

Power of exclusion for parentage verification and probability of match for identity in American kennel club breeds using 17 canine microsatellite markers

S. DeNise*, E. Johnston*, J. Halverson[†], K. Marshall*, D. Rosenfeld*, S. McKenna[‡], T. Sharp[§] and J. Edwards[¶]

*MMI Genomics, Davis, CA, USA. [†]QuestGen Forensics, Davis, CA, USA. [‡]Cognosci, Inc., NC, USA. [§]American Kennel Club, NC, USA.

[¶]Danbury, NC, USA

Summary

DNA analysis of microsatellite markers has become a common tool for verifying parentage in breed registries and identifying individual animals that are linked to a database or owner. Panels of markers have been developed in canines, but their utility across and within a wide range of breeds has not been reported. The American Kennel Club (AKC) authorized a study to determine the power to exclude non-parents and identify individuals using DNA genotypes of 17 microsatellite markers in two panels. Cheek swab samples were voluntarily collected at Parent Breed Club National Specialty dog shows and 9561 samples representing 108 breeds were collected, averaging 88.5 dogs per breed. The primary panel of 10 markers exceeded 99% power of exclusion for canine parentage verification of 61% of the breeds. In combination with the secondary panel of seven markers, 100% of the tested breeds exceeded 99% power of exclusion. The minimum probability match rate of the first panel was 3.6×10^{-5} averaged across breeds, and with the addition of the second panel, the probability match rate was 3.2×10^{-8} ; thus the probability of another random, unrelated dog with the same genotype is very low. The results of this analysis indicated that, on average, the primary panel meets the AKC's needs for routine parentage testing, but that a combination of 10–15 genetic markers from the two panels could yield a universal canine panel with enhanced processing efficiency, reliability and informativeness.

Keywords canine microsatellite, genetic analysis, markers, parentage analysis.

Introduction

DNA analysis provides a powerful tool for verifying the parentage and identification of individual animals. Human forensics led the way by developing microsatellite marker panels that are highly informative among human sub-populations. Marker panels developed for domesticated and highly selected animals are uniquely challenging for these purposes, particularly in dogs, because of founder effects and breeding bottlenecks that could reduce the utility of markers among varieties and breeds as they diverged through time.

Modern breeds of dogs were domesticated from wolves (Vilà *et al.* 1997) and phylogenetic evidence suggests ori-

gins in East Asia approximately 15 000 years ago (Savolainen *et al.* 2002). Through genetic drift caused by small founder populations and interbreeding with other canids, a wide phenotypic and, thus, genetic, variety of dogs comprise our present day breeds. A number of microsatellite markers have been identified in canines (Ostrander *et al.* 1993, 1995; Halverson *et al.* 1995; Francisco *et al.* 1996) but the power of these markers to discriminate among parents and provide evidence of identity among and within dog breeds has not been reported in the literature.

The purpose of this study was to evaluate 17 microsatellite markers on their information content, ability to discriminate among individuals within a breed and across breeds and to determine the power of the panels to exclude non-parents within a breed. The results of this study can be used to develop powerful and efficient genetic tools for breeders and breed registries to provide parentage verification and to match individual dogs when identity is lost.

Address for correspondence

S. DeNise, MMI Genomics, 1756 Picasso Ave., Davis, CA 95616, USA.
E-mail: sue.denise@mmigenomics.com

Accepted for publication 21 October 2003

Materials and methods

The study was authorized by the American Kennel Club (AKC) Board of Directors to determine the ability of a collection of microsatellite markers to discriminate among parents of pups and to allow identification of individual animals. The study was also designed to test the high throughput genotyping system that would allow the AKC to monitor matings, parentage and identity for breeds in a timely and accurate manner.

