

Developmental Validation of Short Tandem Repeat Reagent Kit for Forensic DNA Profiling of Canine Biological Material

Aim To develop a reagent kit that enables multiplex polymerase chain reaction (PCR) amplification of 18 short tandem repeats (STR) and the canine sex-determining Zinc Finger marker.

Methods Validation studies to determine the robustness and reliability in forensic DNA typing of this multiplex assay included sensitivity testing, reproducibility studies, intra- and inter-locus color balance studies, annealing temperature and cycle number studies, peak height ratio determination, characterization of artifacts such as stutter percentages and dye blobs, mixture analyses, species-specificity, case type samples analyses and population studies.

Results The kit robustly amplified domesticated dog samples and consistently generated full 19-locus profiles from as little as 125 pg of dog DNA. In addition, wolf DNA samples could be analyzed with the kit.

Conclusion The kit, which produces robust, reliable, and reproducible results, will be made available for the forensic research community after modifications based on this study's evaluation to comply with the quality standards expected for forensic casework.

Melody Dayton¹, Mikko T Koskinen², Bradley K Tom³, Anna-Maria Mattila², Eric Johnston⁴, Joy Halverson⁵, Dennis Fantin⁴, Sue DeNise⁴, Bruce Budowle^{6,7}, David Glenn Smith^{2,8,9}, Sree Kanthaswamy^{2,8,9}

¹The Jan Bashinski DNA Laboratory, California Department of Justice, Richmond, Calif, USA

²Finnzymes Diagnostics, Espoo, Finland

³Graduate Group in Forensic Science, University of California–Davis, Davis, Calif, USA

⁴MMI Genomics, Inc., Davis, Calif, USA

⁵QuestGen Forensics, Davis, Calif, USA

⁶Department of Forensic and Investigative Genetics, University of North Texas Health Science Center, Fort Worth, Tex, USA

⁷Institute of Investigative Genetics, University of North Texas Health Science Center, Fort Worth, Tex, USA

⁸Department of Anthropology, University of California–Davis, Davis, Calif, USA

⁹California National Primate Research Center, University of California–Davis, Davis, Calif, USA

Received: April 2, 2009

Accepted: April 29, 2009

Correspondence to:

Sree Kanthaswamy
Department of Anthropology
University of California
One Shields Avenue
Davis, CA 95616, USA
skanthaswamy@ucdavis.edu

Dogs (*Canis familiaris*) were domesticated in East Asia almost 15 000 years ago (1). Since then, domestic dogs have been selectively bred to be outdoor working animals as well as household companions (2). Seventy-two million pet dogs are estimated to live in the USA (3), and because many dog owners live in close proximity to their pets, canine DNA evidence may be associated with crimes. Shed dog hairs can be transferred among individuals involved in a crime and can link a suspect to a crime scene, a suspect to a victim, or a victim to a crime scene (4,5). Dog evidence can also be used in dog attack cases and abuse cases (6).

Several studies have generated genotype and sequence information for multiple canine short tandem repeat (STR) markers (7-12), and more recently, various laboratories have multiplexed different canine STR markers for co-amplification in a single polymerase chain reaction (PCR). Initially, multiplexes usually contained from 3 to 10 primer pairs that required multiple amplifications to type 15 or more markers (4,8,13,14). Such protocols are very similar to that of the human DNA analyses Profiler and COfiler kits (Applied Biosystems, Foster City, CA, USA); a sample of forensic interest would be amplified by both kits in order to type all 13 Combined DNA Index System (CODIS) core loci. However, when a forensic sample is limited in quantity, it is much more efficient, as well as cost- and time-effective, to amplify all polymorphic markers of interest in a single multiplex PCR.

Similar to the human Identifiler® kit (Applied Biosystems) designed to co-amplify 15 human STR loci and the amelogenin locus (15), a canine beta-version STR reagent kit was developed that enables amplification of 18 STR loci and the canine sex-linked Zinc Finger marker (16). As a first evaluation of the proposed multiplex kit, similar developmental validation studies were carried out as described for validation of the Identifiler® kit (15). The performance of the reagent kit is discussed in comparison with that of previous typing kits utilized for forensic casework, primarily Identifiler® (15), Stockmark® Canine I and II panels (Applied Biosystems) (4), the Eichmann et al (13) canine multiplexes, and the Menotti-Raymond et al (17) "Meowplex."

METHODS

The canine STR reagent kit is manufactured by Finnzymes Oy (Espoo, Finland). All experiments reported in this study are on a pre-commercial "beta-test" version of the kit. Based on the findings reported herein, the manufacturer will make final modifications and improvements to the kit

prior to its commercial release to the forensic science community.

The canine reagent kit contains 19 primer pairs and associated reagents to amplify 18 STR loci and one gender determination locus. The canine genome contains 38 matched pairs of autosomal chromosomes and 2 sex chromosomes. A larger number of loci than typically included in comparable panels, such as the 13 CODIS markers used for humans, was employed in anticipation that inbreeding in full-breed dogs would require a greater exclusion probability to achieve satisfactory effectiveness. Table 1 presents a list of these markers and information on repeat motifs, size ranges, map location, and the fluorescent dye tags attached to the 5' end of the forward primers, as they appear during laser excitation using filter set G5 (Applied Biosystems). A few markers are located on the same chromosome. The FH3377 and FH2107 loci are on chromosome 3; the FH2054 and PEZ05 loci are on chromosome 12; the FH2017 and FH2088 loci are on chromosome 15; and the PEZ16 and vWFX loci are on chromosome 27. A separate study by Kanthaswamy et al (18) has demonstrated that alleles of these physically linked loci segregate independently, ie, there was no detectable linkage disequilibrium for these syntenic marker pairs.

Three different types of biological samples were typed to evaluate the 19-plex reagent kit. Fifty three liquid blood samples, representing 34 different dog breeds, were kindly provided by a West Sacramento veterinary diagnostic center. Eleven buccal samples, representing 7 different breeds, were collected from dogs throughout the state of California, primarily from the Los Angeles area, the San Francisco Bay area, and the Sacramento area. Plucked and shed hair samples were collected from 6 dogs, 4 of which had also provided buccal samples. A total of 67 dogs were typed during the duration of the study.

DNA from liquid blood samples was extracted using the Qiagen BloodMini kit following the kit's liquid blood extraction protocol (Qiagen, Hilden, Germany). DNA from buccal samples was extracted using the Epicenter® Catch-All (Epicenter, Madison, WI, USA) swabs according to the manufacturer's recommendations. Hair samples from 6 of the dogs were extracted using Epicenter QuickExtract DNA Extraction solution buffer (Epicenter) according to the manufacturer's instructions. Hair samples from 6 dogs were also extracted using an unpublished California Department of Justice Organic Extraction protocol, with the exception that an Amicon® Ultra-4 (Millipore

TABLE 1. Marker information for the Finnzymes canine short tandem repeat reagent kit

Marker	Repeat motif	Dye color*	Size range (bp) [†]	Map location [‡]	Reference
PEZ02	(GGAA) _n	blue	104-145	cfa 17 (13276076-13276209) [§]	6
Zinc Finger	Not applicable	blue	159-164	cfa X/Y	13
PEZ17	(GAAA) _n	blue	190-225	cfa 4 (71904833-71905038)	18
FH2017	AGGT _(m) AGAT _(n) GATA _(o)	blue	256-276	cfa 15 (37914470-37914741)	18
FH2309	(GAAA) _n	blue	339-428	cfa 1 (85772974-85773377)	18
PEZ05	(TTTA) _n	green	92-117	cfa 12 (60326434-60326541)	18
FH2001	(GATA) _n	green	118-160	cfa 23 (50961325-50961475)	7
FH2328	(GAAA) _n	green	171-213	cfa 33 (19158127-19158477)	8
FH2004	(AAAG) _n	green	232-326	cfa 11 (32161381-32161621)	18
FH2361	GAAA	green	322-439	cfa 29 (19723594-19723782)	18
PEZ21	(AAAT) _n	yellow	83-103	cfa 2 (36438658-36438751)	9
FH2054	(GATA) _n	yellow	139-177	cfa 12 (37914504-37914739)	7
FH3377	GAAAA	yellow	183-305	cfa 3 (78748898-78749090)	18
FH2107	(GAAA) _n	yellow	290-426	cfa 3 (83830247-83830574)	7
FH2088	(TTTA) _(m) (TTCA) _(n)	red	94-138	cfa 15 (53905651-53905779)	7
vWF.X	(AGGAAT) _n	red	151-187	cfa 27 (41977918-41978074)	11
FH2010	(ATGA) _n	red	221-243	cfa 24 (5196383-5196605)	7
PEZ16	(GAAA) _n	red	280-332	cfa 27 (10305692-10305995)	18
FH3313	GAAA	red	340-446	cfa 19 (24606038-24606459)	18

*Dye colors are listed as they appear in electrophoresis with filter set G5: 6FAM (blue), VIC (green), NED (yellow), PET (red), and LIZ (orange, size standard).

[†]Allelic marker size ranges determined by evaluating the genotypes of 667 different dogs from the United States.

[‡]The chromosome locations for all markers were confirmed in the University of California Santa Cruz Canine Genome database (<http://genome.ucsc.edu>) and on the National Center for Biotechnology Information Basic Local Alignment Search Tool (BLAST, <http://www.ncbi.nlm.nih.gov>)

[§]cfa refers to *Canis familiaris*.

Bedford, MA, USA) centrifugal filter was used instead of a Centricon®-100 (Millipore) filter. Five of the 7 dog samples were extracted using both methods.

