TECHNICAL NOTE

Marcia LaFountain,[1] *M.S.; Margaret Schwartz,*[1] *Ph.D.; Julie Cormier,*[1] *and Eric Buel,*[1] *Ph.D.*

Validation of Capillary Electrophoresis for Analysis of the X-Y Homologous Amelogenin Gene*

REFERENCE: LaFountain M, Schwartz M, Cormier J, Buel E. Validation of capillary electrophoresis for analysis of the X-Y homologous amelogenin gene. J Forensic Sci 1998;43(6): 1188–1194.

ABSTRACT: Capillary electrophoresis (CE) is a versatile technology with tremendous potential applications in forensic science. A refinement of the traditional slab-gel, separation occurs in a liquid polymer contained within a capillary. The capillary electrophoresis unit is computer-driven, semi-automating the loading and analysis of samples.

As with any new technology, a full performance assessment is required before it may be implemented with confidence. This paper focuses on validation of CE technology for analysis of a portion of the X-Y homologous amelogenin gene used in gender determination. Typical Y/X ratios of peak heights for male samples were determined. Instrument linearity, sample resolution and reproducibility were examined. Samples subjected to contamination, extreme environmental conditions or extracted from a variety of substrates were also tested by CE. All samples typed correctly. Genetic material from a number of common non-primate animals was amplified with amelogenin primers. Some species yielded no product. Products derived from the animal samples that did amplify produced peaks on CE analysis readily distinguishable from those of human origin.

KEYWORDS: forensic science, DNA typing, amelogenin, capillary electrophoresis, validation

Capillary electrophoresis is a separation technology that uses the basic principles of traditional gel-based methods (1,2). The separation occurs within a polymer-filled capillary monitored by a laser. Fluorescently tagged PCR products are detected as they migrate past a window within the capillary. Features such as semi-automated sample loading and analysis offer some key advantages over the current gel-based technologies. Repeat runs or modifications of run conditions are quickly initiated and do not require the pouring of additional gels. Internal sizing standards run with each sample insure accurate sizing and help to compensate for minor run-to-run variations. Allelic ladders provide a reference allowing for comparisons with other methods and laboratories (3).

[1] Vermont Forensic Laboratory, Waterbury, VT 05676.

* Supported in part under grant 96-IJ-CX-0076 from the Office of Justice Programs, National Institute of Justice, Department of Justice. Points of view in this document are those of the authors and do not necessarily represent the official position of the U.S. Department of Justice.

Received 30 Dec. 1997; and in revised form 5 March 1998; accepted 10 March 1998.

Forensic samples offer a challenge to nearly any analytical method. Sample age or origin may be undetermined and available sample volume small. Material submitted for testing may be deposited on any imaginable surface, or it may be incidentally or intentionally contaminated or obliterated. Biological material brought to the forensic lab may have come from more than one individual. Additionally, the outcome of forensic testing can have serious consequences within the criminal justice system. With these considerations in mind, validation of methods or instrumentation for forensic employment must be rigorous using samples comparable to those found in the field.

The X-Y homologous amelogenin gene codes for an integral protein in mammalian tooth enamel development (4). It has been studied in a number of species, and large areas in both the coding and non-coding regions of the gene have been found to be conserved throughout evolution (5–9). This gene has been used extensively for sex-typing of forensic materials (10–13). Amplification of a small portion of the human amelogenin gene produces distinct fragments from the X and the Y chromosomes and consequently, information regarding the gender of the DNA donor (14–16). Since the sequence required for amplification is relatively small, even highly degraded DNA may yield typeable results.

In the past, sex-typing methods using the amelogenin gene have primarily been gel-based. As a test system for capillary electrophoresis validation, amelogenin offers the simplicity of just two variables and ready availability of samples of known phenotype. What follows is a description of the validation work for CE analysis of amelogenin amplicons.

Materials and Methods

Samples for the studies were extracted using proteinase K (Life Technologies, Gaithersberg, MD) digestion with subsequent extraction with phenol/chloroform with the aqueous extract subjected to Microcon-100 (Amicon Inc., Beverly, MA) dialysis (6). Samples were quantitated by mini-gel electrophoresis employing the DNA dye DAPI (17). For low-concentration extracts that fell below the 4-to-8-ng detectable level for DAPI, the mini-gel was placed in a gel buffer solution containing a 1:10 000 dilution of SYBR™ Green I (Molecular Probes Inc., Eugene, OR) and gently rocked for 20 min. DNA was visualized using ultraviolet (UV)light and evaluated by comparison of relative fluorescence with a series of human concentration standards ranging from 125 to 0.12 ng per lane. Sensitivity of this SYBR Green I technique was approximately 0.25 ng.