Cheek swab samples were collected at national breed club specialty dog shows held in 1998, 1999 and 2000, through the Parent Breed Club Program. The AKC registered dogs, representing 108 breeds, were sampled for a total of 9561 animals. An average of 88.5 dogs were sampled per breed ranging from 11 (Norfolk Terrier) to 218 (German Shepherd). No screening of data occurred; thus, families of dogs could be represented. To determine the level of inbreeding and relatedness among the animals within a breed, a four-generation pedigree was generated for each dog sampled and inbreeding coefficients and coancestry relationships were calculated. These data (see *Supplementary material* Table S1) indicate that the degree of relationship among sampled individuals within any one breed was low. Average inbreeding coefficient and relationship was less than 0.10 through four generations with the exception of the following breeds: American Water Spaniel, Canaan Dog, Dandie Dinmont Terrier, Finnish Spitz, Greyhound, Ibizan Hound, Komondor, Pharaoh Hound and Wirehaired Pointing Griffon. Based on these data, no adjustment to allele frequency was made to account for degree of relatedness within a breed.

Seventeen microsatellite markers were selected for the study based on five criteria: reproducibility, information content, ease of scoring, multiplex assay robustness and low mutation rates. These markers were optimized to allow evaluation in two multiplex reactions: a 10 marker panel and a seven marker panel. The 10 marker multiplex system became the commercial StockMarks[®] for Dogs Parentage Typing Kit from Applied Biosystems (ABI, Foster City, CA, USA). To enhance the power of parentage and identity, a seven marker panel was developed in our laboratories and optimized in a single multiplex reaction. Table 1 lists the markers, panel numbers and map location of the marker, if known. The FHC markers are public domain markers identified by the Fred Hutchinson Cancer Research Center (Francisco *et al.* 1996; GenBank accession numbers: L78580, L78579, L78590, L78589, L78595 and L78596). The PEZ markers are identified in Halverson *et al.* (1995).

Markers were amplified in a 10- μ l reaction as recommended by ABI (K-9 1 Version 3, 4307481C and a customized kit) and fragments were analysed on ABI PRISM[®] 377 DNA Sequencers and ABI PRISM[®] 3100

Table 1 Microsatellite markers, description of the markers, and map location.

Locus ¹	Type of repeat	Range in size (bp)	Panel	Canine map location ²
CATA ₁	tetra	95–136	Canine I	CFA7
PEZ03	tri	95–154	Canine I	CFA13/19
PEZ05	tetra	97–121	Canine I	CFA12
PEZ06	tetra	166–215	Canine I	CFA16
PEZ08	tetra	230–260	Canine I	S16
PEZ12	tetra	250–317	Canine I	CFA3
PEZ20	tetra	152–202	Canine I	Unmapped
FHC2010	tetra	220–248	Canine I	S5
FHC2054	tetra	140–184	Canine I	CFA12
FHC2079	tetra	263–299	Canine I	S5
PEZ10	tetra	230–330	Canine II	S9
PEZ11	tetra	123–175	Canine II	CFA8, CFA9
PEZ13	tetra	171–322	Canine II	S1
PEZ15	tetra	193–284	Canine II	CFA16
PEZ16	tetra	263–334	Canine II	CFA16
PEZ17	tetra	196–245	Canine II	S1
PEZ21	tetra	71–109	Canine II	Unmapped

¹PEZ markers are identified in US patent 05874217; FHC markers were identified by the Fred Hutchinson Center (Francisco *et al.* 1996).

²From Werner *et al.* (1999).

Genetic Analyzers. Each animal in the study was genotyped for the entire group of 17 markers.

Data were summarized by marker and breed for a number of parameters. Allele frequencies and observed heterozygosity were calculated by direct counting of alleles from the sample of individuals for each breed. Expected heterozygosity was estimated from allele frequencies assuming Hardy–Weinberg equilibrium. Polymorphic information content (PIC, Botstein *et al.* 1980) was calculated as

$$1 - \sum_{i=1}^n p_i^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^n 2p_i^2 p_j^2$$

where p_i is the population frequency of the i th allele and n is the number of alleles per marker.