Canine species specificity of the 19-plex was assessed by amplifying 5 ng non-canine DNA samples in duplicate. Non-canine DNA samples from various species were provided by the Jan Bashinski DNA Laboratory, Method Development Group of the California Department of Justice, Richmond, CA, USA. DNA was extracted from cat, pig, horse, cow, and chicken whole blood and quantified by an electrophoretic method in a 1% agarose gel containing ethidium bromide. Other previously quantified genomic DNA samples provided by the Jan Bashinski laboratory included: fish, monkey, rat (Zyagen Laboratories, San Diego, CA, USA), mouse (Promega, Madison, WI, USA), *Bacillus subtilis*, *Staphylococcus epidermidis*, *Candida albicans* (ATCC, Manassas, VA, USA), *Escherichia coli*, and *Clostridium perfringens* (Sigma, St. Louis, MO, USA) (19). A human buccal sample was extracted using the Epicenter QuickExtract protocol (Epicenter). Five nanograms of DNA from each species' sample were amplified in duplicate using the canine STR reagent kit. The F-863 canine positive

DNA control originating from a female Cocker Spaniel was provided in the reagent kit. Although stated on the tube that the canine control was a 1 ng/μL concentration, quantification with PicoGreen® (Invitrogen, Carlsbad, CA, USA) indicated that the control was at 0.1 ng/μL concentration.

Seventeen samples from a dog bite case were used for this study. Thirteen of the samples were single shed hairs found on the clothing of the victim. All shed hairs were extracted using the chelex extraction method. Four of the samples were saliva stains extracted from cuttings around tooth marks found in the victim's denim jeans. All saliva stains were extracted using Qiagen spin columns (Qiagen).

All DNA samples were quantified using Quant-iT PicoGreen® (Invitrogen). Prior to quantification, a solution was prepared containing 15 μL of PicoGreen® and 15 mL TE⁻⁴ buffer. One-hundred microliters of PicoGreen® solution were added to dilutions of 4 μL DNA sample and 96 μL of sterile water in a fluoroplate. K562 human DNA standards (Invitrogen) for 5 ng/μL, 2.5 ng/μL, 1.5 ng/μL, 0.75 ng/μL, 0.25 ng/μL, and 0.0 ng/μL were also pipetted in aliquots of 4 μL into 96 μL of sterile water and combined with 100

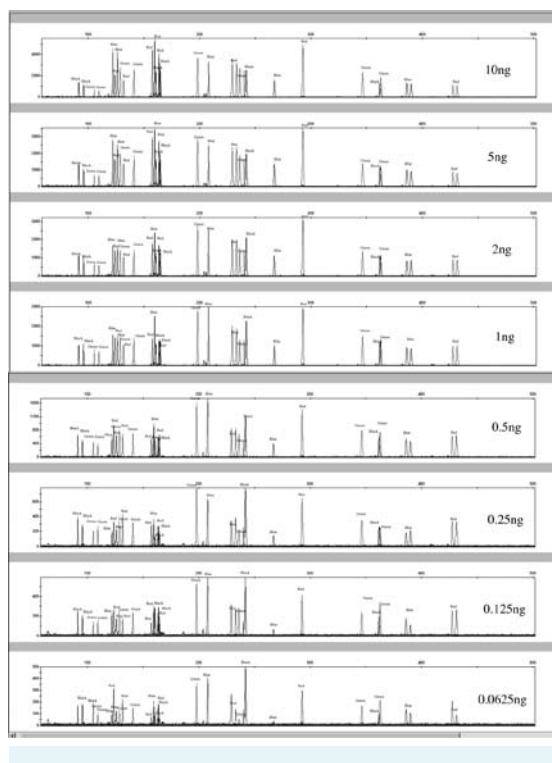
μL of Picogreen solution in the fluoroplate. A Fluoroskan Ascent (Thermo Scientific) fluorometer was used to detect and compare the fluorescence values of the known standards and those of the samples to determine DNA concentrations.

All PCR amplifications were carried out using the canine STR reagent kit. Each reaction contained 9 μL of Primer Mix, 9 μL of Master Mix (containing a modified version of the Phusion™ Hot Start DNA Polymerase, Finnzymes Oy), and 2 μL of DNA template. PCR conditions for the standard sample study, reproducibility study, sensitivity study, mixture study, peak height ratio/stutter percentage study, species specificity study and population study were: a 3-minute initialization step at 98°C, followed by 30 cycles of 3-step PCR (15 seconds of 98°C denaturation, 75 seconds of 60°C primer annealing step, 30 seconds of 72°C extension), and a final 5-minute extension at 72°C, as recommended by the manufacturer. The amplifications for these studies were performed on a GeneAmp® 9700 PCR thermocycler (Applied Biosystems). The Phusion polymerase makes use of double-stranded DNA binding domain (Sso7d) that is covalently linked to a *Pyrococcus*-like DNA polymerase domain (20). The DNA binding domain increases the processivity of the DNA polymerase by approximately 10 times (20). Unlike the *Thermus aquaticus* (Taq) polymerase, which is applied by all other STR typing kits, the Phusion polymerase also has 3' to 5' exonuclease activity (proofreading activity) that eliminates non-templated nucleotide additions during the PCR cycling process and makes the long final extension period redundant (21-23).

The sensitivity of the canine 19-plex was assessed by amplifying varying amounts of template DNA. Samples from 2 different dogs were used for the study. DNA was extracted from whole blood samples from a Golden Retriever and an Akita using the Qiagen BloodMini kit (Qiagen). Serial dilutions of both samples (10, 5, 2, 1, 0.5, 0.25, 0.125, and 0.0625 ng) were amplified in duplicate (Figure 1). The serial dilutions were quantified using Quant-iT PicoGreen® (Invitrogen). All PCR products were diluted 1:10 prior to loading 2 μL of sample into 10 μL of 0.02% LIZ/formamide solution for detection on the 3130 Genetic Analyzer. This was done to reduce off scale data for the 10 ng template amplified samples.

For the reproducibility study, 5 canine International Society of Animal Genetics (ISAG) standards were obtained as pre-quantified genomic DNA samples (http://www.isag.org.uk/ISAG/all/ISAG2006_CompanionAnimals.pdf). The

Figure 1.

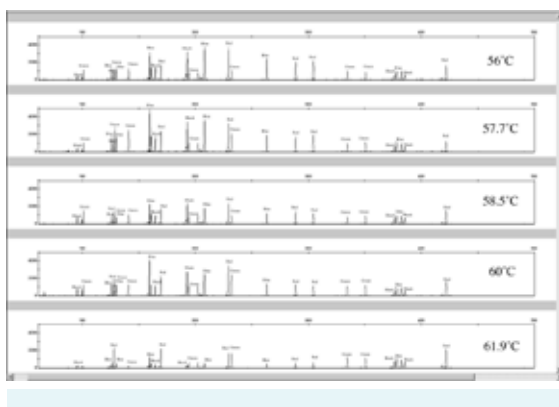


Sensitivity study using 10, 5, 2, 1, 0.5, 0.25, 0.125, and 0.0625 ng of a Golden Retriever sample.

samples were amplified and run at both the Molecular Anthropology Laboratory (MAL) at the University of California, Davis, CA, USA and at the Jan Bashinski DNA Laboratory. At the MAL, an Eppendorf Mastercycler epgradient thermocycler (Eppendorf) was used to amplify the ISAG samples, and an AB 3130 Genetic Analyzer containing POP-7 (Applied Biosystems) was used to resolve the amplified fragments. At the Jan Bashinski DNA Laboratory, a GeneAmp® 9700 PCR thermocycler (Applied Biosystems) was used to amplify the samples and an AB 3130 Genetic Analyzer containing POP-4 (Applied Biosystems) was used to resolve the amplified fragments. All samples were amplified in duplicate. All sizes were calculated using a global southern method with the 75, 100, 139, 150, 160, 200, 300, 350, 400, and 450 bp LIZ peaks.

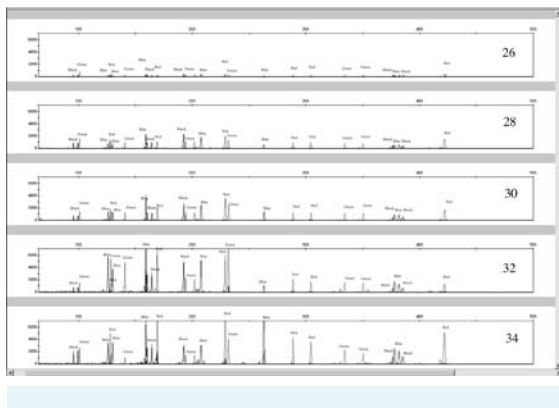
To assess the annealing temperature range of the canine 19-plex, DNA from 2 blood-extracted and one buccal-extracted samples were amplified on an Eppendorf Mastercycler epgradient thermocycler (Eppendorf, Hamburg, Germany) that creates an annealing tem-

Figure 2.



Representative electropherograms for the annealing temperature studies. Annealing temperatures of 56°C, 57.7°C, 58.5°C, 60°C, and 61.9°C were performed in duplicate amplifications for 4 canine DNA samples.

Figure 3.



Representative electropherograms of a cycle number study for 26, 28, 30, 32, and 34 cycle numbers. Duplicated amplifications of 3 different 1-ng canine DNA samples were performed at each cycle number.

perature gradient across the plate head. Exact annealing temperatures tested with all 3 samples (in duplicate) were 56°C, 57.7°C, 58.5°C, 60°C, and 61.9°C. The Eppendorf Mastercycler thermocycler (Eppendorf) was also used in the Reproducibility study using a 60°C annealing temperature. It is important to validate how performance prone a reagent kit is to variation in annealing temperature, because thermal cyclers may exhibit some variation between and within their thermal blocks. Amplification success of the entire suite of primers in the Finnzymes canine STR kit as a function of annealing temperature was investigated (Figure 2). For this study, 3 DNA samples (from 2 bloods and 1 buccal swab) were amplified in duplicate. One mi-

TABLE 2. Peak height ratios (PHR) for the Finnzymes canine 19-plex loci peak heights >500 relative fluorescence units)

Locus	Mean \pm SD*	N†	Lower range PHR‡	Lowest PHR observed
PEZ02	91.52 \pm 6.37	36	72.41	76.56
Zinc. Finger	51.20 \pm 7.29	27	29.33	35.76
PEZ17	87.34 \pm 7.56	29	64.66	67.79
FH2017	51.70 \pm 38.74	24	N/A	8.32
FH2309	82.66 \pm 10.63	42	50.77	59.38
PEZ05	86.31 \pm 7.90	29	62.61	65.96
FH2001	96.14 \pm 4.12	32	83.78	78.79
FH2328	92.51 \pm 7.00	41	71.51	62.92
FH2004	84.28 \pm 12.94	33	45.46	57.71
FH2361	86.85 \pm 9.78	39	57.51	47.43
FH2054	92.91 \pm 5.43	39	76.62	79.62
FH3377	83.98 \pm 9.38	38	55.84	59.62
FH2107	90.81 \pm 7.81	25	67.38	66.88
PEZ21	84.24 \pm 8.63	28	58.35	65.58
FH2088	78.57 \pm 12.12	29	42.21	48.62
vWF.X	92.51 \pm 4.94	24	77.69	82.56
FH2010	92.21 \pm 5.77	35	74.90	72.41
PEZ16	89.43 \pm 7.67	35	66.42	57.08
FH3313	85.10 \pm 10.50	39	53.60	46.06

*Standard deviation.

