88.954	36	0.000
FHC2010	56.179	38	0.029
FHC2054	68.747	38	0.002
FHC2079	52.294	36	0.039
Total	1248.917	640	0.000

* Degrees of freedom.

TABLE 8—*Estimates of population substructure parameters f, F, and q for each locus averaged over breeds.*

Locus	f	F	θ
PEZ1	0.139	0.236	0.112
PEZ3	0.079	0.161	0.088
PEZ5	0.152	0.274	0.143
PEZ6	0.157	0.222	0.077
PEZ8	0.128	0.208	0.092
PEZ10	0.086	0.358	0.298
PEZ11	0.117	0.176	0.067
PEZ12	0.090	0.168	0.085
PEZ13	0.159	0.223	0.076
PEZ15	0.081	0.139	0.064
PEZ16	0.134	0.202	0.078
PEZ17	0.112	0.151	0.044
PEZ18	0.136	0.268	0.153
PEZ21	0.257	0.357	0.135
FHC2010	0.103	0.186	0.092
FHC2054	0.061	0.115	0.057
FHC2079	0.128	0.247	0.136
Overall	0.123	0.216	0.106
99% Upper	0.153	0.262	0.151
99% Lower	0.099	0.178	0.078

populations (F), and among individuals of a population (θ) were computed following standard methods (30) and are summarized per each locus in Table 8. Also reported are upper and lower bootstrap confidence limits on the overall estimates. With 99% bootstrap probability, the θ value across all loci lies between 0.078 and 0.151, suggesting a considerable level of inbreeding.

Discussion

Microsatellite markers for forensic DNA identification in any species must have the following attributes. They must be unlinked, highly informative, robust in multiplexes, and easy to score. Microsatellites for parentage and identification in domestic animals should be applicable to the various breeds so that economies of scale can be achieved. In addition, marker sets should be available to the forensic community in a quality-controlled, convenient kit format. Lastly, population data should be available to evaluate locus informativeness and to quantify population substructure.

Ease of Use

The Stockmarks® for Dogs Canine 1 microsatellites are currently available in a PCR multiplex of 1 trinucleotide and 9 tetranucleotide repeats. The primer sequences, size ranges, and chromosomal locations are listed in Table 2. The Canine II kit, consisting of 7 additional tetranucleotide repeat loci, is no longer manufactured by ABI. The vast majority of parentage testing in dogs is successfully resolved with the Canine I loci so the demand for the

Canine II kit was too low to justify its manufacture. The Canine II kit was originally formulated by the first author, and the multiplex is now used for forensic casework. The PCR multiplexes are robust for DNA samples such as buccal swabs (the routine sample type for parentage testing in dogs), blood samples, semen, and hair root bulbs. They have also proven reliable for analysis of samples with a low amount of DNA template. Although a few loci have intermediate alleles due to imperfect repeat structures, the alleles are easily scored with automated software such as Genotyper and GeneMapper. The use of a known canine DNA profile to offset allele bins was validated during development of the canine parentage program for the American Kennel Club. This method has provided sufficient control of electrophoresis variables to allow parentage comparisons of samples tested years apart.

Breed Comparisons

Canine breeds have had varied histories and geographic origins. Some trace their roots to ancient China (e.g., Chinese Shar Peis, Shih Tzus), Tibet (e.g., Lhasa Apsos), Japan (e.g., Akitas), and Egypt (e.g., Greyhounds). Others were developed only in the last few centuries. Whippets, for example, were reportedly bred in the late 1700s by working class people in Great Britain for hunting rabbits and recreational racing. Some breeds have encountered dramatic genetic bottlenecks; others have been out-crossed to gain a particular attribute and then assiduously inbred to fix it. Dog breeds were compared using mean allele number and summary statistics derived from allele frequencies. With the exception of allele numbers in Australian Shepherds, the measures show a similar gradation of breed diversity. In the Zoogen study, the breeds ranked from more diverse to less as follows: Australian Shepherds, Chihuahuas, Poodles, English Springer Spaniels, Lhasa Apsos, Pomeranians, Yorkshire Terriers, Pit Bull Terriers, Dachshunds, Shar Peis, Shih Tzus, Siberian Huskies, Golden Retrievers, Dalmatians, Greyhounds, Akitas, and Whippets.