Exclusion power (EPR), the probability of excluding a random individual from the population as a potential parent of an animal based on the genotype of one parent and offspring, was calculated from Jamieson (1965) as

$$\sum_{i=1}^n p_i(1 - p_i)^2 - \sum_{i>j=1}^n p_i^2 p_j^2 (4 - 3p_i - 3p_j)$$

where p is the population frequency of the i th or j th allele and n is the number of alleles per marker. Match probability ratio (MPR) was calculated as the square frequency of the most common allele to provide the most conservative estimate of match rate within a breed. Cumulative MPR was estimated as the sequential product of MPR at each microsatellite locus.

Results and discussion

Summaries of marker characteristics within each breed and averaged across all breeds are presented in the *Supplementary material* (Tables S2 and S3) because of the length of the tables. The total number of alleles identified per marker across all breeds are: *CATA*₁: 10 alleles; *PEZO*₃: 17 alleles; *PEZO*₅: 7 alleles; *PEZO*₆: 22 alleles; *PEZO*₈: 14 alleles; *PEZ*₁₂: 24 alleles; *PEZ*₂₀: 10 alleles; *FHC*₂₀₁₀: 6 alleles; *FHC*₂₀₅₄: 12 alleles; *FHC*₂₀₇₉: 10 alleles; *PEZ*₁₀: 28 alleles; *PEZ*₁₁: 17 alleles; *PEZ*₁₃: 31 alleles; *PEZ*₁₅: 27 alleles; *PEZ*₁₆: 16 alleles; *PEZ*₁₇: 10 alleles; and *PEZ*₂₁: 11 alleles. The mean number of alleles per marker within breed ranged from 2.9 (Skye Terrier) to 8.4 (Australian Shepherd) with an average number of alleles per locus per breed of 5.7. *PEZO*₃ was the most informative marker in the first panel across breeds with an average number of 6.7 alleles per breed, and *PEZ*₁₀ was the most informative in the second panel with 9.3 alleles per breed.

Only four of the 17 markers were monomorphic in specific breeds, and most of those breeds had the fewest numbers of sampled animals. *FHC*₂₀₇₉ was monomorphic in Boxer, Bull Terrier, English Toy Spaniel and Skye Terrier. *FHC*₂₀₁₀ was monomorphic in Norfolk Terrier and Skye Terrier. *PEZ*₂₀ was monomorphic in Scottish Deerhound and Skye Terrier. *PEZO*₅ was monomorphic in the Pharaoh Hound.

The average observed heterozygosity of the first panel was 0.57 and of the second panel was 0.64. The deviations of observed and estimated heterozygosity ranged from 0.01 to 0.03 for all 17 markers averaged across all breeds. For breeds with the highest average deviation for observed and estimated heterozygosity, there were few markers that contributed large deviations. For example, two breeds (Soft Coated Wheaten Terriers and Bedlington Terriers) had average deviations of 0.09 across the 17 markers. Two markers contributed the greatest deviations for Soft Coated Wheaten Terriers, and three markers contributed the greatest deviations for Bedlington Terriers. However, most markers were in Hardy–Weinberg equilibrium in all breeds.

Breeds differed in PIC at individual loci but average PIC values across the markers were high. Average PIC value of the first panel was 0.53 and of the second panel was 0.61. The primary panel of 10 markers was able to reach the 99% power of exclusion for canine parentage verification for 61% of the breeds (*Supplementary material* Table S2). In combination with the secondary panel of seven markers, 100% of the tested breeds reached the 99% power of exclusion. The MPR of the panels was calculated using the most likely homozygote genotype probability for each marker. The first panel MPR average across breeds was 3.6×10^{-5} ; with the addition of the second panel, the MPR was 3.2×10^{-8} ; thus the probability of another random, unrelated dog with the same genotype is very low. Basset Hounds required only the

first six markers to reach the 1×10^{-6} MPR; while all 17 markers were not informative enough to reach that same probability for Skye Terriers.