†N = number of heterozygote genotypes analyzed for each locus.

‡Lower range PHR is assessed as the mean - 3SD.

croliker of PCR product was added to 10 μ L of 0.02% LIZ/Formamide solution for electrokinetic injection.

To determine the optimal number of cycles for the canine STR kit, PCRs (1 ng template in a 20- μ L reaction) were carried out at 26, 28, 30, 32, and 34 cycles on a GeneAmp 9700 thermocycler (Applied Biosystems). DNA from 2 blood samples and one buccal sample was amplified in duplicate. The manufacturer (Finnzymes Oy) recommended 30 cycles for amplifying canine DNA. As expected, the overall signal intensity for most peaks increased with cycle number (Figure 3).

For the peak height ratio (PHR) study, 61 different dog samples were amplified to calculate PHRs for heterozygous allele pairs. Only samples that amplified peaks exceeding 500 relative fluorescence units (RFU) were used in the calculations. The mean PHRs and standard deviation for each marker are listed in Table 2.

All PCR products were separated by capillary electrophoresis on an Applied Biosystems 3130 Genetic Analyzer using 36 cm capillary filled with POP-4 polymer (Applied Biosystems). Unless otherwise stated, 1 μ L of PCR product was

combined with 0.21 μ L GeneScan-500 LIZ[®] Size Standard (Applied Biosystems) and 9.79 μ L Hi-Di Formamide (Applied Biosystems) (10 μ L of 0.02% LIZ/formamide solution) and pipetted into an injection plate. The plate was covered with Genetic Analyzer septa (Applied Biosystems) and the contents were thoroughly mixed and spun down. Samples were denatured at 95°C for 3 minutes then snap-cooled on a cold block for 3 minutes. A 5-second electrokinetic sample injection was performed with an injection voltage of 3 kV followed by a run voltage of 15 kV. The data collection software was set to detect G5 dye chemistry (see last footnote of Table 1). After collection, the data were analyzed with GeneMapper ID Version v3.2.1 Software (Applied Biosystems) using 50 RFU peak amplitude detection thresholds for all colors, excluding LIZ peaks which were set to a threshold of 100 RFU. Three standard deviations of the detectable “noise” RFU signal was used to establish the 50 RFU level of detection (LOD) of the ABI 3130xl instruments (Applied Biosystems) at the Jan Bashinski DNA Laboratory.

Three different types of peak balance were evaluated to assess the overall Finnzymes canine STR kit performance (15) and to identify possible primer pairs that amplify poorly in relation to the rest of the multiplex. Intracolor balance, intercolor balance, and PHRs were calculated using Microsoft Excel. Intracolor balance was calculated to assess the overall color balance within a dye lane. First, homozygous and heterozygous peaks within a dye lane had to be normalized; this was done by averaging the RFU values for each allele of heterozygous genotypes of polymorphic loci and by dividing the RFU of each homozygous peak by 2. The lowest normalized RFU value was then divided by the highest normalized RFU value and expressed as a percentage. Intercolor balance was calculated by first normalizing all of the loci peaks within a sample, then the lowest normalized RFU in the sample was divided by the highest normalized RFU in the sample (regardless of dye color) (15). PHRs were calculated for heterozygous genotypes for all markers by comparing the RFU values for each allele and dividing the lower RFU value by the higher value. The ratio was multiplied by 100 to express PHR as a percentage.

RESULTS AND DISCUSSION

Sensitivity study

For both the Golden Retriever and Akita samples amplified in duplicate, all peaks were detected when 10-0.125 ng of template DNA were amplified. In 1 of the 2 reactions using

0.0625 ng template DNA, the Golden Retriever sample exhibited a true homozygous peak at the FH2017 locus that was below the interpretational threshold of 50 RFU. In the duplicated Golden Retriever sample amplification, 1 of 2 heterozygous peaks at the FH2004 locus was below the interpretational threshold. All peaks were called (ie, RFU > 50) for the duplicated Akita sample amplifications at every template amount (including the 0.0625 ng samples). Because a 1:10 dilution was made for all PCR products prior to loading on the Genetic Analyzer, on a separate injection plate, 1 μ L of non-diluted PCR products was pipetted into 10 μ L of 0.02% LIZ/formamide solution for all template amplification amounts. Similar peak dropout results were obtained when no PCR product dilutions were made; on average the 1:10 diluted samples had RFUs 52% of that of the non-diluted samples with a 9.7% standard deviation. When no PCR product dilutions were made, increased pull-up artifacts were seen for the 10 ng template samples. Blue into green pull-up peaks ranged from 2031 to 3774 RFUs. Increased noise (peaks less than 50 RFU) were also present in the green and yellow dye channels for the 10-ng template samples and non-template adenylation (27.4% of the parent peak) was observed for marker FH2001 in the Akita sample. Blue into green pull-up artifacts were also seen for the 5-ng template samples. The off-scale and pull-up peaks indicate that 10 ng of high-quality DNA is too much template for the Finnzymes canine STR kit.

The PHRs for several markers (FH2328, FH2309, PEZ05, FH2010, vWF.X, FH3313, and FH2088) started to fall below 60% when DNA template amounts of 0.25 ng and below were amplified. Noticeable stochastic fluctuations were apparent for the FH2004 locus, which experienced allelic drop out in one of the 0.0625 ng amplifications; the other amplifications with 0.0625 ng (2 separate amplifications of the Akita sample and 1 amplification of the Golden Retriever sample) exhibited a PHR range between 41.5%-63.2% ($n=3$).

To verify that the template amounts prepared in serial dilutions were the amounts used in the experiment, the diluted samples were re-quantified using PicoGreen[®]. The minimum Akita sample (0.0625 ng) was 0.01 fluorescent units higher than the Golden Retriever sample when quantified using K562 DNA standards. The 0.125-ng samples were quantified at the same fluorescent values, meaning the serial dilution precision for the Golden Retriever and Akita samples was better for the 0.125 ng-DNA template compared with the 0.0625 ng-DNA template dilutions. At 0.125 ng-template, the Finnzymes canine

reagent kit was capable of detecting all profile peaks at >50 RFU (all but one peak at the FH2017 locus [63 RFU] was above 100 RFU) when no PCR product dilutions were made prior to adding sample to the LIZ/Formamide solution. This level of sensitivity is comparable with that of previous STR multiplexes developed. The "Meowplex" produces full STR profiles from 125 pg of template DNA (17). The 3 canine STR multiplexes (MP1, MP2, and MP3) developed by Eichmann et al (13) are capable of producing full profiles from 100 pg for MP1 and MP3, and 250 pg for MP2. A similar input DNA template range of 0.5-1.25 ng is also seen for the Identifiler® kit (15). However, a significant difference is that the Finnzymes canine multiplex includes a total of 19 markers, while the largest number of markers co-amplified in the other abovementioned studies was 15.

Reproducibility

As expected, the samples resolved with POP-7 migrated more rapidly through the 3130 capillary (75 bp LIZ peak resolved at ~1500 data points) compared with the samples resolved with POP-4 (75 bp LIZ peak resolved at ~2600 data points). However, more importantly, variations from 2-5 bp in size occurred between ISAG samples run at the different laboratories. Therefore, normalization of results is imperative for effective data comparison (24).

A binning program called Flexibin (25) was used to determine allele calls for the ISAG and kit provided positive control samples run at the different laboratories. Each set of data from the University of California and the Richmond DNA laboratory were analyzed separately with the Flexibin software. The canine control DNA included in the kit was used to calibrate the allele sizes observed. DNA amplification products from duplicate PCRs of the control animal sample subjected to electrophoresis at different time intervals demonstrate the reproducibility and precision of the assay within a particular laboratory (see "Run to run sizing;" <http://www.cstl.nist.gov/biotech/strbase/>). The level of reproducibility and precision might suggest that there is no need for an allelic ladder in inter-laboratory and within laboratory comparisons. However, these samples were typed in close time proximity. Due to the presence of only 1-2 alleles per locus, calibration may be inaccurate when manually comparing the binned data with the national canine allele database, particularly if the alleles of a locus span a large size range (18,26). The entire data set of raw and binned alleles used in their study is available at <http://www.cstl.nist.gov/biotech/strbase/>.

Eventually, an allelic ladder would mitigate the effects of migration variations between laboratories. Accurate concordance data (binned within 1 bp) are difficult until the allelic ladder is fully developed and utilized in subsequent analyses. Currently, a canine allelic ladder is being developed for the Finnzymes canine reagent kit, but without the use of a ladder, comparison of genotyping results between laboratories may be difficult.

Intracolor and intercolor balance

The intracolor balance percentages were calculated for 61 samples. Products in the red channel consistently had the best intracolor balance at 45.18% with a standard deviation of 12.85 (n=61) among the loci. All other dye colors had the following intracolor balances: green – $32.77 \pm 10.5\%$, blue – $33.63 \pm 14.18\%$, and yellow – $38.69 \pm 12.92\%$. For comparison, the amplification of non-degraded and uninhibited samples of human DNA using the Identifiler® kit was reported to attain an average intracolor balance of 50% or higher (15). Intercolor balance for 61 samples was estimated to be $21.97 \pm 7.67\%$, which is within the lower range of intercolor balance for the Identifiler® kit (20%-40%). Certain loci tended to have the lowest normalized RFU values within dye colors (Table 3).