Comparison of the Zoogen Study and the AKC Parent Club Study

Population data on the Stockmarks® canine loci were acquired in two studies with very different sampling designs. The Zoogen study consisted of randomly selected dogs from a veterinary diagnostic lab and from a sample collection in which dogs were pre-screened to be unrelated for at least 3 generations. The AKC Parent Club Breed Study samples were collected without regard to relatedness between individuals at Parent Club dog shows (special shows held nationally once or twice annually). Parent Club dog shows attract the champion dogs of their breed; some genetic lines may be overrepresented in this data set. If the two studies are substantially similar in their measures of population variation, it would indicate that both data sets provide adequate estimates of the actual allele frequencies of the canine population.

In order to compare the two studies we compared measures of locus informativeness in breeds common to both studies. The mean number of alleles, the mean Expected Heterozygosity (EH), and mean Power of Exclusion (PE) (averaged across breeds) follow similar trends. The mean number of alleles (averaged across breeds) is slightly higher in the AKC study, probably due to the far greater sample number. The mean EH and PE, however, are slightly greater in the Zoogen study, although the difference is only significant in some breeds. In five out of 15 breeds, the mean EHs from the AKC study are significantly less than are seen in the Zoogen study. Several explanations could account for the differences in these breeds. In the case of Greyhounds, the AKC study had a small sample number ($n = 13$) compared to other breeds in the study (average $n = 88$). The sampling protocol for the Zoogen study allowed the inclusion of "pet quality" purebred dogs and, possibly, dogs with the breed misidentified. In contrast, the AKC study did not exclude close relatives and was comprised entirely of dogs participating in dog shows; some genetic lines may have been overrepresented.

Zoogen Database Compared to Human Databases

In order to compare the discrimination potential of the Stockmarks® canine microsatellites with human forensic microsatellites, two human CODIS databases (from the FBI and from the Pennsylvania State Police Bureau of Forensic Sciences and Criminal Identification) were compared with the Zoogen database. The human CODIS loci, the "gold standard" for forensic DNA identification, have been validated extensively by the human forensic community and are widely accepted for DNA identification in court. They are composed of tetranucleotide repeat motifs and are available in convenient, quality-controlled kits.

The frequency of genotype sharing between individuals in populations provides an indication of the discrimination power of forensic microsatellites. The Zoogen canine and two human databases were compared by determining the number of locus genotypes shared between all individuals in each database divided by the total number of comparisons to determine a frequency for one locus shared, two loci shared, etc. In both human databases, the frequency of individuals sharing genotypes at four or fewer loci was 0.998. In the Zoogen database, the frequency of individuals sharing genotypes at four or fewer loci was 0.996. Despite the higher level of inbreeding found in canine populations, the frequencies of individuals with shared locus genotypes in the canine and human databases are very similar (Table 5).

In the appellate court ruling for the *State of Washington v. Kenneth Leuluaialii and George Tuilefano*, the authors expressed concern over the fact that 5 out of 18 possible alleles were shared between the dogs in the *State of Washington v Kenneth L. and George Tuilefano* and *Crown v. William Faulconer*. The authors neglected the fact that the likelihood of encountering the same allele in two profiles is inversely proportional to the number of alleles and their frequencies in a population, whereas the likelihood of encountering the same genotype is much lower since there are many more possible genotypes at a locus than alleles. Although the two profiles in these specific cases did share 5 out of 18 alleles, they shared a genotype at only one out of nine loci, placing the two dogs at the lower end of the distribution seen with the Zoogen database and two human forensic databases.

Population Substructure

The Hardy Weinberg proportion describes the relationship between the allele frequencies and genotype frequencies at a single locus. The Zoogen study database involves 17 loci in 17 AKC-recognized breeds and mixed-breed dogs. This entails 306 independent tests of the Hardy-Weinberg equilibrium. For statistical tests at the five-percent significance level, approximately 5% (15 tests) would be expected to show disequilibria just by chance alone. Instead, there were 57 tests showing disequilibria; this is far more than one would expect if the loci were in Hardy-Weinberg equilibria in these populations.

