Comparing two commercial domestic dog (*Canis familiaris*) STR genotyping kits for forensic identity calculations in a mixed-breed dog population sample

S. Kanthaswamy*†‡, R. F. Oldt*†‡, M. Montes* and A. Falak*

*School of Mathematics and Natural Sciences, Arizona State University (ASU) at the West Campus, Glendale, AZ 85306, USA. †California National Primate Research Center, University of California, One Shields Ave, Davis, CA 95616, USA. ‡Evolutionary Biology Graduate Program, School of Life Sciences, Arizona State University, Tempe, AZ 85281, USA.

Summary

Almost half of all US households own a dog (*Canis familiaris*). Though these household pets can attack humans and other animals, they are also frequently victims of cruelty, neglect and theft. In human-oriented investigations, the tendency of domestic dogs to leave behind physical traces (such as hair) can serve as valuable links between crime scenes and suspects/victims. This demonstrated utility of canine biological evidence has created demand for genotyping marker sets for canine forensic genetic testing. Through research and casework, short tandem repeat (STR) panels have been shown to be very efficient for identity and parentage testing in dogs. However, there is an absence of comparative studies between different canine forensic identification kits. The Thermo Fisher Scientific Canine Genotypes™ Panel 1.1 and 2.1 Kits were originally designed and developed for routine and forensic use respectively, although both kits can be used for either purpose. In this study, we evaluated both STR panels to determine how critical forensic genetic metrics are affected by panel-to-panel variation in marker composition and design. Our results show that although STR panel composition can influence estimates such as inbreeding, combined power of discrimination and combined probability of exclusion, greater average allele number values exhibited across all markers in Panel 2.1 facilitated significantly more precise estimates of random match probability (RMP) and combined probability of identity. Furthermore, we demonstrate that a theta (θ) correction of 0.09 can be used to conservatively adjust RMPs generated from a small reference database of fewer than 50 samples, confirming that Panel 2.1 is a more robust forensic genotyping system than is Panel 1.1. for domestic dogs. We also demonstrate that opportunistic local sampling of fewer than 50 mixed-breed dogs can produce sufficient discriminatory and exclusionary power with either genotyping kit.

Keywords canine DNA, forensic analysis, STR profiling

Domestic dogs (*Canis familiaris*) are one of the most common companion animals in the world. In the US, approximately 90 million dogs are owned as pets with over 60 million households across the country owning at least one (ASPCA.org). As 4–5 million people a year are bitten by dogs (Gilchrist *et al*. 2008), it is critical to be able to accurately identify the individual dog responsible based on a biological sample, such as saliva left on a bite wound (Kanthaswamy *et al*. 2009). This capability is also necessary in cases for which a dog is not the aggressor but when canine evidence—such as hair, feces, or blood—is uncovered in a criminal or civil investigation involving humans (Clarke & Vandenberg 2010; Amorim & Budowle 2017). Before the discovery and widespread use of short tandem repeat (STR) markers in forensic identity testing, information obtained from canine samples was limited to species designation or, at best, breed assignment (Peabody *et al*. 1983; Carracedo *et al*. 1987; Amorim & Budowle 2017). Presently, STR markers are successfully employed in both identity and parentage testing in dogs (Zajc *et al*. 1994; Shutler *et al*. 1999; Padar *et al*. 2001; DeNise *et al*. 2004; Kanthaswamy *et al*. 2009). Like in...
humans, many well-studied dog STR markers are tetranucleotide and are affected by population subdivision (Zaic et al. 1997; Morera et al. 1999; Irlon et al. 2003; Kanthaswamy et al. 2009). Although the efficacy of these STR markers has been demonstrated both in research and in casework (DeNise et al. 2004; Hulverson & Basten 2005; Clarke & Vandenberg 2010), there is an absence of comparative studies between different canine forensic identification kits.

The Canine Genotypes™ Panel 1.1 Kit (catalog number: F860S) and the Canine Genotypes™ Panel 2.1 Kit (catalog number: F864S), manufactured by Thermo Fisher Scientific, are two commercial STR amplification kits that are currently available for identity testing and parentage verification in domestic dogs (Kanthaswamy 2015). Panel 1.1 was developed for service laboratories that perform routine parentage and identity testing. The 18 STR markers in Panel 1.1 are among the core markers recommended by the International Society for Animal Genetics (ISAG) parentage verification in dogs (http://www.isag.us/Docs/consignments/2005ISAGPanelDOG.pdf). Panel 2.1 was developed by the University of California, Davis and Finnzymes Oy (Espoo, Finland) specifically for forensic analysis of dog material, including identity and parentage testing (Budowle et al. 2005; Dayton et al. 2009; Kanthaswamy et al. 2009; Tom et al. 2010). The suites of STR markers that constitute Panels 1.1 and 2.1 are listed in Table S1, which also includes a description of each marker’s repeat motif and chromosomal location.