The loci that exhibit the lowest or highest normalized RFU values are possible candidates for further primer concentration studies to improve intracolor and intercolor balance. In the blue channel, locus FH2017 demonstrated the lowest RFU value for 36 out of 61 samples, followed by locus FH2309, which had the lowest RFU value for 16 samples. The low RFU values for locus FH2017 might be partially explained by a suspected primer binding site mutation. The Zinc Finger locus usually exhibits the highest RFU value within the blue channel, but locus FH2017 also had the highest RFU value in 11 samples (presumably there is no primer binding site variant in these samples). By removing locus FH2017 from the intracolor balance calculations, an intracolor balance of 50.44 ± 13.94 was obtained for the blue channel. If the primer binding site mutation and overall peak signal for FH2017 can be addressed, then it is possible to increase the intracolor balance for the blue dye channel.

The locus with the lowest RFU value in the green channel is PEZ05. In well-balanced primer mixtures, primer pairs that produce the smallest PCR product lengths would be expected to produce the highest RFU signal because the polymerase is more efficient for amplifying smaller products (27). Because locus PEZ05 (the smallest PCR product

TABLE 3. Number of times a particular locus had the lowest and highest normalized relative fluorescence units (RFU) values for each dye channel*

Locus	Dye color channel	No. of times marker	No. of times marker	No.
		had the lowest normalized RFU	had the highest normalized RFU	
PEZ02	blue	5	4	61
Zinc Finger	blue	0	45	61
PEZ17	blue	4	0	61
FH2017	blue	36	11	61
FH2309	blue	16	1	61
PEZ05	green	53	2	61
FH2001	green	1	19	61
FH2328	green	1	4	61
FH2004	green	2	19	61
FH2361	green	4	17	61
PEZ21	yellow	14	0	60
FH2054	yellow	0	52	60
FH3377	yellow	0	8	60
FH2107	yellow	46	0	60
FH2088	red	32	4	61
VWF.X	red	3	5	61
FH2010	red	0	47	61
PEZ16	red	4	4	61
FH3313	red	22	1	61

*Trends in overall normalized peak intensity for each marker in the canine short tandem repeat kit were investigated because the intracolor balances for each dye channel were fairly low (ranging from 32.77 ± 10.5% for the green dye channel to 45.18% ± 12.85% for the red dye channel) in comparison with other commercial kits like Identifier® (which has intracolor balances for all colors 50% or higher).

in the green channel) exhibits the lowest RFU value, the current primer pairs for locus PEZ05 are not performing optimally with the other primers in the same dye channel. Interestingly, 3 different markers (FH2001, FH2004, and FH2361) competed for the highest RFUs in the green dye color. When the RFU values for locus PEZ05 are removed from the intracolor peak balance calculations, the balance for the green channel greatly improves to 70.45 ± 16.86%. Thus, if the signal from locus PEZ05 can be increased in relation to the other markers labeled in green, the intracolor balance can be improved.

For the yellow channel, locus FH2107 tends to produce the lowest normalized signal while locus FH2054 produces the highest signal, the difference being 1000-1700 RFU. By removing FH2107 from the intracolor calculations for yellow, the balance becomes 51.49 ± 16.20%; if the peak intensity of locus FH2107 can be addressed then the balance in the yellow channel can be improved.

The red channel exhibits the best averaged intracolor balance of 45.18%. Loci FH2088 and FH3313 tend to produce the lowest RFU signal, 1000-2000 RFU lower than normalized peaks produced by locus FH2010 (with the highest RFU peaks). When locus FH2010 is excluded from the balance calculations, intracolor balance is improved further to 54.60 ± 17.16%. By either increasing or decreasing the signal of loci creating either the lowest or highest RFUs respectively, it is possible to obtain intracolor peak balances greater than 50%.

Amplification – annealing temperature

The best overall intracolor balance occurred at 60°C, with average color balances being 42.41 ± 12.38% for blue, 52.99 ± 11.30% for green, 53.31 ± 12.49% for yellow, and 54.98 ± 17.71% for red. A 60°C annealing temperature also provided the best intercolor balance at 28.98 ± 12.01%, compared with 27.00 ± 8.25% for 58.5°C and 22.00 ± 3.28% for 57.7°C. Between 56°C and 60°C, all allele peaks were called and the apparent allele lengths for all but one sample sized within one basepair when amplified at all annealing temperatures. One of the blood extracted DNA samples that was amplified at the 57.7°C annealing temperature contained a third peak at 87.48 bp in the PEZ21 size range. This was only observed in one of the 2 amplifications at 57.7°C for that sample, and not observed in any of the other annealing temperature amplifications, suggesting that the peak is a possible artifact of the kit. The average PHR for each annealing temperature was 82.88 ± 13.84% for 56°C, 84.33 ± 13.46% for 57.7°C, 85.91 ± 11.95% for 58.5°C, 86.76 ± 11.68% for 60°C, and 87.63 ± 8.47% for 61.9°C. Although 61.9°C contained the best peak height ratios for all primer pairs that successfully amplified DNA, the PEZ05 locus dropped out in one amplification and had RFU values below 100 for the other 5 amplifications (compared with RFUs of 500-4000 for all other loci). For all samples tested, the average normalized peak heights across all loci were 882.01 RFU for 56°C, 954.20 RFU for 57.7°C, 835.30 RFU for 58.5°C, 892.21 RFU for 60°C, and 514.37 RFU for 61.9°C. The manufacturer of the kit recommends an annealing temperature of 60°C, which is congruent with the results herein.

Amplification – number of cycles

To determine the optimal number of cycles for the canine STR kit, PCRs (1 ng template in a 20 µL reaction) were carried out at 26, 28, 30, 32, and 34 cycles on a GeneAmp 9700 thermocycler (Applied Biosystems). DNA from 2 blood samples and one buccal sample was ampli-

fied in duplicate. The manufacturer (Finnzymes Oy) recommended 30 cycles for amplifying canine DNA. As expected, the overall signal intensity for most peaks increased with cycle number (Figure 3).

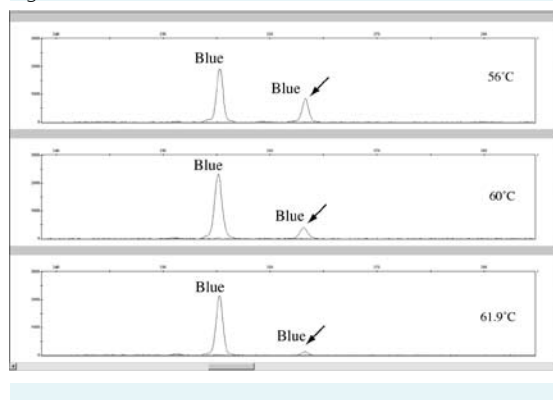
Four of the 6 amplifications did not yield product at locus FH2017 after only 26 PCR cycles, but all other cycles provided peaks above 50 RFU. The best overall intracolor and intercolor balances were seen at 28 and 30 cycles. At 34 cycles, certain loci had considerably higher signal than others. Peaks for the Zinc Finger, FH2017, FH2088, FH2010, and vWFX loci had the highest signals (5000-8000 RFU), while most other peaks had signals below 2000 RFU. There is no apparent reason why these specific markers showed higher intensity peaks at 34 cycles, although certain markers (Zinc Finger, vWFX, and FH2010) did have higher average RFU values in the PHR study of 61 animals, with 4271 RFU for Zinc Finger, 3362 RFU for vWFX, and 3687 RFU for FH2010. Most other markers had average peak heights between 1032 RFU (PEZ05) to 2874 RFU (FH2001). There also is not a clear correlation between the markers with higher intensity peaks at 34 cycles and the sensitivity study.

No shouldering from non-template nucleotide addition was seen at any of the cycle numbers tested, as expected because unlike Taq polymerase, Phusion Polymerase has 3' to 5' exonuclease activity that removes the non-templated nucleotide additions.

Peak height ratios (PHRs)

PHRs (excluding locus FH2017 and Zinc Finger) averaged between 78.57 ± 12.12 to 96.14 ± 4.12 . The FH2017 locus exhibited a standard deviation in PHR over 3 times that of the locus with the next highest standard deviation in PHR. Also, the low heterozygosity (0.49 for $n=61$) for locus FH2017 in relation to other markers in the panel may suggest a primer binding site mutation causing allelic imbalance. Across 13 breed categories of dogs, this locus exhibited a highly significant departure from Hardy-Weinberg expectations ($P < 0.01$) (18). Another indication of a primer binding site mutation is the increase of PHR for locus FH2017 with a decrease in annealing temperature (27). As an example, a sample from a Labrador Retriever used in the annealing temperature study exhibited this phenomenon (Figure 4). Eleven other samples also produced low PHRs ranging from 22.41% to 8.32%. One solution to address low peak signal caused by primer binding site mutations is to design a degenerate primer for locus FH2017 to add to the PCR (27,28). The primer binding regions

Figure 4.



A possible primer binding site variant at locus FH2017 illustrated by varying heterozygous peak signals and peak height ratios at different annealing temperatures. One nanogram of genomic canine DNA was amplified in duplicate at 56°C, 60°C, and 61.9°C. The lower relative fluorescence units heterozygous peak increases in signal as the annealing temperature decreases supporting a primer binding site mismatch (14).

for animals demonstrating FH2017 binding mutations should be sequenced to determine the location and base composition of the mutation. Because both intracolor and intercolor balance is poor at 56°C, reducing the stringency of the PCR by reducing the annealing temperature is not a feasible option.

The Zinc Finger locus had a lower PHR average (51.20 ± 7.29) with the male specific peak (~164 bp) amplifying 1000-2000 RFU less than the X specific peak (~159 bp). All samples that contained sex information about the donor animal were typed correctly by the Zinc Finger locus ($n=67$).

