Evaluation of a rapid single multiplex microsatellite-based assay for use in forensic genetic investigations in dogs

Leigh Anne Clark, BS; Thomas R. Famula, PhD; Keith E. Murphy, PhD

Objective—To develop a set of microsatellite markers, composed of a minimal number of these markers, suitable for use in forensic genetic investigations in dogs.

Sample Population—Blood, tissue, or buccal epithelial cells from 364 dogs of 85 breeds and mixed breeds and 19 animals from related species in the family *Canidae*.

Procedure—61 tetranucleotide microsatellite markers were characterized on the basis of number and size of alleles, ease of genotyping, chromosomal location, and ability to be coamplified. The range in allele size, number of alleles, total heterozygosity, and fixation index for each marker were determined by use of genotype data from 383 dogs and related species. Polymorphism information content was calculated for several breeds of dogs.

Results—7 microsatellite markers could be coamplified. These markers were labeled with fluorescent dyes, multiplexed into a single reaction, and optimized for resolution in a commercial genetic analyzer. The multiplex set was used to identify sires for 2 mixed litters. The test was not species specific; genotype information collected for wolves, coyotes, jackals, New Guinea singing dogs, and an African wild dog could not distinguish between these species.

Conclusions and Clinical Relevance—This set of 7 microsatellite markers is useful in forensic applications (ie, identification of dogs and determination of parentage) in closely related animals and is applicable to a wide range of species belonging to the family *Canidae*. (*Am J Vet Res* 2004;65:1446–1450)

Microsatellite markers are tandem repeats of 1 to 6 bp that are abundant and evenly distributed across vertebrate genomes. Errors during DNA replication (slippage) occur when polymerase loses its place and causes loss or gain of tandem repeats, resulting in microsatellites that are highly polymorphic. This polymorphic nature, in conjunction with their strict Mendelian inheritance and the ease with which genotypes can be collected, has made microsatellites the markers of choice for use in forensic genetic investigations (ie, determination of parentage and identification of individuals).[1]

Received August 6, 2003.
Accepted September 23, 2003.
From the Department of Pathobiology, College of Veterinary Medicine, Texas A&M University, College Station, TX 77843-4467 (Clark, Murphy) and the Department of Animal Science, College of Agricultural and Enviromental Sciences, University of California, Davis, CA 95616 (Famula).
Supported in part by Genetic Savings and Clone.
The authors thank Sarah Zienko, Jessica Henske, and Edward Cargill for technical assistance.
Address correspondence to Dr. Murphy.

Studies[2-7] of microsatellites in dogs reveal their variability within and among breeds as well as their utility in identifying individuals and determining parentage. High sequence conservation within the family *Canidae* also allows comparative studies of canine microsatellites across species.[8]

It is desirable to minimize the number of **polymerase chain reaction (PCR)** assays required to collect genotype information because the amount of available DNA is often limited; materials, reagents, and resolution of PCR products by use of a genetic analyzer are costly; and data can be generated more efficiently. Until recently, it has been difficult to assemble a panel of microsatellite markers suitable for use in canine forensic genetic investigations because purebred dogs have been highly inbred and line bred, resulting in decreased allelic diversity.[9]

In 2001, Richman et al[10] characterized a set of 172 microsatellite markers (**minimal screening set-1 [MSS-1]**) that was suitable for use in canine whole genome screens. All markers of the MSS-1 were genotyped on panels of 17 reference families or radiation hybrid cell lines and were selected on the basis of high informativeness and inclusion in linkage groups.[10-13] To enhance the utility of the MSS-1, our group multiplexed 155 markers into 48 multiplex sets.[14] Multiplexing is the simultaneous amplification and resolution of markers.

Much of the debate regarding the use of DNA in forensic investigations has focused on issues of statistics and population genetics. Part of the concern is about the accuracy and validity of the data collected and issues of population sampling. The focus of forensic genetics is the computation of a match probability, an expression that requires the knowledge of allelic frequencies and **fixation index (F_{ST})**, a measure of population diversity.[15,16] Implicit in the calculation of a match probability is the effect of population subdivision on the independence of alleles in the population as a whole.