Linkage disequilibrium refers to associations of alleles or genotypes between loci. If there is equilibrium across loci, then the genotypic frequencies can be multiplied to get a profile frequency. Out of 2482 tests of linkage equilibrium performed, 103 tests with

A. Genotypes of evidence samples and standard with allele frequencies from the combined Zoogen Study (n = 558).

Locus	Jacket A	Jacket B	Jeans	Chief	Freq allele1	Freq allele 2
PEZ 1	AC	-	-	AC	0.0469	0.2604
PEZ 3	BG	BG	BG	BG	0.0368	0.0632
PEZ 5	AB	AB	AB	AB	0.4842	0.1842
PEZ 6	DF	DF	DF	DF	0.1771	0.2292
PEZ 8	BC	BC	BC	BC	0.1579	0.1316
PEZ 12	GH	GH	-	GH	0.2211	0.1474
PEZ 20	CD	CD	-	CD	0.4096	0.2872
FHC 2010	BD	BD	BD	BD	0.4096	0.234
FHC 2054	BH	BH	BH	BH	0.1579	0.1
FHC 2079	AF	AF	AF	AF	0.2917	0.0208

B.

Likelihood Ratio at different Theta values			
Item	theta=0.0	theta=0.1	theta=0.15
Jacket A	4.18E+12	4.82E+09	7.72E+08
Jacket B	3.17E+11	5.65E+08	1.05E+08
Jeans	4.86E+09	1.51E+07	3.43E+06

C.

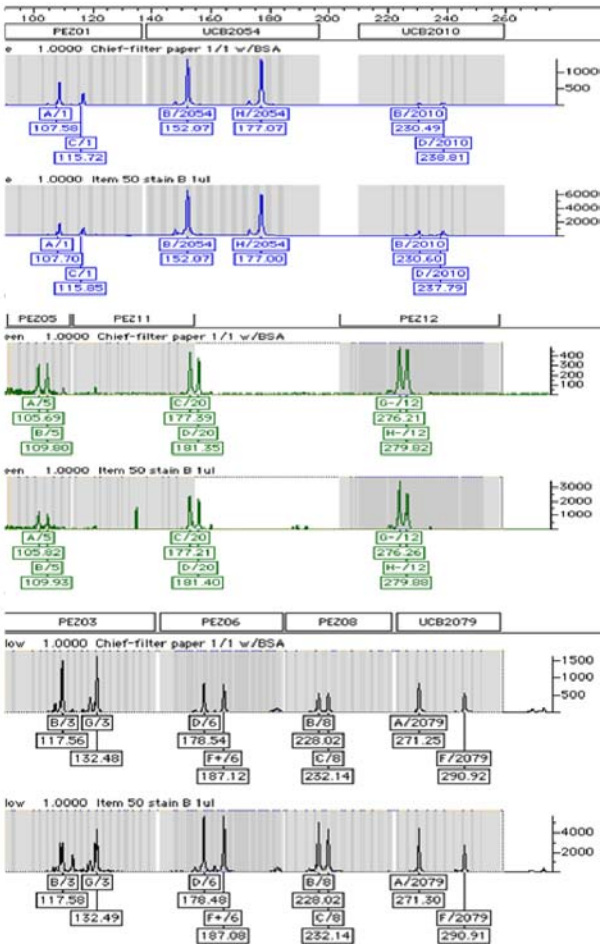

FIG. 4—*The sample genotypes (a), the likelihood ratios at three θ values (b), and an electropheragram (c) comparing evidence Item 50 (Jacket A) with the standard Chief from* State of Washington v. Kenneth Leuluaialii and George Tuilefano. *The samples were amplified with the Stockmarks Canine I loci. The first two rows of the electropheragram show the alleles resulting from amplification with the FAM-labeled primers for PEZ1, FHC2054, and FHC2010. The next two rows show products for the JOE-labeled primers for PEZ5, PEZ20, and PEZ12. The last two rows show products for the NED-labeled primers for PEZ3, PEZ6, PEZ8, and FHC 2079.*

an exact test at a p-value less than 0.05 were observed. This is very close to the expectation of 122 significant tests due to chance alone and argues strongly against associations between loci. It is appropriate, therefore, to multiply the genotype frequencies to compute the overall probability of a DNA test, provided that corrections are made for the allelic disequilibria.