Stutter percentages

The stutter percentages were calculated by dividing the RFU value of the stutter peak ($n-4$) by the RFU value of the parent peak (n). Only parent peaks that were equal to or greater than 1000 RFUs were used for the study. Peaks with a heterozygous allele in the $n-4$ stutter position were not used to determine the mean stutter percentages for the locus. Table 4 is a summary of $n-4$ stutter percentages for all loci excluding Zinc Finger. Locus PEZ17 contained both $n-2$ and $n-4$ stutter peaks. The average $n-2$ stutter percentage for the PEZ17 locus was 12.10% with a standard deviation of 3.59% ($n=75$). There were stretches of CT repeat regions in the amplified product of the PEZ17 locus that may explain the presence of $n-2$ stutter. Locus FH2309 sometimes exhibits $n+4$ stutter peaks ($n=3$). The height of the FH2309

parent peak was typically above 2000 RFU when n+4 stutter was seen, with one exception of a parent peak at 900 RFU that contained an n+4 stutter artifact. In all cases, n-4 stutter was present in the 3 samples containing n+4 stutter peaks, and the n+4 stutter percentages for all animals ranged from 3.9% to 4.6%. For the heterozygous animals with n+4 stutter present, it is unlikely that the extra peak in the n+4 stutter position is the result of a tri-allelic pattern, but sequencing of these animals may better explain the nature of the 3-peak pattern.

TABLE 4. N-4 stutter percentages for 18 short tandem repeat markers (parent allele peak heights >1000 relative fluorescence units)

Locus	Mean ± standard deviation	N	Minimum stutter % observed	Maximum stutter % observed
PEZ02	6.43 ± 2.76	69	2.53	25.14
PEZ17	12.49 ± 3.92	78	7.26	30.87
FH2017	3.70 ± 4.60	45	1.23	19.86
FH2309	10.65 ± 3.00	72	4.81	21.72
PEZ05	3.15 ± 1.46	41	1.70	10.42
FH2001	0.73 ± 1.09	68	0.59	5.34
FH2328	2.85 ± 1.32	68	0.91	8.36
FH2004	6.24 ± 1.26	59	4.02	9.8
FH2361	5.71 ± 3.37	72	1.94	23.73
FH2054	3.04 ± 1.55	72	1.30	10.27
FH3377	3.60 ± 1.41	71	1.52	16.24
FH2107	5.03 ± 2.00	53	2.23	13.49
PEZ21	2.69 ± 1.05	52	1.39	5.66
FH2088	5.18 ± 1.13	57	3.01	8.94
vWfX	1.26 ± 1.01	52	0.49	4.02
FH2010	0.69 ± 0.84	61	0.67	2.79
PEZ16	0.57 ± 1.62	65	0.49	12.32
FH3313	9.30 ± 3.74	79	2.89	19.83

Artifacts (dye blobs)

Several artifacts were seen in the reagent blanks and in samples containing peaks with low RFU values. An artifact at ~123 bp is seen in the blue dye channel, at ~117 bp in the green channel, at ~96 bp in the yellow channel, and at ~112 bp and ~118 bp in the red channel. The artifact in the yellow channel sometimes interferes with the allele calls in marker PEZ21 if the peak height for the allele is at or below 100 RFU.

Mixture study

A binning program called Flexibin (25) was used to determine allele calls for the animals of the mixture study (18). The "Repeats" column is not an accurate reflection of how many tetranucleotide repeats are present in the actual PCR product but the values are operationally defined. Based on this binning method the operational allele lengths were determined for the mixture ratio (Table 5).

Two genomic canine samples were mixed in the following proportions so that the total DNA input for each reaction was 1.0 ng: 20:1, 10:1, 5:1, 2:1, 1:1, 1:2, 1:5, 1:10, and 1:20. The loci in Table 5 were selected because the alleles from both mixed samples were non-overlapping. The minor component robustly amplifies (all peaks above 50 RFU) at a 1:5 mixture ratio. This may be partly due to lower intracolor balances across all dye channels. The Identifiler® kit typically has an intracolor balance of 50% for all color channels when samples are not degraded or inhibited, which may explain the typing kit's ability to robustly amplify the minor component in mixture ratios of 1:10 (15). Better detection of minor component DNA in canine mix-

TABLE 5. Mixture ratio study comparison of markers that did not have any allele overlaps*

Mixture ratio	Locus								
	PEZ17	FH2309	PEZ05	FH2328	FH2054	FH3377	FH2107	PEZ16	FH3313
20:1	208	-, -	104, -	-	-, 164	239	-	292	-, 430
20:1	208	-, -	-, -	198	160, 164	239	362	292	-, -
10:1	208	-, -	-, 108	198	160, 164	239	362	292	426, 430
10:1	208	-, 390	104, -	198	-, 164	239	362	292	-, -
5:1	208	386, 390	104, 108	198	160, 164	239	362	292	426, 430
2:1	208	386, 390	104, 108	198	160, 164	239	362	292	426, 430
1:2	199, 225	367, 397	100	190, 194	152, 172	301	377, 381	300, 312	395, 422
1:5	199, 225	367, 397	100	190, 194	152, 172	301	377, 381	300, 312	395, 422
1:10	199, 225	367, -	100	190, 194	152, 172	301	377, 381	300, 312	-, 422
1:10	199, 225	367, 397	100	-, 194	152, 172	301	-, -	300, 312	395, 422
1:20	199, 225	-, -	100	190, 194	152, 172	301	-, -	300, 312	-, -
1:20	199, 225	-, -	100	-, 194	-, -	301	-, -	300, 312	395, -

*The alleles for the minor component are listed. A minimum height cut off of 50 relative fluorescence units is used.

tures might be attained if the intracolor balance were improved.

Overlapping marker size ranges

Size range overlap among loci was detected in the electropherograms of 67 different California dog samples. A marker range overlap of 14 base pairs occurred between the FH3377 and FH2107 loci in the Akita sample from the sensitivity study. The overlap was only detected because marker FH3377 appeared to contain no peaks, while marker FH2107 had 3 peaks. Also, the 3 peaks in the FH2107 marker range were composed of 2 peaks that are of significantly less intensity than the third peak which was closest to the FH3377 size range determined before the mixture study. Of 67 animals, the Akita sample was the only example of an allele size in the FH3377/FH2107 overlapping region. The analyst should be cautious when making an interpretation of an allele in the overlapping range for loci FH3377 and FH2107 because both markers contain an allele ~295.6 bp. Otherwise, the FH3377 locus is a pentanucleotide repeat and locus FH2107 is a tetranucleotide repeat; all other alleles that type in the overlap region should be more easily distinguished with the binning properties of the different length repeat motifs (unless a partial repeat occurs, which is not common). Another overlapping allele range is between markers FH2004/FH2361 that share a 3-4 base pair overlap, but perfect repeat allele bins in both marker ranges are offset by one base pair from each other.

Microvariants seen in FH3313 and FH2361

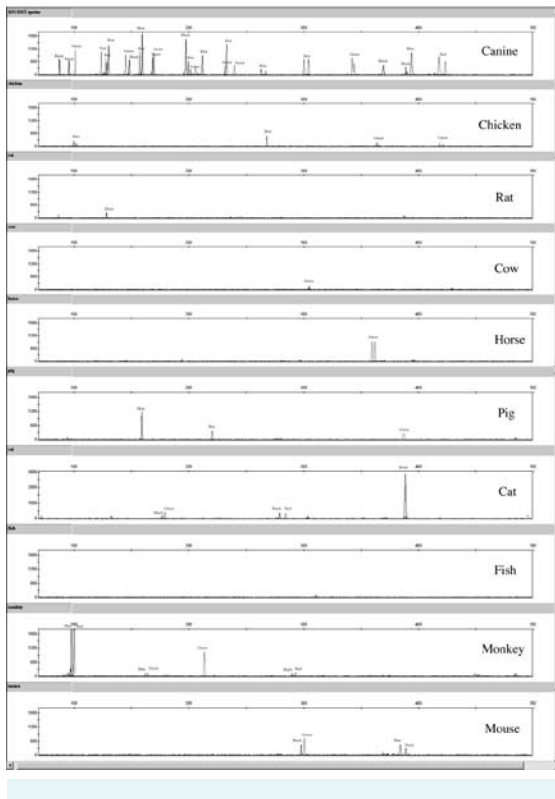
Three animals typed in this study exhibited a microvariant allele for locus FH2361. One animal had heterozygous alleles 1 base pair apart and 2 other animals had heterozygous alleles 2 base pairs apart. The peaks that were a single base pair apart were resolved clearly. Five animals contained microvariant alleles in the FH3313 locus. Three animals contained well resolved FH3313 microvariants which appeared to be 2 base pairs apart, but 2 other canine samples had broad peaks. The samples containing broad peaks were confirmed by a second amplification and 2 separate injections (Figure 5). The morphology of these peaks might be caused by STR alleles of the same length on either autosomal chromosome containing slightly difference sequence variations from each other, causing different migration rates during electrophoresis (25,29). Microvariants were also observed at loci FH2107, FH2309, and FH3377 when an independent larger sample set of 667 dogs was analyzed (26).

Figure 5.



Five different canine samples exhibiting microvariant alleles for marker FH3313. Two of the animals (the bottom frames) may have alleles that are the same length but are composed of slightly different sequences, causing slight differences in electrophoretic mobility and resulting in a broad peak.

Figure 6.



Representative electropherograms for canine, chicken, rat, cow, horse, pig, cat, fish, rhesus monkey, and mouse samples amplified with the Finnzymes Canine 2.1 Panel β Lot reagent kit.