Specifically, F_{ST} measures the amount of genetic variation in the entire population of dogs that can be attributed to differentiation among subpopulations (ie, breeds), such that when $F_{ST} = 0$, there are no genetic differences among subpopulations. Algebraically, $F_{ST} = (H_T - H_S)/H_T$, where H_T is the measure of the total heterozygosity for a locus (ie, the probability that 2 gametes chosen at random from the total population will carry different alleles) and H_S is the subpopulation heterozygosity (ie, the mean heterozygosity among subpopulations).

Although **polymorphism information content (PIC)** values (measurements of variability at a locus) are

not involved in the calculation of match probabilities, many forensic geneticists use PIC values as a means of characterizing the diversity of allelic and genotypic frequencies for a given locus. This statistic is bound by 0.0 and 1.0, such that the closer the value is to 1.0, the greater the amount of polymorphism, and therefore linkage information, captured by this locus.[17,18]

The purpose of the study reported here was to develop a set of microsatellite markers, composed of a minimal number of these markers, suitable for use in forensic genetic investigations in dogs, and with advantages over multiplexed parentage tests already available.[19,20]

Materials and Methods

Selection of microsatellite markers—Sixty-one tetranucleotide microsatellite markers with high (> 0.5) PIC values were selected from the MSS-1 for analysis. To characterize these microsatellite markers, preliminary data including the number of alleles and peak morphology were collected from multigenerational pedigrees of German Shepherd Dogs, a family of mixed-breed dogs, and several groups of purebred dogs. A subset of 11 microsatellite markers was selected on the basis of the number and range of allelic sizes, ease of genotyping, and chromosomal location. Marker compatibility was determined via PCR amplification of various combinations of the 11 markers until coamplification of a maximum number of markers was obtained.

Samples for genotyping—Blood, tissue, or buccal epithelial cell samples collected from dogs and related species were donated by owners, collaborating laboratories, and breeders and exhibitors at dog shows. Related species included the gray wolf (*Canis lupus*), Mexican wolf (*Canis lupus baileyi*), red wolf (*Canis rufus*), maned wolf (*Chrysocyon brachyurus*), coyote (*Canis latrans*), jackal (*Canis adustus*, *Canis mesomelas*, and *Canis aureus*), African wild dog (*Lycaon pictus*), and New Guinea singing dog (*Canis hallstromi*). Genotyping was also performed to determine the paternity of 2 potentially mixed litters of puppies. Genomic DNA was isolated from blood and buccal epithelial cells by use of a commercial DNA isolation kit[a] and from tissue by use of a commercial tissue DNA isolation kit.[b] Working solutions of 50 ng of DNA/µL were prepared.

Table 1—Variables associated with 7 microsatellite markers* coamplified from specimens from 383 dogs and related species.

Marker	Chromosomal location	Fluorescent label	Concentration of primer (µmol)	Range in size of alleles	No. of alleles (bp)	H_T	F_{ST}
FH2309	CFA01	Dye 1†	0.60	342–474	34	0.93	0.20
FH2132	CFA02	FAM	0.74	152–370	52	0.92	0.27
FH2137	CFA03	Dye 3†	0.30	154–316	27	0.90	0.16
FH2263	CFA09	Dye 2†	0.89	175–503	50	0.96	0.16
FH2293	CFA10	Dye 1†	0.74	183–527	49	0.91	0.16
FH2321	CFA17	Dye 2†	0.60	276–396	42	0.93	0.20
FH2001	CFA23	FAM	0.45	115–163	13	0.83	0.21

*Markers selected on the basis of number and size of alleles, ease of genotyping, chromosomal location, and ability to be coamplified. †For dyes 1, 2, and 3 please see footnote d.
H_T = Total heterozygosity. F_{ST} = Fixation index. FAM = 6-carboxy-fluorescein.