Apart from the tetranucleotide repeat locus FH2054, which Panel 1.1 shares with Panel 2.1, all other STR markers in Panel 1.1 contain dinucleotide repeat motifs (Table S1). With the exception of VWF.X, a mostly hexameric marker, and FH3377, a mostly pentameric marker, all other STRs in Panel 2.1 are tetrameric (Table S1). Compared to markers with dinucleotide repeat motifs, those with repeat units of four or more tend to generate reduced stutter artifacts (Carracedo et al. 1998; Dayton et al. 2009). On the other hand, the complex repeat sequences of the larger alleles of the hexameric and pentameric markers in Panel 2.1 contain sequence variation within repeat units that lead to different repeat motifs in these alleles (Tom et al. 2010). Moreover, the order of units within these alleles can be different, and even if these alleles have uniform lengths, their sequence complexity can prevent allelic bins from being designated by whole integers (Gill et al. 1996; Tom et al. 2010). The characterization of these complex alleles has been described by Tom et al. (2010) by means of complete sequence analyses.

It is known that variation in marker composition among different kits may give different genotyping results that bias statistical parameters in population genetic analyses (Westen et al. 2014). Although each genotyping kit was developed for different end users, both kits are applicable for forensic and non-forensic identity and parentage testing in dogs. Due to their overlapping utility, our goal was to evaluate if one kit performed better than the other in regard to identity and parentage testing. Given that both panels cost approximately $14 per reaction, a better understanding of these kits’ performance will help laboratories avoid redundant reagent expenses.

Most animal forensic laboratories do not have the resources to establish large-scale population databases for forensic casework (Kanthaswamy 2015). Population sampling efforts for building dog STR databases tend to be restricted to opportunistic sampling from local sources such as veterinary clinics and dog owners (Verscheure et al. 2013). Not only have local dog samples been shown to be representative of a much broader geographic sampling (Himmelberger et al. 2008), but also estimates of genetic diversity based on population samples proximate to the crime scene tend to exhibit a more similar genetic structure represented by the forensic sample than do those collected from different and diverse geographic regions (Himmelberger et al. 2008; Kanthaswamy et al. 2009).

No comparative studies examining the genotyping results from both these commercial kits have been published. Although Kanthaswamy et al. (2009) have validated Panel 2.1 with a DNA database comprising mixed and purebred dogs from a wide geographic sampling to assess the distribution of genetic variation across the US, little is known about the samples on which Panel 1.1 was evaluated. As most forensic cases involving canine evidence are local in nature, the broad geographic sampling strategy employed by Kanthaswamy et al. (2009) to create a national database may not be necessarily relevant for local caseloads. Differences in sample sets will yield variation in genotyping data, which can influence estimates of population genetic metrics. As such, this study was aimed at investigating the descriptive statistics and discriminatory power of both kits for parentage and individualization based on a set of 40 random unrelated (i.e. unrelated to the third generation) local mixed-breed dogs, as well as 25 and 30 individuals can provide estimates of allele frequencies for accurately characterizing the genetic composition of a population (Hale et al. 2012). When a local dog forensic STR database is not readily available for casework, then this sampling range is an easily achievable target for a typical forensic case involving dogs (Scharnhorst & Kanthaswamy 2011). Himmelberger et al. (2008) and Kanthaswamy et al. (2009) have argued that samples from mixed-breed dogs represent the majority of forensic caseloads because material from these dogs would likely be present far more often than would biomaterial from purebred dogs. Furthermore, breed information is not critical for routine casework; Wictum et al. (2013) have demonstrated that high estimates of heterozygosity (~86%) can be obtained from sample sets irrespective of breed.