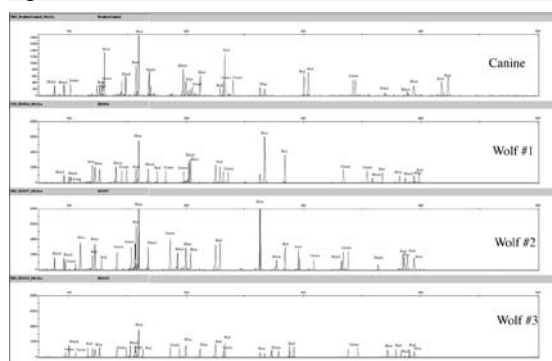
Species specificity

Several peaks were observed when DNA samples from chicken, mouse, rat, horse, cow, pig, cat, fish, and monkey were typed (Figure 6). Only low level peaks below 100 RFU were present for both amplifications with rat and fish DNA. Peaks below 100 RFU were seen in one of 2 amplifications of pig DNA. Most peaks fell outside of the canine allele size ranges. Using the Flexibin program and offsetting the allele sizes using the positive control sample, a few peak sizes were close to canine allele positions. To better assess whether these peaks truly fall into a canine allele bin, the various non-canine species (fish, pig, rat, cat, mouse, chicken, and monkey) should be run again with an allelic ladder, once it is developed, and it may also be worthwhile to perform mixture studies between canine and non-canine samples. Subsequent BLAT and BLAST searches of the 19 sets of primers revealed that some primers had 95%-100% identity with the genomic information for certain species (Table 6). No 2 primers (forward and reverse) for a locus had such high homology. Primers for the FH2361R, vWFX, and FH2017F loci have the most homology with a number of non-canine species. The PEZ17R primer has 100% identity with 100 locations in the cat genome. But, the primer binding sites discovered with the BLAT and BLAST searches for the 19-plex primers do not assess the possibility for non-target binding of the 3' end of the primers. For instance, primer binding to chicken, cow, and pig samples was not confirmed by searching for similar DNA sequences between the Finnzymes 19-plex primer panel and the genome information available on the BLAST and BLAT Web sites. Similarly, several canine primers shared 95%-100% homology with multiple locations in the human genome, but amplification peaks were not present in duplicate 19-

plex reactions of human DNA (data not shown); so the ability to assess cross-reactivity of a primer to a particular genome is not always possible with a BLAST or BLAT search. Also, none of the peaks exhibited in any of the non-canine species appeared to exhibit the morphology of STR products (all peaks above 500 RFU lack stutter peaks), which could be helpful in evaluating if non-canine species amplification has occurred in a DNA mixture. Of course, species cross-reactivity does not make a kit unreliable. It is important to perform such studies to better understand the limitations of an assay. No full 19 locus profile was observed with any of the above mentioned species.

The Finnzymes canine STR kit amplified reproducible STR profiles in 5 wolf (*Canis lupus*) samples (Figure 7), as would be expected given the evolutionary history of the domes-

Figure 7.



Canine F-863 control sample and 3 different wolf sample amplification with the Finnzymes 19-plex. The following amplifications used 1 ng template DNA.

TABLE 6. Primers that show 95%-100% identity with non-canine species based on BLAT and BLAST searches*

Primer	Species chromosome map location				
	cat	mouse	horse	human	rhesus
FH2309R†	–	8, 13	–	–	–
FH2361R	151419	–	14+	8-	11+, 19+, 20-
FH2004R	199169	–	–	–	–
vWFX	–	4-, 19+	2+, 3-	4+, 22-	11-
PEZ17R	[100 BLAST hits]	–	–	–	–
FH2107R	–	–	–	10+	–
FH2017F	135908	–	Un+/-, 1+, 16+/-, 17-, 29	2+,3+,5+,13-,X-	1+, 7-, 13+
ZFR	202367	–	X-	–	–

*All primer sequences utilized in the Finnzymes canine 19-plex were searched in both the BLAT (<http://genome.ucsc.edu/>) database and the BLAST (<http://genome.ucsc.edu/>) database to assess whether certain primers anneal to non-canine species. The chromosome map locations for each species that showed sequence homology with particular primers are indicated in the table above.

†Sequence homology only confirmed by the BLAST search tool.

TABLE 7. Peak sizes and relative fluorescence units (RFU) values for peaks at the canine 19-plex loci for 13 single hair extraction samples (samples 1-13) and 4 saliva stain extractions (samples 14-17). Next to the marker name, alleles are described in raw base lengths and are followed by a backslash (/) and the height of the peak is listed in RFUs

Sample	Locus	Peak 1 (bp)/RFU	Peak 2 (bp)/RFU	Peak 3 (bp)/RFU
1	Zinc Finger	159.31/70	163.77/55	
2	-	-	-	
3	-	-	-	
	PEZ02	125.86/28	-	
	Zinc Finger	159.21/67	-	
4	FH2328	190.35/55	-	
	FH2004	239.67/27	-	
	FH2361	350.13/28	-	
5	-	-	-	
	Zinc Finger	159.21/67	-	
6	PEZ17	207.63/50	-	
	FH3377	192.70/30	-	
	PEZ02	125.83/46	-	
	Zinc Finger	159.22/67	163.67/32	
7	PEZ05	101.09/59	-	
	FH2001	148.92/79	-	
	FH2361	354.23/77	-	
	PEZ02	125.76/54	129.87/50	
	Zinc Finger	159.29/173	163.75/50	
	PEZ17*	199.39/45	203.57/49	215.79/47
	FH2309	355.04/92	-	
	PEZ05	101.07/234	-	
	FH2001	140.71/155	148.92/132	
	FH2328	182.46/87	193.96/60	
	FH2004	239.60/73	-	
8	FH2361	346.16/57	350.24/92	
	PEZ21	87.39/102	-	
	FH2054	152.20/44	164.34/70	
	FH3377	197.38/46	246.90/40	
	FH2107	380.37/54	-	
	FH2088	131.89/224	-	
	vWF.X	157.09/115	-	
	FH2010	236.78/120	-	
	PEZ16	304.25/112	-	
	FH3313	402.70/53	414.37/55	
9	PEZ05	101.12/50	-	
	Zinc Finger	159.14/47	163.72/36	
	PEZ05	105.11/35	-	
	FH2001	128.34/30	144.61/34	
10	FH2328	197.91/38	-	
	FH2361	354.15/31	-	
	PEZ21	87.30/27	-	
	FH2054	168.43/37	172.34/31	

TABLE 7. Peak sizes and relative fluorescence units (RFU) values for peaks at the canine 19-plex loci for 13 single hair extraction samples (samples 1-13) and 4 saliva stain extractions (samples 14-17). Next to the marker name, alleles are described in raw base lengths and are followed by a backslash (/) and the height of the peak is listed in RFUs – continued

Sample	Locus	Peak 1 (bp)/RFU	Peak 2 (bp)/RFU	Peak 3 (bp)/RFU
	FH3377	246.56/46	-	
	FH2088	123.49/50	131.75/35	
	FH3313	383.98/51	-	
	Zinc Finger	159.15/25	-	
	PEZ05	101.01/53	-	
	FH2001	144.80/56	148.87/26	
11	FH2328	182.47/30	-	
	FH2361	345.97/25	-	
	PEZ21	87.41/49	-	
	FH2054	152.09/39	-	
	FH2088	131.75/49	-	
	Zinc Finger	159.16/38	-	
	FH2001	148.90/37	-	
	FH2004	239.43/53	-	
12	FH2361	346.12/31	-	
	FH2054	164.36/43	-	
	FH3377	197.41/53	-	
	FH2010	236.69/46	-	
	PEZ02	121.74/40	-	
	Zinc Finger	159.08/62	-	
	PEZ17	203.55/68	207.53/35	215.73/35
	FH2017	266.77/86	-	
	FH2001	140.70/31	-	
13	FH2361	350.06/30	-	
	PEZ21	87.37/52	-	
	FH2054	152.12/28	164.32/64	
	FH2010	236.70/42	-	
	PEZ16	288.36/27	299.70/43	
	FH3313	402.71/48	-	
	PEZ02	130.00/45	-	
	Zinc Finger	159.23/53	163.70/73	
	PEZ17	203.50/137	215.73/64	
	PEZ05	101.18/137	-	
	FH2001	140.71/101	148.86/131	
14	FH2328	182.50/69	193.87/93	
	FH2004	239.49/79	-	
	PEZ21	83.39/83	87.42/49	
	FH2054	152.21/63	164.29/34	
	FH3377	187.92/44	197.44/44	
	FH2088	131.81/97	-	
15	-	-	-	
16	-	-	-	
17	PEZ02	129.91/64	-	

TABLE 7. Peak sizes and relative fluorescence units (RFU) values for peaks at the canine 19-plex loci for 13 single hair extraction samples (samples 1-13) and 4 saliva stain extractions (samples 14-17). Next to the marker name, alleles are described in raw base lengths and are followed by a backslash (/) and the height of the peak is listed in RFUs – continued

Sample	Locus	Peak 1 (bp)/RFU	Peak 2 (bp)/RFU	Peak 3 (bp)/RFU
	Zinc Finger	159.23/113	163.77/62	
	PEZ17	215.57/92	-	
	PEZ05	101.08/126	-	
	FH2001	140.75/73	-	
	FH2328	193.94/29	-	
	FH2004	239.62/68	-	
17	PEZ21	83.40/64	87.49/54	
	FH2054	152.15/187	-	
	FH3377	197.44/95	-	
	FH2088	131.93/199	-	
	vWF.X	157.10/94	-	
	FH2010	229.02/55	-	
	PEZ16	288.36/54	-	

*PEZ17 had 3 allele calls in 2 different samples. The low RFU values make it difficult to determine which allele calls are true peaks.

tic dog and wolf (2). The peaks obtained in wolves were concordant with the allele spectrum of the domestic dog for each locus. Indeed, STR loci are well-known to cross-amplify even among more genetically differentiated species than the dog and wolf (30). For example, the Identifier® kit developed for human profiling generates partial profiles for other primate species (chimpanzee, orangutan, macaque, and gorilla) (15) and the “Meowplex” generates PCR products for ocelot and puma samples (17).

Case-type samples. Samples from a recent dog attack were amplified in duplicate using the Finnzymes 19-plex. Sample numbers 1 through 13 were single shed hair samples found on a sweat shirt and samples 14 through 17 were saliva stain samples found on the victim’s sweatshirt. All samples were quantified at levels below 0.1 ng/μL, the limit of detection for the Fluoroskan Ascent (Thermo Scientific) when using the PicoGreen® (Invitrogen) quant method. Two microliters of sample was amplified for each reaction. Although not recommended for routine case-work, the detection threshold was lowered to 25 RFU for all colors to collect as much data as possible (Table 7). Using the California Department of Justice casework peak threshold standards, only peaks called above 50 RFU were considered “alleles,” and single peaks at a locus above 200

RFU were considered a homozygous genotype for the random match probability calculations. A 10-second injection time at 3 kV was used while running the sample on the ABI 3130CE. Table 7 presents a combined summary of the peaks called from duplicate PCRs; all alleles are described as “raw” lengths in bases. Many of the case work samples had peaks below 50 RFU and many markers only had a single peak called, possibly indicating peak drop-out due to low input DNA amounts amplified in the PCRs. The detection threshold was initially lowered to 25 RFU to determine if low input DNA peaks were present, but interpreting these peaks may be difficult due to stochastic effects and the possibility of allele drop-in due to low copy number DNA being amplified (27,28,31-33).