Table 2—Variables associated with 7 microsatellite markers coamplified from specimens from unrelated and related populations of 8 groups of purebred dogs and a group of gray wolves.

Unrelated populations	Miniature Schnauzer (n = 14)			Labrador Retriever (13)			Golden Retriever (12)			Boxer (11)			Gray wolf (7)		
Marker	PIC	No. of alleles	H_S	PIC	No. of alleles	H_S	PIC	No. of alleles	H_S	PIC	No. of alleles	H_S	PIC	No. of alleles	H_S
FH2309 (0.82*)	0.78	7	0.84	0.85	10	0.90	0.67	7	0.84	0.38	2	0.52	0.82	8	0.92
FH2132 (0.76)	0.70	12	0.76	0.82	11	0.88	0.58	4	0.67	0.73	6	0.79	0.90	12	0.98
FH2137 (0.60)	0.79	10	0.85	0.82	10	0.87	0.75	6	0.83	0.63	4	0.72	0.87	10	0.95
FH2263 (0.92)	0.83	11	0.88	0.87	12	0.92	0.85	9	0.90	0.45	4	0.54	0.79	8	0.98
FH2293 (0.87)	0.72	10	0.78	0.82	9	0.87	0.81	9	0.87	0.70	6	0.79	0.89	11	0.98
FH2321 (0.84)	0.82	12	0.88	0.89	14	0.94	0.74	8	0.80	0.23	3	0.25	0.85	9	0.93
FH2001 (0.62)	0.49	5	0.56	0.73	7	0.80	0.67	6	0.75	0.52	3	0.64	0.69	5	0.79

Related populations	American Eskimo Dog (45)			Dalmatian (44)			German Shepherd Dog (31)			Miniature Schnauzer (7)		
Marker	PIC	No. of alleles	H_S	PIC	No. of alleles	H_S	PIC	No. of alleles	H_S	PIC	No. of alleles	H_S
FH2309 (0.82)	0.59	6	0.65	0.65	5	0.71	0.36	4	0.55	0.58	3	0.69
FH2132 (0.76)	0.54	4	0.62	0.09	3	0.09	0.44	3	0.54	0.58	3	0.68
FH2137 (0.60)	0.70	6	0.75	0.63	6	0.70	0.53	3	0.64	0.5	3	0.60
FH2263 (0.92)	0.87	14	0.89	0.72	8	0.76	0.58	4	0.66	0.58	3	0.70
FH2293 (0.87)	0.39	3	0.47	0.79	8	0.76	0.47	4	0.60	0.46	3	0.59
FH2321 (0.84)	0.65	7	0.71	0.68	6	0.73	0.29	4	0.37	0.58	3	0.70
FH2001 (0.62)	0.55	4	0.62	0.39	3	0.50	0.62	4	0.44	0.28	2	0.36

*Published PIC value.[11]
PIC = Polymorphism information content. H_S = Subpopulation heterozygosity.
Values for German Shepherd Dogs are means of 3 unrelated families.

DNA amplification—The 5' end of the forward primer[e] for each microsatellite marker selected was labeled with 1 of 4 fluorescent dyes (6-carboxy-fluorescein and 3 other commercially available dyes[d]; **Table 1**). Labels were chosen on the basis of observed allelic sizes to prevent overlap of dye types. Concentrations of reagents for a 13.45-µL PCR volume were 0.37mM total dNTPs, 3.7 ng of genomic DNA/µL, 2.2mM magnesium chloride, 0.74X commercial DNA polymerase buffer,[e] 0.74X PCR enhancer,[f] and 0.06 U of DNA polymerase/µL. Concentrations of primers varied for each microsatellite marker. A commercial thermal cycling instrument[g] was used. Thermal cycling conditions for DNA amplification were 2 minutes at 95°C; 5 cycles of 30 seconds at 95°C, 15 seconds at 58°C, and 10 seconds at 72°C; 30 cycles of 20 seconds at 95°C, 15 seconds at 56°C, and 10 seconds at 72°C; and a final extension of 5 minutes at 72°C.