Overall the statistical results are consistent with populations in equilibrium between loci but in disequilibrium within loci. This is consistent with the high levels of inbreeding in the dog populations. From the Zoogen study, the best estimate of θ in dogs is 0.106, which is about ten times the conservative estimate from the human population. Our common approach in casework today is to report likelihood computed with θ values of 0.1 and 0.15.

Case Study: State of Washington v. Kenneth Leuluaialii and George Tuilefano

In the *State of Washington v. Kenneth Leuluaialii and George Tuilefano* case, the DNA extracted from the bloodstains on the suspect's clothing amplified at all ten loci and clearly matched the victims' dog, Chief (Fig. 4). Using the θ value of 0.1, the likelihood for the prosecution's explanation of the match (that the dog Chief was the source of the stains on the clothing of George Tuilefano and Kenneth Leuluaialii) ranged from 1.5×10^7 for the jeans to 4.82×10^9 for Jacket A.

Case Study: Crown v. Daniel McGowan

In the *Crown v. Daniel McGowan* case, DNA profiles from dog hairs found on the victim's clothing matched the dog of Daniel McGowan, the owner of a van allegedly used during the abduction (Fig. 5). The profiled hairs exhibited allelic dropout as shown by the circled loci in Fig. 5*a* and the arrow in Fig. 5*c*. Using a θ value of 0.11 and the equation for probability of a heterozygote with allelic dropout for locus FHC 2079, the likelihood that McGowan's dog was the source of hairs 1–14 found on the victim was 5.27×10^{14} more likely than another unrelated dog.

Conclusion

Microsatellite analysis is a powerful tool for the identification of dogs. In the cases described, the identification of individual dogs, using bloodstains and hairs, has established important links between victims and suspects. The Stockmarks® for Dogs Canine 1 kit is a commercially available, quality-controlled PCR multiplex that includes one trinucleotide locus and nine tetranucleotide loci. Another PCR multiplex, consisting of seven

A. Genotypes of evidence samples and standard with allele frequencies from the combined Zoogen Study (n = 558).

Locus	Hair 1-14	Hair 2-1	Hair 2-3	Duchess	Freq allele1	Freq allele 2
PEZ 1	AA	(A)	AA	AA	0.0469	0.0469
PEZ 3	NN	NN	NN	NN	0.0105	0.0105
PEZ 5	(AC)	-	-	AC	0.4842	0.2368
PEZ 6	-	-	-	EH	0.2344	0.0833
PEZ 8	CC	-	CC	CC	0.1316	0.1316
PEZ 12	pp	-	pp	pp	0.0158	0.0158
PEZ 20	EE	-	EE	EE	0.117	0.117
FHC 2010	BB	BB	BB	BB	0.4096	0.4096
FHC 2054	DG	(D)	(DG)	DG	0.1158	0.1
FHC 2079	(C)	BC	BC	BC	0.4167	0.2031
PEZ 10	HK	(H)	HK	HK	0.1292	0.1685
PEZ 11	DG	-	DG	DG	0.2135	0.0573
PEZ 13	AM	-	AM	AM	0.2278	0.2333
PEZ 15	FP	(P)	FP	FP	0.2135	0.0056
PEZ 16	EE	EE	EE	EE	0.163	0.163
PEZ 17	EE	(E)	(E)	(E)	0.2111	0.2111
PEZ 21	BB	-	BB	BB	0.0976	0.0976

B.

Likelihood Ratio with different Theta values			
Item	theta=0.0	theta=0.11	theta=0.15
Hair 1-14	1.43E+29	5.27E+14	7.73E+12
Hair 2-1	2.04E+10	1.05E+05	2.45E+04
Hair 2-3	1.46E+27	3.28E+13	7.02E+11

C.