Samples 8 and 13 had 3 peaks in the marker range for PEZ17, some of which were below 50 RFU. Sample 8, a single shed hair extracted using the Chelex method, provided the best profile overall with the most peaks called above 50 RFU. Hair samples tend to have low levels of nuclear DNA; additional technical modifications will be required to amplify samples below 0.125 ng, such as increasing the cycle number and investigating alternative extraction protocols for telogen hair samples (31-33). Based on the positive control sample run on the same injection plate as the non-probative samples, sample 8 allele calls were obtained by binning (for peaks above 50 RFU, Table 8).

Assuming independent segregation of loci and using a conservative F_{st} value of 0.09 [The F_{st} and the allele frequencies were determined from a 667 animal population study from canines across the United States (18), the random match probability for the sample 8 profile was calculated at 1.00×10^{-17} . Two saliva stain samples (14 and 17), provided partial profiles containing peaks from 10 and 12 loci, respectively. Using the F_{st} value of 0.09, the random match probability for the sample 14 profile was and 3.18×10^{-3} . Sample 17 had a match probability of 1.04×10^{-3} .

As seen in Table 7, the sex typing marker (Zinc Finger) had varying success when typing telogen hair samples. Of the 10 hair samples that had amplified peaks, 3 samples typed X and Y associated peaks, but one of those samples did not type the Y peak above 50 RFU. Four hair samples only had the X peak type above 50 RFU and 3 samples had X peaks below 50 RFU but above 25 RFU. As seen with the PHR study, the ~159 base X peak types at much higher RFUs (typically 2 times higher) than the ~164 base Y peak in male canines and therefore seems to amplify more robustly at input DNA levels below 125 pg. This was

TABLE 8. Allele calls and frequencies for the Finnzymes 19-plex kit's positive control animal (F-863) and for non-probative sample 8. "Raw" and "binned" values are in bases

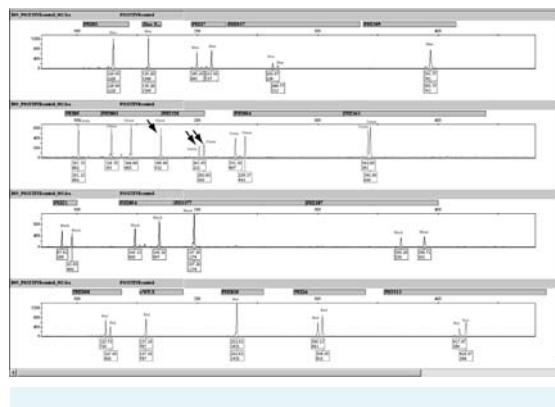
Locus	Positive control		Non-probative sample		National frequencies
	raw	binned	raw	binned	
FH2001	128.44	127.01	-	-	0.2339
	-	-	140.71	139.37	0.2864
	144.76	143.49	-	-	0.2211
	-	-	148.92	147.61	0.0652
FH2004	231.51	232.82	-	-	0.1319
	239.63	241.22	239.6	241.22	0.2789
FH2010	232.79	234.26	-	-	0.2459
	-	-	236.78	238.46	0.2519
FH2017	262.64	264.29	-	-	0.6214
	266.76	268.09	-	-	0.2406
FH2054	148.08	147.41	-	-	0.2136
	-	-	164.34	164.05	0.1432
FH2088	168.26	168.21	-	-	0.0915
	123.65	122.28	-	-	0.072
FH2107	127.58	126.24	-	-	0.2939
	-	-	131.89	130.2	0.2489
FH2107	369.77	369.84	-	-	0.1522
	-	-	380.37	381	0.1439
	389.23	388.44	-	-	0.0405
FH2309	-	-	355.04	355.02	0.0322
	393.9	393.42	-	-	0.1289
FH2328	169.79	171	-	-	0.0007
	-	-	182.46	182.52	0.1724
	-	-	193.96	194.04	0.1514
	205.47	205.56	-	-	0.0285
FH2361	342.1	342.7	-	-	0.2324
	343.99	342.7	-	-	0.2324
	-	-	346.16	346.7	0.2571
	-	-	350.24	350.7	0.1724
FH3313	-	-	402.7	403.01	0.0457
	-	-	414.37	414.65	0.1004
	418.21	418.53	-	-	0.1259
	423.84	422.41	-	-	0.0525
FH3377	197.4	197.11	-	-	0.1874
vWF.X	157.06	157.04	157.09	157.04	0.4843
PEZ02	-	-	125.76	124.36	0.3238
	129.8	128.36	129.87	128.36	0.2646
PEZ05	101.04	100.4	101.07	100.4	0.4528
PEZ16	300.32	300.3	-	-	0.3171
	304.19	304.22	304.25	304.22	0.1087
PEZ17	199.47	199.38	-	-	0.1844
	211.56	211.98	-	-	0.2211
PEZ21	87.44	87.06	87.39	87.06	0.3126
	95.68	95.14	-	-	0.4423

also observed for one of the samples used in the sensitivity study. When 1 ng of template DNA was amplified, the average Zinc Finger PHR for both amplifications was $53.42 \pm 2.69\%$. The same sample produced a Zinc Finger PHR of $35.64 \pm 12.05\%$ when 62.5 pg of DNA was amplified. Furthermore, with Zinc Finger peak heights ranging from 247 RFU to 67 RFU for samples amplified with less than 125 pg, it is possible that the Y peak could fall below the 50 RFU detection threshold. For the 2 saliva stain samples that generated profiles, both samples amplified both the X and Y peaks above 50 RFU.

Positive control animal

It was noted during the validation that positive control sample (F-863) provided with the canine multiplex had a concentration of 0.1 ng/ μ L instead of the 1 ng/ μ L reported on the tube. Two different tubes of F-863 were quantified with PicoGreen[®] in triplicate (for a total of 6 separate quantifications of the positive control sample.) All 6 quantifications yielded a 0.1 ng/ μ L concentration. Interestingly, the positive control F-863 produces a profile that contains 3 peaks for marker FH2328 (Figure 8). To confirm that the third allele is attributed to locus FH2328 and not a third allele due to allele overlap with neighboring loci PEZ05 and FH2001, singleplex reactions were carried out. In triplicate reactions, 4.5 μ L of each of the forward and reverse primers (500 nm total) were combined with 9 μ L of the Finnzymes Reaction Mixture and 2 μ L of each singleplex. Based on these results, the positive control animal contained a single peak at ~ 100.4 bases for locus PEZ05, the FH2001 locus contained 2 peaks at ~ 128.01 bases and ~ 144.8 bases, and the FH2328 locus contained 3 peaks at ~ 169.9 bas-

Figure 8.



Electropherogram for the positive control animal F-863. Three peaks (indicated by the arrows) are present in marker FH2328.

es, ~201.6 bases, and ~205.5 bases. In 2 of the 3 amplifications, the FH2328 locus had a third peak at ~201.6 bp. All of the amplifications of the positive control animal were performed using only 0.2 ng of DNA in a 20- μ L reaction. Possible stochastic fluctuations from amplifying less than the ideal template amount of 1 ng may explain why the third peak was only seen in 2 of the 3 single plex amplifications using the FH2328 primers. All of the amplifications of the positive control animal using the Finnzymes 19-plex throughout the validation ($n = 13$) contained 3 peaks. This third FH2328 peak could possibly be an artifact peak generated from the 19-plex or it could be a tri-allelic pattern inherent to the cell line used for the positive control, with the latter explanation being more likely.

The Zinc Finger sex typing marker successfully typed the positive control animal as female in all ten amplifications. Typically, the ~159 bp X peak was the highest peak in the blue color channel for the positive control.

CONCLUSIONS

DNA extracted from blood and buccal cells can reliably produce full profiles when amplified with the canine STR kit. The multiplex is able to type 0.125 ng of DNA template in a 20 μ L reaction volume and obtain full profiles. Furthermore, the Finnzymes multiplex has a high level of sensitivity, ability to detect low level mixture ratios, and power of discrimination while amplifying case-type samples comparable with those of other forensic typing systems.

The overall PHRs for most markers (excluding the Zinc Finger and FH2017) ranged between 78.57 and 96.14%. Intracolor balance was not as robust as in Identifier® (between 32%-45% depending on color channel); Intercolor balance was 22%, averaged over 61 samples evaluated. These lower values are due to 1 or 2 loci per color yielding too low or too much product. To obtain better intracolor balance, increasing the signal of markers that tended to produce the lowest RFU values per color (FH2017, PEZ05, FH2107, and FH2088) and reducing the signal in markers that tended to generate the highest RFU values (FH2010, Zinc Finger, and FH2054) could be investigated.

The proposed reagent kit is also capable of genotyping wolf (*C. lupus*) samples, which can be useful in wild-life population and conservation studies. Certain primers had high sequence homology with non-canine species, but this should not be problematic. The FH2017 forward primer has 95%-100% homology to numerous locations in

the horse, human, and rhesus monkey genomes. Also, it is probable that the primers designed for locus FH2017 bind to a region of DNA that contains primer binding site variants, which is apparent in the large standard deviation for the PHRs calculated. A degenerate primer could be added to the reaction to improve performance. The positive control animal sample provided in the kit is at a lower concentration (0.1 ng/ μ L) than what is reported on the tube (1 ng/ μ L). Three peaks were seen in locus FH2328 for the positive control. This peak could be a tri-allelic pattern associated with the cell line used for positive control F-863. Only the positive control sample exhibited this 3-peak pattern at locus FH2328.