Products of DNA amplification reactions were separated by use of a commercial genetic analyzer[h] and sized relative to an internal size standard.[i] Genotypes were determined by use of standard software programs.[j,k]

Calculation of H_T, H_S, F_{ST}, and PIC values—Calculations of H_T, H_S, and F_{ST} were performed by use of public domain software.[21] Total heterozygosity and F_{ST} were calculated for each microsatellite marker by use of genotype data for all dogs and related species.

Calculation of PIC values from the estimated allelic frequencies was performed by use of software developed by our group.[l] The PIC values and H_S were calculated for several groups of purebred dogs.

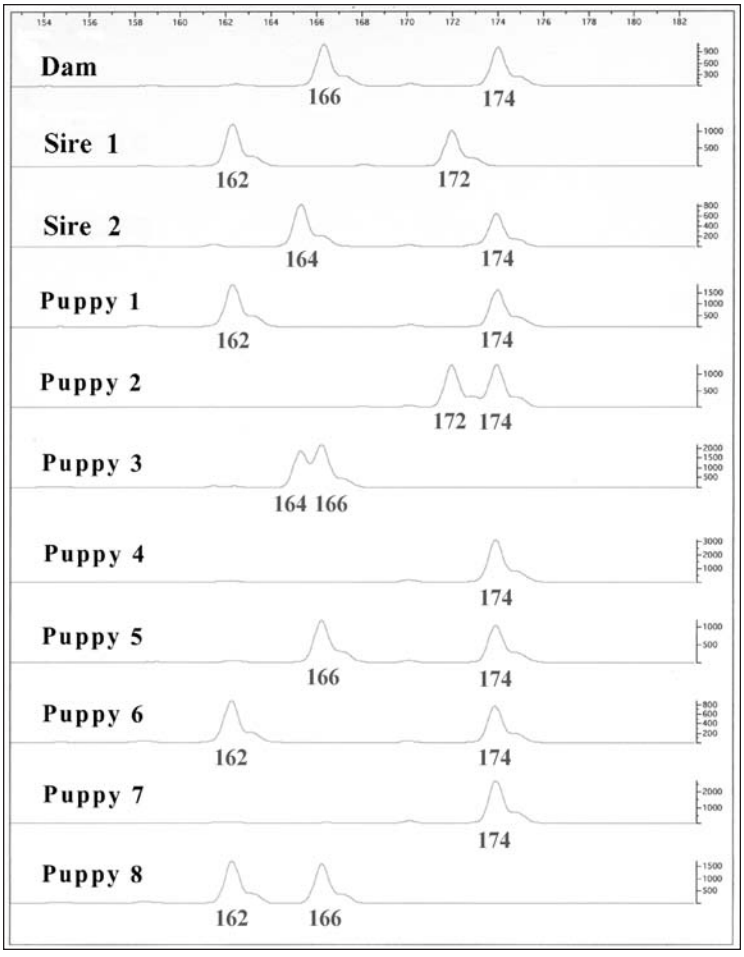

Figure 1—Genotype data for marker FH2137 for 11 Blue Tick Hounds (dam, 2 potential sires, and 8 puppies). The number below each peak is the size of the allele (bp). Neither the dam nor sire 2 had alleles 162 or 172, excluding sire 2 as a potential sire of puppies 1, 2, 6, or 8. Sire 1 did not have alleles 164 or 174, excluding that dog as a possible sire of puppies 3, 4, 5, or 7.

Results

Eleven microsatellite markers were selected on the basis of ease of resolution of products, number and range in size of alleles, and chromosomal location. Seven markers could be coamplified: FH2309, FH2263, FH2293, FH2321,[11] FH2132, FH2137, and FH2001[22] (Table 1). All loci were on different chromosomes.