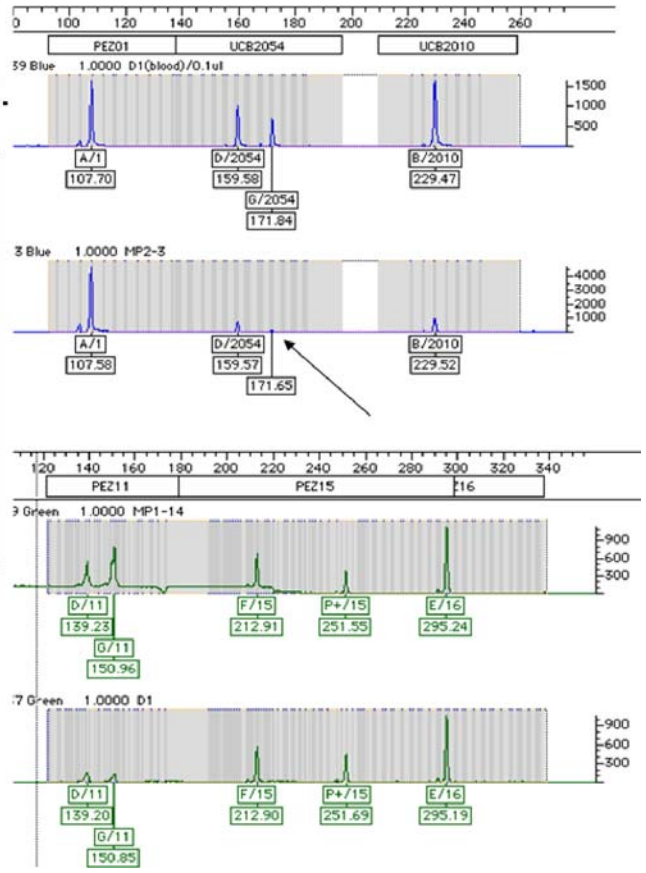

FIG. 5—*The sample genotypes (a), the likelihood ratios at three θ values (b), and an electropheragram (c) comparing evidence hair MP1-14 with the standard Duchess from* Crown v. Daniel McGowan. *The samples were amplified with the Stockmarks Canine I and Canine II loci. The first two rows of the electropheragrams show the alleles resulting from amplification with the Canine I FAM-labeled primers for PEZ1, FHC2054, and FHC2010. The last two rows show products for the Canine II VIC-labeled primers for PEZ11, PEZ15, and PEZ16. Circled loci indicate allele dropout in evidence hair profiles.*

additional tetranucleotide loci, provides additional power for identification in forensic cases. The loci described amplify robustly and are scored easily. Although designed for ideal DNA samples, such as blood samples and buccal swabs, the multiplexes have often provided results with samples containing low amounts of DNA template, such as telogen hairs.

For the Zoogen database, allele frequencies and distributions at 17 loci from 558 dogs of pure and mixed breeds were used to determine measures of locus informativeness (number of alleles, expected heterozygosity, observed heterozygosity, and power of exclusion) and population substructure. Measures of locus informativeness in 15 breeds were also compared to the same breeds in a large study conducted by the American Kennel Club. The mean expected heterozygosities were significantly different in 5 out of 15 breeds (likely due to sampling strategy), but the trends among loci were very similar. The Zoogen database also was compared to two CODIS databases by examining the frequency of individuals sharing matching locus genotypes. Despite the inbreeding evident in canine populations, the Stockmarks® canine microsatellites have a similar discriminatory power to identify individual dogs as human forensic loci to identify people.

State of Washington v. Kenneth Leuluaialii and George Tuilefano was the first case in the United States to use a canine DNA match as evidence. Since that time, similar evidence has been admitted (several with Frye hearings) in nine assault and homicide trials nationwide and in Great Britain. A recent appellate court ruling on *State of Washington v. Kenneth Leuluaialii and George Tuilefano* cautioned that the use of a canine DNA match as evidence was inappropriate without a Frye hearing. The court questioned whether the canine microsatellites used had been characterized and validated sufficiently for forensic investigation. The published AKC study demonstrated the quality of the Stockmarks® canine microsatellites for DNA identification and parentage. This paper provides further validation of these microsatellites and proposes that the Zoogen database, based on randomly selected samples, can be used confidently to determine the statistical significance of forensic canine STR matches.