These seventeen genetic markers have been shown to be reproducible, informative and robust, and have the information content necessary for powerful parentage oversight and identity programmes. The results of this analysis indicated that, on average, the primary panel meets the needs of the AKC breeds for routine parentage testing and the addition of the second panel provides a high probability of resolution in those cases requiring additional markers. Because all of these markers have similar powers of discrimination, we propose that a combination of 10–15 genetic markers from the two panels could yield a universal canine panel with enhanced processing efficiency, reliability and informativeness given enhancements in the technology of fragment analysis systems. Selection of the specific markers that would form the universal panel depends on the accuracy required of the panel balanced against the length of the amplicons and ability to form a successful multiplex reaction.

Although this study focused on genetic identity and parentage, it demonstrates the utility of a DNA testing programme in a large breed registry. As such, the greatest impact may be as an effective precursor or paradigm of DNA testing for canine genetics and health.

Acknowledgements

The authors wish to thank Robert Slay and the Board of Directors of The American Kennel Club for their support of the project. We also wish to thank the participating Parent Breed Clubs for their cooperation and support, and the individual participating dog owners for their willingness to supply samples for the study. We also wish to acknowledge the dedicated participation of Leslie Bickel and the genotyping group at MMI Genomics for analysis of the genotypes used in this study.

Supplementary material

The following material is available from: <http://www.blackwellpublishing.com/products/journals/suppmat/AGE/AGE1074/AGE1074sm.htm>

Table S1 Number of sampled dogs per breed and average inbreeding coefficient and coancestry relationship of sampled dogs based on a four-generation pedigree.

Table S2 Number of animals tested by breed, alleles per marker, expected (EH) and observed (OH) heterozygosity, polymorphic information content (PIC), exclusion power (EPR) and cumulative inverse match power (Inverse MPR) of 17 microsatellite markers.

Table S3 Mean number of alleles segregating per marker, mean expected (EH) and observed heterozygosity (OH),

mean polymorphic information content (PIC), mean exclusion power (EPR) and standard errors (SE) averaged for 17 microsatellite markers by breed; and number of loci required to reach EPR of 0.99 and number of loci required to reach the match power (MPR) in the AKC represented breeds.

References

- Botstein D., White R.L., Skolnick M.H. & Davies R.W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* **32**, 314–31.
- Francisco L.V., Langston A.A., Mellersh C.S., Neal C.L. & Ostrander E.A. (1996) A class of highly polymorphic tetranucleotide repeats for canine genetic mapping. *Mammalian Genome* **7**, 359–62.
- Halverson J., Dvorak J. & Stevenson T. (1995) Microsatellite sequences for canine genotyping. *US Patent* 05874217.
- Jamieson A. (1965) The genetics of transferrin in cattle. *Heredity* **20**, 419–41.
- Ostrander E.A., Sprague G.F., Jr & Rine J. (1993) Identification and characterization of dinucleotide repeat (CA)_n markers for genetic mapping in dog. *Genomics* **16**, 207–13.
- Ostrander E.A., Mapa F.A., Yee M. & Rine J. (1995) One hundred and one simple sequence repeat-based markers for the canine genome. *Mammalian Genome* **6**, 192–5.
- Savolainen P., Zhang Y., Luo J., Lundeberg J. & Leitner T. (2002) Genetic evidence for an East Asian origin of domestic dogs. *Science* **298**, 1610–13.
- Vilà C., Savolainen P., Maldonado J.E., Amorim I.R., Rice J.E., Honeycutt R.L., Crandall J., Lundeberg K.A. & Wayne R.K. (1997) Multiple and ancient origins of the domestic dog. *Science* **276**, 1687–89.
- Werner P, Mellersh C.S., Roducha M.G., De Rose S., Acland G.M., Prociuk U., Wiegand N., Agurre G.D., Henthorn P.S., Patterson D.F. & Ostrander E.A. (1999) Anchoring of canine linkage groups with chromosome-specific markers. *Mammalian Genome* **10**, 814–23.