Genotype data for the panel of markers were generated for 364 dogs that represented 82 breeds recognized by the American Kennel Club, including 16 breeds with the highest registration rates in 2002. Breeds not recognized by the American Kennel Club included the Boykin Spaniel, Blue Tick Hound, and Long Hair Weimeraner. Forty-eight of 364 dogs were mixed-breed dogs. Genotype data were also collected for 7 gray wolves from Alaska, Canada, Sweden, Spain, and Oman; 1 Mexican wolf; 1 red wolf; 3 coyotes representing 3 populations in the United States; 3 species of jackals from Africa; 2 New Guinea singing dogs; 1 maned wolf; and 1 African wild dog.

For all 7 microsatellite markers, no 2 genotypes were identical, even among closely related individuals. Total heterozygosity and F_{ST} were calculated for each marker by use of all 383 genotypes. Total heterozygosity ranged from 0.83 to 0.96 with a mean ± SD of 0.91 ± 0.04 for all 7 markers. Fixation index ranged from 0.16 to 0.27 with a mean ± SD of 0.19 ± 0.04 (Table 1).

Polymorphism information content and H_S values were calculated for 7 breeds of dogs and the gray wolf (**Table 2**). Four purebred groups (Miniature Schnauzer, Boxer, Golden Retriever, and Labrador Retriever) were comprised of unrelated, randomly selected dogs. Polymorphism information content and H_S values for 3 breeds (American Eskimo Dog, Dalmatian, and 3 unrelated families of German Shepherd Dog) for which extended pedigrees were available were calculated in these multigenerational families. The Miniature Schnauzer, Golden Retriever, and Labrador Retriever groups had PIC values similar to published values, which were calculated by use of 17 reference pedigrees of purebred dogs.[11] As expected, PIC values calculated within families were generally lower than published values. Among the unrelated dogs analyzed, Boxers had the lowest PIC values for 5 of 7 markers.

We tested the utility of our multiplexed set of microsatellite markers via 2 applications of the technique. A breeder of Blue Tick Hounds suspected that 2 males had bred a bitch and the litter of 8 puppies was mixed. Genotype data for the panel of markers were collected for the dam, both potential sires, and all puppies. Genotypes for FH2137 (**Figure 1**) and other data revealed that each male sired 4 puppies. Puppies 4 and

7 had only 1 peak at allele 174 and were homozygous for this locus. Because each parent must have contributed 1 of these alleles each, sire 1 can be excluded as the sire of these puppies. Puppies 3, 5, and 8 had allele 166, which could only have been contributed by the dam. Because the remaining alleles for puppies 3 and 5 were 164 and 174, respectively, sire 1 can again be excluded as the sire of these puppies. Puppies 1, 2, 6, and 8 had a copy of either allele 162 or 172, excluding sire 2 as the sire of these puppies.

We also used our set of microsatellite markers to determine the paternity of a litter of 4 Miniature Schnauzers. The intended breeding was a repeat breeding, but a male from the dam's first litter had also bred the dam. Genotype data were collected for the dam, both potential sires, and the 4 puppies. Polymorphism information content and H_S values were calculated within this population (**Table 2**). Two microsatellite markers (FH2293 and FH2321) excluded the dam's previous offspring as the sire of 3 puppies, and 1 marker (FH2263) excluded the intended sire as the sire of the fourth puppy.