Amplifications were performed using amelogenin primers generously provided by PE Applied Biosystems. These primers are identical to those supplied in the multiplex STR kit AmpFISTR™ Green I (PE Applied Biosystems, Foster City, CA), which is designed to amplify three STR loci in addition to the amelogenin locus. Approximately 2 ng of extracted DNA were added to eight picomoles of forward and reverse primers, 0.5 μL of AmpliTaq Gold™ polymerase (5 U/μL) and AmpFISTR™ PCR reaction mix for a final volume of 25 μL. The amplifications were performed using a GeneAmp® PCR System 9600 thermocycler (PE Applied Biosystems) using the following conditions:

95°C for 11 min, then 29 cycles at:
94°C for 1 min
59°C for 1 min
72°C for 1 min

followed by a 30 min hold at 60°C

The capillary electrophoresis unit employed to evaluate the amelogenin amplicons was an ABI PRISM™ 310 Genetic Analyzer (PE Applied Biosystems). An uncoated 50 μm inside diameter capillary (Perkin-Elmer), 47 cm in length with a 37 cm read length (distance from injection end to laser), was used for the separation. The separation medium was the proprietary polymer product Performance Optimized Polymer 4 (PE Applied Biosystems). The run buffer was a 1:10 dilution of 310 Genetic Analyzer Buffer with EDTA (PE Applied Biosystems). Samples were routinely electrokinetically injected at 15 kV for 5 s, followed by a 24 min run at a constant voltage of 15 kV and a constant temperature of 60°C. Prior to each injection the column was back-flushed with new polymer by the instrument.

Amplified amelogenin products were prepared for CE analysis, as was an allelic ladder from the AmpFISTR™ Green I kit, by mixing 1 μL with 0.5 μL of Genescan® 350 ROX internal lane standard (PE Applied Biosystems) and 12 μL of deionized formamide. Formamide (Amresco, Solon, OH) was deionized by using a molecular grade mixed-bed resin (Sigma, St. Louis, MO).

Validation Testing

Stains for testing or for reference samples were prepared from EDTA anti-coagulated blood spotted on cotton cloth or suitable substrate and allowed to air dry. Stains that were used for contamination studies had the appropriate contaminant added directly to the stain. Those samples that were subjected to degradative environmental conditions were placed in a sunny window or heated for the designated time prior to testing. Animal blood dried on cotton or paper, or muscle tissue, was used for the non-human study. Tissues from a single donor were used to determine tissue specificity. Reproducibility was assessed by multiple testings of the same samples. Linearity and minimum sample requirements were evaluated by using the undigested male DNA supplied in the DNA standard reference material 2390 kit (National Institute of Standards and Technology, Gaithersburg, MD). Male and female samples were combined for the mixing studies. Unless otherwise indicated, approximately 2 ng of template DNA was used in each amplification.

The average base pair size, range of sizes and standard deviations were obtained from the data set resulting from the analysis of 107 male and 49 female samples. The variability in quantity of the X and Y products derived from males was determined by examining 82 amplified, standard, male blood samples and calculating the Y/X ratio for each. To identify some of the variables which could influence this ratio, approximately 2 ng of male DNA were amplified in five separate reactions. Each of the five resulting amplifications was split to make two samples that were then processed separately for CE analysis. Processing involved addition of formamide and the internal lane standard followed by heat denaturation, as described above. Finally, these samples were each run three times through the capillary column.

The linearity of the system was examined by varying the concentration of template DNA and also by examining fluorescence versus injection time. Concentrations tested ranged from 0.05 to 10 ng of template DNA. Injection times examined were from 1 to 10 seconds.

Resolution achieved for the amelogenin system was calculated using the following equation as detailed by Luckey et al. 1993 (18):

$$R = [2 \ln 2]^{1/2} * [(T_2 - T_1)/(HW_1 + HW_2)]$$

where T_1 and T_2 are the elution times for peaks 1 and 2 under evaluation and HW_1 and HW_2 are the peak widths measured at one-half the peak height. The resolution in base pairs was determined by dividing the difference between the base pair sizes obtained for the X and Y amplicons by the calculated resolution value.

Results and Discussion

Validation studies for the CE were structured to establish instrument limits such as fluorescence linearity and resolution as well as to develop acceptance guidelines based on data gathered from samples collected from individuals of known gender.

At amplification, fluorescently tagged primers are incorporated into the PCR product. Consequently, the quality and concentration of the template material at amplification and the injection time of the sample at the time of CE analysis affect the level of fluorescence detected by the CE.