Acknowledgments

The assistance of Christine S. Tomsey, DNA Laboratory Manager, Pennsylvania State Police Bureau of Forensic Science and Criminal Identification, Forensics Division in providing data from a CODIS database is gratefully acknowledged. The DNA samples provided by Candy Gaiser, Ph.D. were a valuable contribution to the work described. The authors would also like to thank Jan Dvorak, Ph.D. and Candy Gaiser, Ph. D. for their invaluable guidance and editing skills.

References

1. Weber JL, May PE. Abundant class of human DNA polymorphisms which can be typed by the polymerase chain reaction. Am J Hum Genet 1989;44:388–96. [PubMed]
2. Ostrander EA, Sprague, Jr. GF, Rine J. Identification and characterization of dinucleotide repeat (CA)n markers for genetic mapping in dog. Genomics 1993;16:207–13. [PubMed]

3. Francisco LV, Langston AA, Mellersh CS, Neal CL, Ostrander EA. A class of highly polymorphic tetranucleotide repeats for canine genetic mapping. Mam Genome 1996;7(5):359–62.
4. Zajc I, Mellersh CS, Sampson J. Variability of canine microsatellites within and between different dog breeds. Mam Genome 1997;8(3): 182–5.
5. Morera L, Barba CJ, Garrido JJ, Barbancho M, de Andres DF. Genetic variation detected by microsatellites in five Spanish dog breeds. J Here d 1999;90(6):654–6.
6. Zajc I, Sampson J. Utility of canine microsatellites in revealing the rela- [PubMed] tionships of pure bred dogs. J Hered 1999;90(1):104–7.
7. Koskinen MT, Bredbacka P. Assessment of the population structure of five Finnish dog breeds with microsatellites. An Genet 2000;31(5): 310–7.
8. Irion DN, Schaffer AL, Famula TR, Eggleston ML, Hughes SS, Pedersen NC. Analysis of genetic variation in 28 dog breed populations with 100 [PubMed] microsatellite markers. J Hered 2003;94(1):81–7.
9. Koskinen MT. Individual assignment using microsatellite DNA reveals unambiguous breed identification in the domestic dog. An Genet 2003;34:297–301.
10. Breen M, Jouquand S, Renier, Mellersh CS, Hitte C, Holmes N, et al. Chromosome-specific single-locus FISH probes allow anchorage of an 1800-marker integrated radiation-hybrid/linkage map of the domestic [PubMed] dog genome to all chromosomes. Genome Res 2001;11(10):1784–95.
11. Mellersh CS, Hitte C, Richman M, Vignaux F, Priat C, et al. An integrated linkage-radiation hybrid map of the canine genome. Mam Genome 2000;11:120–30.
12. Werner P, Mellersh, CS, Raducha MG, DeRose S, Acland GM, et al. Anchoring of canine linkage groups with chromosome-specific markers. Mam Genome 1999;10:814–32.
13. Neff NW, Broman KW, Mellersh CS, Ray K, Acland GM, Aguirre, et al. A second-generation genetic linkage map of the domestic dog *Canis familiaris*. Genet 1998;151:803–20.
14. Todhunter RJ, Bliss SP, Casella G, Wu R, Lust G, Burton-Wurster NI, et al. Genetic structure of susceptibility traits for hip dysplasia and microsatellite informativeness of an outcrossed canine pedigree. J Hered [PubMed] 2003;94(1):39–48.
15. The American Kennel Club. Pure-Bred Dogs Sect. AKC News. 1996 Sep;9.
16. Halverson J, Dvorak J, Stevenson T, inventors. Metamorphix, Inc., assignee. Microsatellite Sequences for Canine Genotyping. US Patent #5,874,217. February 1999.
17. Primmer CR, Matthews ME. Canine tetranucleotide repeat polymorphism at the VIAS-D10 locus. An Genet 1993;24(4):332.