DNA was collected and extracted from buccal swabs using the Qiagen QuiaAmp DNA Blood Mini Kit, following the manufacturer’s protocol. The extracts were quantified...
with a Qubit 3.0 Fluorometer (Thermo Fisher Scientific), and PCR was amplified following the manufacturer’s protocol for each kit using a ProFlex PCR thermocycler system (Thermo Fisher Scientific). Amplified STR fragments were run on a 3500 Genetic Analyzer (Applied Biosystems) with POP-4 polymer and the LIZ 500 Sizing Standard (Thermo Fisher Scientific) in accordance with Panels 1.1 and 2.1 user manuals. STR profiles were analyzed with GENEMAPPER® ID-X software v1.4 (Life Technologies), and allele sizes were binned into corresponding categories using the FLEXIBIN v2.0 program and methods described by Amos et al. (2007).

The exact probability test in ARLEQUIN v3.5.1.2 (Excoffier et al. 2007) was used to test for the presence of linkage disequilibrium (LD, or the non-random association of genotypes occurring at different markers) between pairs of STR markers in each panel. As the sample number was not large (n = 40), only pairwise LD analyses were performed, because linkage equilibria at three or more markers would have caused the associated frequency table to become too sparse for any meaningful conclusions. Unbiased estimates were made through randomization (1000 iterations; 10,000 dememorizations), and the Markov-chain method was used to find pairs of markers that violated linkage equilibrium conditions at the 0.05 level of probability. A sequential Bonferroni-type procedure was used to correct for multiple significance tests. The number of markers significantly linked to a given marker, compared to the 17 total possible pairs in each panel, is reported as the LD proportion. In addition to examining LD between markers, deviation from Hardy-Weinberg equilibrium (HWE), average allele number per locus (n_a), inbreeding coefficient (F_is) (Wright 1978) and observed and expected heterozygosity within each locus (h_o and h_e respectively) as well as averaged across loci (Ho and He respectively) were analyzed using ARLEQUIN software. Analysis of molecular variance (AMOVA; Excoffier et al. 2007) was also performed by means of the same software program to ascertain each panel’s ability to detect the levels of extant genetic structure among the canine samples. Statistical significance of the AMOVA values was estimated by a permutation test (10,000 permutations) at the 0.05 level of probability. Modified versions of ONPOP 200.1 (http://www.cstl.nist.gov/strbase/populationdata.htm) and POWERSTATS v1.2 (Promega) were used to calculate random match probability (RMP; the probability of a random individual having that exact DNA profile) (National Research Council 1996), combined probability of identity (CPI; the probability that two individuals chosen randomly will have the same DNA profile) (Sensabaugh 1982; Waits et al. 2001), combined probability of exclusion (CPE; the probability that a random individual would be excluded as a DNA profile contributor in a mixture analysis) (Jumieson & Taylor 1997; Ladd et al. 2001) and combined power of discrimination (CPD; the probability of differentiating random individuals) (Fisher 1951; National Research Council 1996).

Full DNA profiles from all 40 dog samples that were generated using Panels 1.1 and 2.1 were used to calculate estimates of RMP, CPI, CPE and CPD. Because canine population structure is known to exist and can bias these population genetic statistics (Kanthaswamy et al. 2009; Wictum et al. 2013), the RMP estimates were calculated using a theta (θ) correction (National Research Council 1996) of 0.09 based on the F_{ST} genetic differentiation value approximated by Kanthaswamy et al. (2009). Incorporating θ into the RMP equation adjusts for inflated homozigosity observed in strongly structured and isolated populations.

Comparisons of genotyping results from Panels 1.1 and 2.1 based on the same dog samples helped reveal critical differences in population genetic metrics that were due primarily to different marker composition and informativeness of those panels. Eleven markers in Panel 1.1 and only five markers in Panel 2.1 were found to be statistically linked to at least one other marker (Table 1). Statistically, when the 5% significance threshold is adopted, a few significant results among the 306 total pairwise comparisons tests could have been expected, even in a situation of total independence. However, Panel 1.1 exhibited more than twice the number of markers (or about two-thirds of its marker composition) in LD than did Panel 2.1. Additionally, 12 markers in Panel 1.1 were found to be within Hardy-Weinberg expectations compared to 11 markers in Panel 2.1. The shared marker between the panels, FH2054, was in HWE in both datasets. A marker-by-marker comparison of genetic diversity metrics for each of the 18 STR panels, such as n_a, h_o, h_e, are shown in Table 1.