Discussion

We developed a single multiplex set of 7 tetranucleotide microsatellite markers suitable for use in forensic genetic investigations in dogs. Only tetranucleotide microsatellite motifs were selected because they are more polymorphic than dinucleotide repeats and typically resolve with fewer stutter bands (extra peaks that result from errors during replication and complicate analysis of products).[22] The polymorphic nature of tetranucleotides is thought to be the result of greater instability associated with longer repeat lengths.[22] Our data support this hypothesis. Of the 7 microsatellite markers, those with the smallest product sizes (FH2001 and FH2137) had the lowest number of alleles (13 and 27, respectively). The remaining markers had larger product sizes and a mean of 45 alleles.. Marker FH2001, which had the fewest alleles overall and was the only marker to have all alleles differ by exactly 4 nucleotides, was the only (GATA) repeat microsatellite. The other 6 markers were (GAAA) repeats known to be markedly polymorphic microsatellites and have variation within the repeat sequence.[22]

In general, all dogs and related species had similar ranges of allelic sizes. The widest range in size of alleles was in FH2293, for which the maned wolf had the smallest allele and the Black-backed jackal had the largest. Marker FH2263 also had a wide range of allelic sizes. Many dogs in the American Eskimo Dog family and 7 unrelated dogs of other breeds had allelic sizes of approximately 500 nucleotides for FH2263; the largest allelic size for this marker in the remaining breeds was 275 nucleotides. Sequence data confirmed that these large alleles were composed of (GAAA) repeats.

Many breeds of dogs were represented by only a small number of dogs. Therefore, PIC values were not calculated for every breed. Total heterozygosity and F_{ST} were calculated for each marker across all dog breeds and species as measures of genetic variation. The mean F_{ST} among the microsatellite markers (0.19) lies between 0.15 and 0.25, indicating marked genetic differentiation among subpopulations.[18] Values for humans range from 0.02 to 0.11.[18]

The PIC values and allelic sizes in 7 unrelated wolves were similar to those in dogs. The similarity of alleles between species indicates that this panel of markers is not appropriate for use in evolutionary studies. However, the high PIC values in the wolf and the high H_S values for the markers indicate that the panel is suitable for forensics applications in this species.

Although the large number of alleles that exist for each marker indicates that the panel is highly polymorphic, it is not an indicator of how polymorphic the markers are within individual breeds. The PIC values calculated for the Miniature Schnauzers, Labrador Retrievers, and Golden Retrievers were similar to published PIC values and indicate that these markers are informative within these breeds.[11] Only the PIC values for Boxers were lower than published values.[11] This was a factor in selecting Boxers to be the first breed of dog to have its genome sequenced.[23]

Empirically, genetic variability in populations of related dogs will be lower than in populations of unrelated dogs.[2] Overall, the PIC values calculated in the American Eskimo Dog, Dalmatian, and German Shepherd Dog families were lower than those of dogs in unrelated populations. Despite a lesser degree of polymorphism, sufficient markers were informative (ie, those with PIC values ≥ 0.5) to identify every dog analyzed.

The 7 selected microsatellite markers were sufficient to determine parentage. In the mixed litter of Blue Tick Hounds, all markers supported the conclusion that each male sired 4 puppies. One marker (FH2137) was sufficiently polymorphic to reveal the paternity of each puppy. The testing involving the Miniature Schnauzers illustrated the ability of the 7 microsatellite markers to identify parentage in closely related individuals. Despite the lower PIC values in this family of Miniature Schnauzers, compared with values for unrelated Miniature Schnauzers, the panel of markers was able to detect the paternity of all 4 puppies.

[a]Puregene DNA isolation kit, Gentra Systems, Minneapolis, Minn.
[b]DNeasy tissue kit, Qiagen, Valencia, Calif.
[c]primers, cPE Biosystems, Foster City, Calif.
[d]6FAM, VIC (dye 1), NED (dye 2), and PET (dye 3), PE Biosystems, Foster City, Calif.
[e]Buffer B, Fisher Scientific, Pittsburgh, Pa.
[f]MasterAmp PCR enhancer, Epicentre Technologies, Madison, Wis.
[g]Eppendorf mastercycler, Eppendorf Scientific Inc, New York, NY.
[h]ABI 3100 capillary-based genetic analyzer, PE Biosystems, Foster City, Calif.
[i]GeneScan LIZ size standard, PE Biosystems, Foster City, Calif.
[j]GeneScan 3.1, PE Biosystems, Foster City, Calif.
[k]Genotyper 2.0, PE Biosystems, Foster City, Calif.
[l]Available from authors upon request.