18. Vila C, Walker C, Sundqvist AK, Flagstad O, Andersone Z, Casullit A, et al. Combined use of maternal, paternal, and bi-parental genetic markers of the identification of wolf-dog hybrids. Hered 2002;90(1):17–24.
19. Dolf G, Schlapfer J, Gaillard C, Randi E, Lucchini V, et al. Differentiation of the Italian wolf and the domestic dog based on microsatellite analysis. Genet Sel Evol (Paris) 2000;32(5):533–41.
20. Shutler GG, Gagnon P, Verret G, Kalyn H, Korkosh S, Johnston E, et al. Removal of a PCR inhibitor and resolution of DNA STR types in mixed human-canine stains from a five year old case. J Forensic Sci [PubMed] 1999;44(3):623–6.
21. Brauner P, Reshef A, Gorski A. DNA profiling of trace evidence: Mitigating evidence in a dog biting case. J Forensic Sci 2001;46(5): 1232–4. [PubMed]
22. Padar Z, Egyed B, Kontadakis K, Furedi S, Woller J, Zoldag L, et al. Canine STR analyses in forensic practice. Observation of a possible mutation in a dog hair. Int J Legal Med 2002;116(5):286–8. [PubMed]
23. Padar Z, Angyal M, Egyed B, Furedi S, Woller J, Zoldag L, et al. Canine microsatellite polymorphisms as the resolution of an illegal animal death in a Hungarian zoological garden. Int J Legal Med 2001;115:79–81. [PubMed]
24. Mueller S, Flekna G, Mueller M, Brem G. Use of canine microsatellite polymorphisms in forensic examinations. J Heredity 1999;90(1): 55–6.
25. Hellmann A, Schleenbecker U, Rohleder U, Demmelmeyer H. DNA Typing of canine samples—allele frequencies and standardization of the STR-marker. Forensic Sci Int 2003;136(suppl.1):66.
26. Lederer T, Betz P, Seidl S. DNA analysis of fingernail debris using different multiplex systems: a case report. Int J Legal Med 2001;114(4–5):263–6. [PubMed]
27. Lu DJ, Sun HY, Chen LX. Genotype reliability of short tandem repeat typing from minute DNA. Fa Yi Za Zhi 2003;19(3):151–3.
28. National Research Council. The evaluation of forensic DNA evidence. Washington, DC: National Academy Press, 1996.
29. Yeh, FC, Yang R-C, Boyle TJB, Ye Z-H, Mao JX. POPGENE, the user-friendly shareware for population genetic analysis. Canada: Molecular Biology and Biotechnology Centre, University of Alberta, 1997.
30. Weir BS. Genetic data analysis II. Sunderland, MA: Sinauer Associates, 1996.
31. Botstein D, White RL, Skolnick MH, Davies RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 1980;32:314–31. [PubMed]
32. Jamieson A. The genetics of transferrins in cattle. Hered 1965;20:419–41.
33. Lewis PO, Zaykin D. Genetic data analysis: Portable version 1.1 for Linux. (Software available at http://lewis.eeb.uconn.edu/lewishome/software.html) 1996.
34. Balding DJ, Nichols RA. DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands. Forensic Sci Int 1994;64:125–40. [PubMed]
35. DeNise S, Johnston E, Halverson J, Marshall K, Rosenfeld D, McKenna S, et al. Power of exclusion for parentage verification and probability of match of identity in American Kennel Club breeds using 17 canine microsatellite markers. An Genet 2004;35:14–7.
36. Budowle B, Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM. Population data on the thirteen CODIS core short tandem repeat loci in African Americans, U.S. Caucasians, Hispanics, Bahamians, Jamaicans, and Trinidadians. J Forensic Sci 1999;44:1277–86. [PubMed]
37. Fisher RA. Statistical methods for research workers, London: Oliver and Boyd, 1932.

Additional information and reprint requests:
Joy Halverson, D.V.M.
QuestGen Forensics
29280 Mace Bld.
Davis, CA 95616
E-mail: questgen@zoogen.biz