Allele numbers per markers are typically highly associated with the size of the sample and also uneven allele frequencies at specific loci across samples. Despite using the same number of samples, the range of observed allele numbers per locus in Panel 2.1 was greater than in Panel 1.1 (i.e. from four to 19 alleles vs. from six to 13 alleles respectively), resulting in a significantly higher N_A estimate (9.94) compared to Panel 1.1 (8.74). Per-marker inbreeding coefficients (F_is) ranged from −0.03 to 0.31 in Panel 1.1 and from 0.03 to 0.26 in Panel 2.1. h_o estimates ranged from 0.52 to 0.82 and from 0.41 to 0.86 per marker for Panels 1.1 and 2.1 respectively; h_e estimates were calculated to be as low as 0.67 and 0.55 to as high as 0.89 or 0.94 per marker in both the panels. Average estimates of genetic diversity metrics across all 18 STR markers in each panel are shown in Table 2a. The coefficients of inbreeding (F_is) were very similar between both panels (i.e. 0.12 vs. 0.11 for Panels 1.1 and 2.1). Panels 1.1 and 2.1 also exhibited comparable H_o estimates (0.70 and 0.71 respectively), whereas the estimates of H_e for both panels were identical (0.79) (Table 2a). The AMOVA for both panels revealed like estimates of variation at the individual level (88% and 89% of the total genetic variation respectively) (Table 2b). Both panels are in agreement that genetic
variation among individuals represents only 11–12% of the total genetic variation observed in this study (Table 2b). The complete AMOVA results are presented in Table S2.

Panel 2.1 yielded an average RMP two orders of magnitude smaller than for Panel 1.1 (7.66 $\times$ 10^{-21} and 1.25 $\times$ 10^{-21} respectively) when no $\theta$ adjustment was incorporated (Table 2c). When a $\theta$ value of 0.09 (Kanthaswamy et al. 2009) was used to calculate a more conservative RMP (i.e. to account for genetic variation attributed to the various combinations of different breeds represented in the mixed-breed sample) the Panel 2.1 RMP increased by one order of magnitude (2.29 $\times$ 10^{-22}), whereas the Panel 1.1 RMP did not change in a significant way (1.45 $\times$ 10^{-21}). Despite the use of fewer samples in this study ($n =$ 40), both Panels 1.1 and 2.1 produced RMPs that were within the range of human RMPs (10^{-12} to 10^{-10}) based on 13 to 20 STR loci (Ng et al. 2018). Both canine panels also yielded far more powerful RMPs than did ‘Meowplex’, a 10-loci domestic cat STR panel producing RMPs on the order of 10^{-7} to 10^{-13}. With a set of 12 STRs, Fernandez et al. (2013) obtained an RMP of 10^{-11} in cattle. The average CPI value for Panel 1.1 was 1.15 $\times$ 10^{-31}, and the overall CPI was 9.92 $\times$ 10^{-24} compared to those based on Panel 2.1, which were 3.85 $\times$ 10^{-15} and 6.37 $\times$ 10^{-25} respectively. Therefore, the CPI value based on Panel 2.1 was four orders of magnitude smaller than the Panel 1.1 estimate, whereas the CPE was only one order of magnitude smaller. Both panels yielded more discriminatory CPI and exclusionary CPE estimates than did species-specific panels that were developed for bovines and equines respectively (Zhang et al. 2010; van de Goor et al. 2011). Both those studies employed 17 STRs, and although Zhang et al. (2010) generated a CPI of 10^{-16}, the most powerful CPE estimate van de Goor et al. (2011) calculated was approximately 10^{-18}. The average CPE estimate produced in the present study did not differ significantly between Panels 1.1 and 2.1 (Table 2c).

The most fundamental difference observed between the two canine genotyping panels was the average numbers of alleles ($N_a$) in Panel 2.1 due to the wider range of alleles captured with tetranucleotide STR loci. As a result, RMP

© 2018 Stichting International Foundation for Animal Genetics, 50, 105–111
and CPI calculations revealed that Panel 2.1 can be used to individualize biological samples slightly more discriminatively than Panel 1.1. The implication of this much greater ability to match a profile to an individual animal is significant given that variation among the individual dogs included in this study was only between 11 and 12%, as demonstrated by the AMOVA results.