References

1. Leopoldino AM, Pena SD. The mutational spectrum of

human autosomal tetranucleotide microsatellites. *Hum Mutat* 2003;21:71–79.

2. Sutton MD, Holmes NG, Brennan FB, et al. A comparative genetic analysis of the Irish Greyhound population using multilocus DNA fingerprinting, canine single locus minisatellites and canine microsatellites. *Anim Genet* 1998;29:168–172.

3. Zajc I, Sampson J. Utility of canine microsatellites in revealing the relationships of pure bred dogs. *J Hered* 1999;90:104–107.

4. Ichikawa Y, Takagi K, Tsumagari S, et al. Canine parentage testing based on microsatellite polymorphisms. *J Vet Med Sci* 2001; 63:1209–1213.

5. Binns MM, Holmes NG, Marti E, et al. Dog parentage testing using canine microsatellites. *J Small Anim Pract* 1995;36:493–497.

6. Fredholm M, Winterø AK. Efficient resolution of parentage in dogs by amplification of microsatellites. *Anim Genet* 1996;27:19–23.

7. Muller S, Flekna G, Muller M, et al. Use of canine microsatellite polymorphisms in forensic examinations. *J Hered* 1999; 90:55–56.

8. Fredholm M, Winterø AK. Variation of short tandem repeats within and between species belonging to the *Canidae* family. *Mamm Genome* 1995;6:11–18.

9. Pihkanen S, Vainola R, Varvio S. Characterizing dog breed differentiation with microsatellite markers. *Anim Genet* 1996;27:343–346.

10. Richman M, Mellersh C, Andre C, et al. Characterization of a minimal screening set of 172 microsatellite markers for genome-wide screens of the canine genome. *J Biochem Biophys Methods* 2001; 47:137–149.

11. Mellersh CS, Langston AA, Acland GM, et al. A linkage map of the canine genome. *Genomics* 1997;46:326–336.

12. Neff MW, Broman KW, Mellersh CS, et al. A second-generation genetic linkage map of the domestic dog, *Canis familiaris*. *Genetics* 1999;151:803–820.

13. Vignaux F, Hitte C, Priat C, et al. Construction and optimization of a dog whole-genome radiation hybrid panel. *Mamm Genome* 1999;10:888–894.

14. Cargill EJ, Clark LA, Steiner JM, et al. Multiplexing of canine microsatellite markers for whole-genome screens. *Genomics* 2002;80:250–253.

15. Evett IW, Weir BS. Statistical genetics. In: *Interpreting DNA evidence: statistical genetics for forensic scientists*. Sunderland, Mass: Sinauer Associates Inc, 1998;132–162.

16. Hartl DL, Clark AG. Population substructure. In: *Principles of population genetics*. 3rd ed. Sunderland, Mass: Sinauer Associates Inc, 1997;111–162.

17. Botstein D, White RL, Skolnick M, et al. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 1980;32:314–331.

18. Liu BH. Single-locus models. In: *Statistical genomics: linkage, mapping, and QTL analysis.* Boca Raton, Fla: CRC Press Inc, 1998; 139–162.

19. Koskinen MT, Bredbacka P. A convenient and efficient microsatellite-based assay for resolving parentages in dogs. *Anim Genet* 1999;30:148–149.

20. Altet L, Francino O, Sanchez A. Microsatellite polymorphism in closely related dogs. *J Hered* 2001;92:276–279.

21. Raymond M, Rousset F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J Hered* 1992; 86:248–249.

22. Francisco LV, Langston AA, Mellersh CS, et al. A class of highly polymorphic tetranucleotide repeats for canine genetic mapping. *Mamm Genome* 1996;7:359–362.

23. National Human Genome Research Institute. Newsroom. Available at:www.genome.gov/11007358. Accessed May 20, 2003.