In conclusion, the Finnzymes 19-plex holds excellent promise as a tool for routine forensic casework involving canine DNA. The data presented in this validation study should prove highly useful for the forensic science community as well as for the manufacturer of the kit in order to make final modifications so that the kit will meet the needs and requirements of the DNA forensic science community.

Acknowledgments

Martin Buoncristiani, William Hudlow, and Mavis Date-Chong of the Method Development group and Ken Konzak and Eva Steinberger from the California Department of Justice, Richmond DNA Laboratory were instrumental in allowing us to complete portions of this research at their facilities. We would also like to thank IDEXX Labs for providing canine blood samples. We would like to thank Dr John Butler (NIST) for his ideas and co-development of the Finnzymes multiplex which made this project possible. We also thank anonymous reviewers for their comments on improving this article. This research was partially funded by the National Institute of Justice (NIJ Grant No. 2004-DN-BX-K007). Parts of the research described in this article have been reported in Melody Dayton's Masters of Science (Forensic Science) thesis which has been accepted submitted to the UC Davis Office of Graduate Studies. Dayton has also presented this work at the 2008 UC Davis Graduate Group in Forensic Science Student Seminar as part of the program's requirement. This research has been presented at the 2008 NIJ Grantees Conference and at the 2008 International Society for Animal Genetics (ISAG) Conference in Amsterdam, the Netherlands.

References

- 1 Savolainen P, Zhang YP, Luo J, Lundeberg J, Leitner T. Genetic evidence for an East Asian origin of domestic dogs. *Science*.

- 2002;298:1610-3. [Medline:12446907](#) [doi:10.1126/science.1073906](#)
- 2 Brewer DJ, Terence C, Phillips A. Dogs in antiquity. Anubis to Cerberus: The Origins of the domestic dog. Warminster (UK): Aris & Phillips; 2002.
 - 3 U.S. pet ownership & demographics sourcebook. Schaumburg (IL): Center for Information Management, American Veterinary Medical Association, 1997.
 - 4 Halverson J, Basten CA. PCR multiplex and database for forensic DNA identification of dogs. *J Forensic Sci.* 2005;50:352-63. [Medline:15813546](#) [doi:10.1520/JFS2004207](#)
 - 5 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.* 1999;44:623-6. [Medline:10408119](#)
 - 6 Eichmann C, Berger B, Reinhold M, Lutz M, Parson W. Canine-specific STR typing of saliva traces on dog bite wounds. *Int J Legal Med.* 2004;118:337-42. [Medline:15480731](#) [doi:10.1007/s00414-004-0479-7](#)
 - 7 Francisco LV, Langston AA, Mellersh CS, Neal CL, Ostrander EA. A class of highly polymorphic tetranucleotide repeats for canine genetic mapping. *Mamm Genome.* 1996;7:359-62. [Medline:8661717](#) [doi:10.1007/s003359900104](#)
 - 8 Hellmann AP, Rohleder U, Eichmann C, Pfeiffer I, Parson W, Schleenbecker U. A proposal for standardization in forensic canine DNA typing: allele nomenclature of six canine-specific STR loci. *J Forensic Sci.* 2006;51:274-81. [Medline:16566760](#) [doi:10.1111/j.1556-4029.2006.00049.x](#)
 - 9 Halverson J, Dvorak J, Stevenson T. inventors. Metamorphix, Inc, assignee. Microsatellite sequences for canine genotyping. US patent 5,874,217. February 23, 1999.
 - 10 Neff MW, Broman KW, Mellersh CS, Ray K, Acland GM, Aguirre GD, et al. A second-generation genetic linkage map of the domestic dog, *Canis familiaris*. *Genetics.* 1999;151:803-20. [Medline:9927471](#)
 - 11 Shibuya H, Collins BK, Huang TH, Johnson GS. A polymorphic (AGGAAT)_n tandem repeat in an intron of the canine von Willebrand factor gene. *Anim Genet.* 1994;25:122. [Medline:8010530](#)
 - 12 Mellersh CS, Langston AA, Acland GM, Fleming MA, Ray K, Wiegand NA, et al. A linkage map of the canine genome. *Genomics.* 1997;46:326-36. [Medline:9441735](#) [doi:10.1006/geno.1997.5098](#)
 - 13 Eichmann C, Berger B, Steinlechner M, Parson W. Estimating the probability of identity in a random dog population using 15 highly polymorphic canine STR markers. *Forensic Sci Int.* 2005;151:37-44. [Medline:15935941](#) [doi:10.1016/j.forsciint.2004.07.002](#)
 - 14 Koskinen MT, Bredbacka P. A convenient and efficient microsatellite-based assay for resolving parentages in dogs. *Anim Genet.* 1999;30:148-9. [Medline:10376306](#) [doi:10.1046/j.1365-2052.1999.00438.x](#)
 - 15 Collins PJ, Hennessy LK, Leibelt CS, Roby RK, Reeder DJ, Foxall PA. Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433, and amelogenin: the AmpFISTR Identifier PCR Amplification Kit. *J Forensic Sci.* 2004;49:1265-77. [Medline:15568700](#) [doi:10.1520/JFS2002195](#)
 - 16 Meiboom M, Murua Escobar H, Winkler S, Nolte I, Bullerdiek J. Molecular characterization and mapping of the canine KRAB zinc finger gene ZNF331. *Anim Genet.* 2004;35:262-3. [Medline:15147412](#) [doi:10.1111/j.1365-2052.2004.01146.x](#)
 - 17 Menotti-Raymond MA, David VA, Wachter LL, Butler JM, O'Brien SJ. An STR forensic typing system for genetic individualization of domestic cat (*Felis catus*) samples. *J Forensic Sci.* 2005;50:1061-70. [Medline:16225210](#) [doi:10.1520/JFS2004317](#)
 - 18 Kanthaswamy S, Tom BK, Mattila AM, Johnston E, Dayton M, Halverson J, et al. A Canine STR Kit and Population Database for Use in Forensic Casework. *J Forensic Sci.* Forthcoming 2009.
 - 19 Hudlow WR, Date Chong M, Swango KL, Timken MD, Buoncristiani MR. A quadruplex real-time qPCR assay for the simultaneous assessment of total human DNA, human male DNA, DNA degradation and the presence of PCR inhibitors in forensic samples: A diagnostic tool for STR typing. *Forensic Sci Int Genet.* 2008;2:108-25. [Medline:19083806](#) [doi:10.1016/j.fsigen.2007.09.001](#)
 - 20 Wang Y, Prosen DE, Mei L, Sullivan JC, Finney M, Vander Horn PB. A novel strategy to engineer DNA polymerases for enhanced processivity and improved performance in vitro. *Nucleic Acids Res.* 2004;32:1197-207. [Medline:14973201](#) [doi:10.1093/nar/gkh271](#)
 - 21 Phusion Hot Start High-Fidelity DNA Polymerase, New England Biolabs, Finnzymes, Version 1.2 September 2006. Available from: http://www.finnzymes.com/pdf/phusion_hs_datasheet_f540sl_1_5_low.pdf. Accessed: May 19, 2009.
 - 22 Nord K, Gunneriusson E, Ringdahl J, Stahl S, Uhlen M, Nygren PA. Binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain. Binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain. *Nat Biotechnol.* 1997;15:772-7. [Medline:9255793](#) [doi:10.1038/nbt0897-772](#)
 - 23 Wikman M, Steffen AC, Gunneriusson E, Tolmachev V, Adams GP, Carlsson J, et al. Selection and characterization of HER2/neu-binding affibody ligands. *Protein Eng Des Sel.* 2004;17:455-62. [Medline:15208403](#) [doi:10.1093/protein/gzh053](#)
 - 24 Klein SB, Wallin JM, Buoncristiani MR. Addressing ambient temperature variation effects on sizing precision of AmpFISTR® Profiler Plus™ alleles detected on the ABI Prism® 310 Genetic Analyzer. *Forensic Science Communication.* 2003;5.
 - 25 Amos W, Hoffman JI, Frodsham A, Zhang L, Best S, Hill AVS. Automated binning of microsatellite alleles: problems and solutions. *Mol Ecol Notes.* 2007;7:10-4. [doi:10.1111/j.1471-8286.2006.01560.x](#)
 - 26 Tom BK, Koskinen MT, Dayton MR, Mattila AM, Johnston E, Fantin D, et al. Development of a nomenclature system for a canine STR multiplex reagent kit. *Forthcoming J Forensic Sci.* 2009.
 - 27 Butler J. Forensic DNA typing: biology & technology behind STR

- markers. San Diego (CA): Academic Press; 2005.
- 28 Leibelt C, Budowle B, Collins P, Daoudi Y, Moretti T, Nunn G, et al. Identification of a D8S1179 primer binding site mutation and the validation of a primer designed to recover null alleles. *Forensic Sci Int.* 2003;133:220-7. [Medline:12787655](#) [doi:10.1016/S0379-0738\(03\)00035-5](#)
- 29 Rosenblum BB, Oaks F, Menchen S, Johnson B. Improved single-strand DNA sizing accuracy in capillary electrophoresis. *Nucleic Acids Res.* 1997;25:3925-9. [Medline:9380518](#) [doi:10.1093/nar/25.19.3925](#)
- 30 Primmer CR, Moller AP, Ellegren H. A wide-range survey of cross-species microsatellite amplification in birds. *Mol Ecol.* 1997;6:101. [Medline:9004524](#)
- 31 Brauner P, Reshef A, Gorski A. DNA profiling of trace evidence – mitigating evidence in a dog biting case. *J Forensic Sci.* 2001;46:1232-4. [Medline:11569571](#)
- 32 Lu DJ, Sun HY, Chen LX. Genotype reliability of short tandem repeats typing from minute DNA [in Chinese]. *Fa Yi Xue Za Zhi.* 2003;19:151-3. [Medline:13677317](#)
- 33 Budowle B, Hobson DL, Smerick JB, Smith JAL. Low copy number – consideration and caution. In: Twelfth International Symposium on Human Identification 2001, Promega Corporation, Madison, Wisconsin, 2001. Available from: <http://www.promega.com/geneticidproc/ussymp12proc/contents/budowle.pdf>. Accessed: May 18, 2009.