The departure from HWE by some of the markers in both panels (i.e. six in Panel 1.1. and seven in Panel 2.1) may be attributed to the use of predominantly mixed-breed dogs in this study (Kanthaswamy et al. 2009). Departure from HWE may also occur due to inbreeding and population substructure. Regionally representative dog datasets have been shown to exhibit very little genetic differentiation (FST = 0.002) (Kanthaswamy et al. 2009). Conversely, inbreeding estimates within geographic regions were high (FIS = 0.10), indicating that there is an increased number of homozygotes in almost all regional dog populations in the US including the one from which the present study samples were collected (Kanthaswamy et al. 2009).

Linkage disequilibrium among markers typically results from their co-segregation due to physical linkage and/or inbreeding. Because the markers on both panels are biologically unlinked, the departure from linkage equilibria may be explained by the high inbreeding rates within the study population (Kanthaswamy et al. 2009). The high overall FIS values of 10% or more obtained in this analysis and also by Kanthaswamy et al. (2009), in conjunction with much lower Ho (~70%) than Ht (~80%) values in both studies, suggest that inbreeding alone is a principal factor that has shaped the population genetic structure of domestic dogs in the US. Despite this, a highly conservative estimate of RMP (2.29 × 10^-32) based on Panel 2.1 was derived after a θ correction of 0.09 was employed. The generation of conservative RMP estimates is driven by a desire not to overstate the strength of a forensic DNA match (National Research Council 1996).

The results from this study are consistent with STR data for humans, for which high allelic diversity at carefully selected markers elevates the power of discrimination. As the STRs in Panel 2.1 are primarily tetranucleotide markers, the interpretation of peak size was expected to be more straightforward than for those of the dinucleotide markers in Panel 1.1. However, both kits performed equally well in yielding genotyping results that could be readily interpreted from the pristine DNA samples used in this study. Although Panel 2.1 has previously undergone developmental validation (Dayton et al. 2009; Kanthaswamy et al. 2009; Tom et al. 2010) and has been used in forensic casework (Ogden et al. 2012), there is no published literature on the use of Panel 1.1 in forensic investigations nor is there any evidence that this panel has been validated for such use.

Based on their genetic diversity metrics, both panels can be regarded as highly useful for canine identity and sample matching. Although diversity estimates such as Ho and Ht derived from the panels were comparable, a CPI of 3.85 × 10^-35 based on Panel 2.1 that is more discriminative than that of Panel 1.1 (1.15 × 10^-31), coupled with conservative RMP results, means that the chance of finding another random, unrelated dog with the same DNA profile is much lower with Panel 2.1. As the CPE for Panel 2.1 was also smaller than 1.1, the likelihood of correctly excluding an individual not bearing the inculpating (or incriminating) alleles is much greater. Our findings demonstrate that although opportunistic, local sampling of fewer than 50 dogs can produce sufficient discriminatory and exclusionary power regardless of genotyping kit, Panel 2.1’s predominantly tetranucleotide marker set and enhanced processing efficiency and reliability (Dayton et al. 2009; Kanthaswamy et al. 2009) is most capable of distinctly identifying individual animals with high precision. Panel 2.1 is, therefore, a highly optimized and well-characterized genotyping system that not only meets the needs and standards for forensic testing but will also be of great use for both routine service-oriented identity and parentage testing.

At the time of this writing, Thermo Fisher, the manufacturer of both canine panels, announced that Panel 1.1 will be discontinued from production as of November 2018 and will be replaced with the Canine ISAG STR Parentage Kit (catalog number: A36139). Priced at $10 to $13 per reaction, the new panel consists of the 22 mostly dinucleotide STRs that were recommended by ISAG in 2014 for parentage testing purposes. Although it is anticipated that the additional markers, including AHTh130, REN105L03 and REN64E19, may probably result in lowering RMP values and improving other forensic metrics (van de Goor et al. 2010; Ng et al. 2018), the new canine parentage panel has yet to be validated for use in forensic investigations.

Acknowledgements

We would like to thank the Arizona State University West and Tempe campuses, the greater Phoenix-area community and especially Dr. Gwyneth Gordon and the Anbar Lab at the ASU School of Earth and Space Exploration for contributing dog buccal swabs. We also thank Mr. Christopher C. Adams at Thermo Fisher Scientific for giving us the Canine Genotypes™ Panels 1.1 and 2.1 Kit on a pro bono basis.

References


**Supporting information**

Additional supporting information may be found online in the Supporting Information section at the end of the article: Table S1 Marker information for Canine Genotype™ Panels 1.1 and 2.1 Kits, Table S2 AMOVA results showing sources of variation in both panels.