Undigested male DNA from the NIST DNA profiling standard 2390 kit was serially diluted, amplified, and tested on the CE to determine concentration-fluorescence linearity. The linear fluorescence range is detailed in Figs. 1*a* and 1*b*. Linearity was seen up to a peak height of approximately 5000 fluorescence units when template DNA was varied from 0.05 to 100 ng. Template amounts of 12.5 ng or more resulted in an abundance of amplified product and fluorescence greater than 5000 units. Results in the median range of 1000 to 4000 fluorescence units were obtained from templates of 1.5 to 6.5 ng. The amplification and subsequent CE analysis of a 0.05 ng sample still yielded peak heights two to three times over background.

The linearity of the system was also examined using injection time as the variable, as seen in Figs. 2*a* and 2*b*. As mentioned before, standard injection time is five seconds; however, this may be increased or decreased to compensate for weak or overly strong fluorescence. When the injection time was varied from one to ten seconds and plotted against either peak height or peak area, the range observed was linear.

Resolution was examined by analyzing the data generated from running ten allelic ladders from the AmpFISTR™ Green I amplification kit. The ladders were run on different days and on different capillaries. Average resolution was found to be 0.931 basepairs for the amelogenin system. The average resolution of the system is a useful measure of the performance of the system and allows tracking to determine when a capillary may need to be replaced or may indicate other issues that must be addressed. For example,

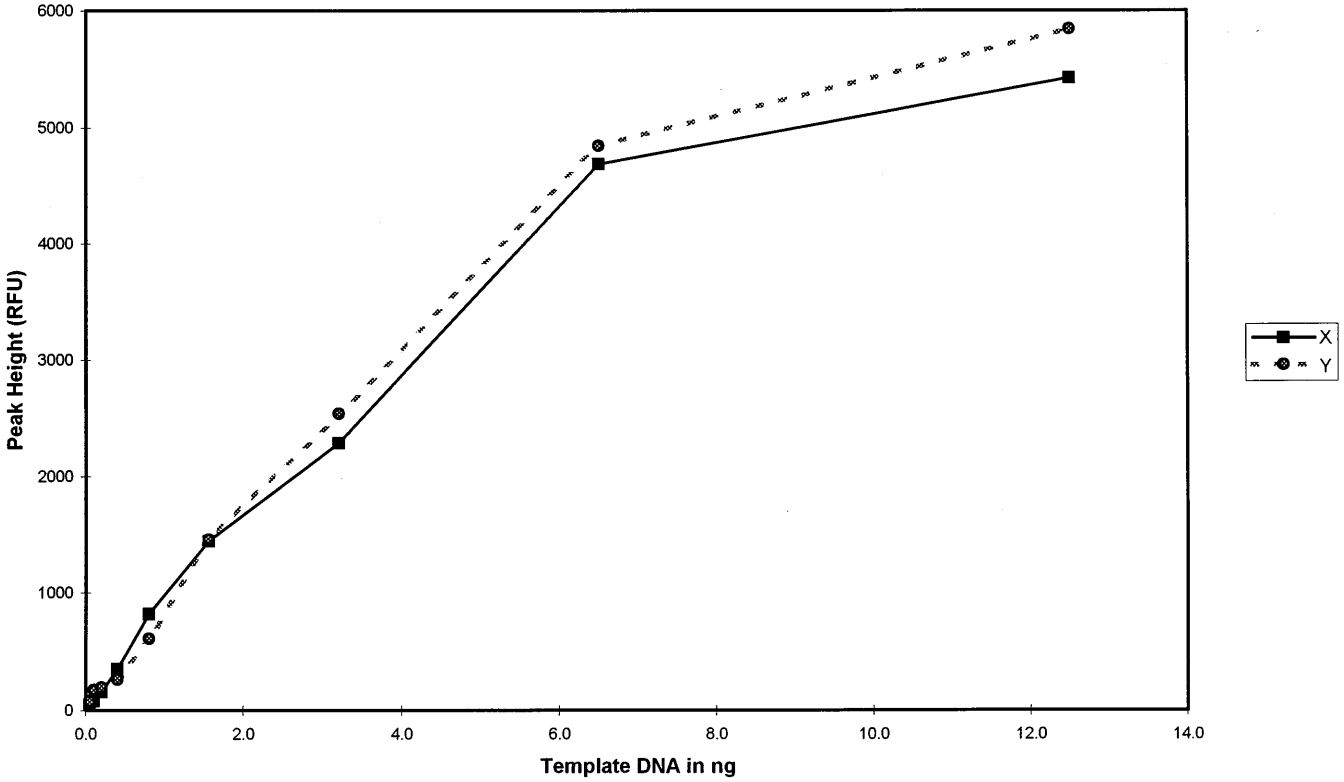


FIG. 1*a*—*Template DNA vs. peak height.*

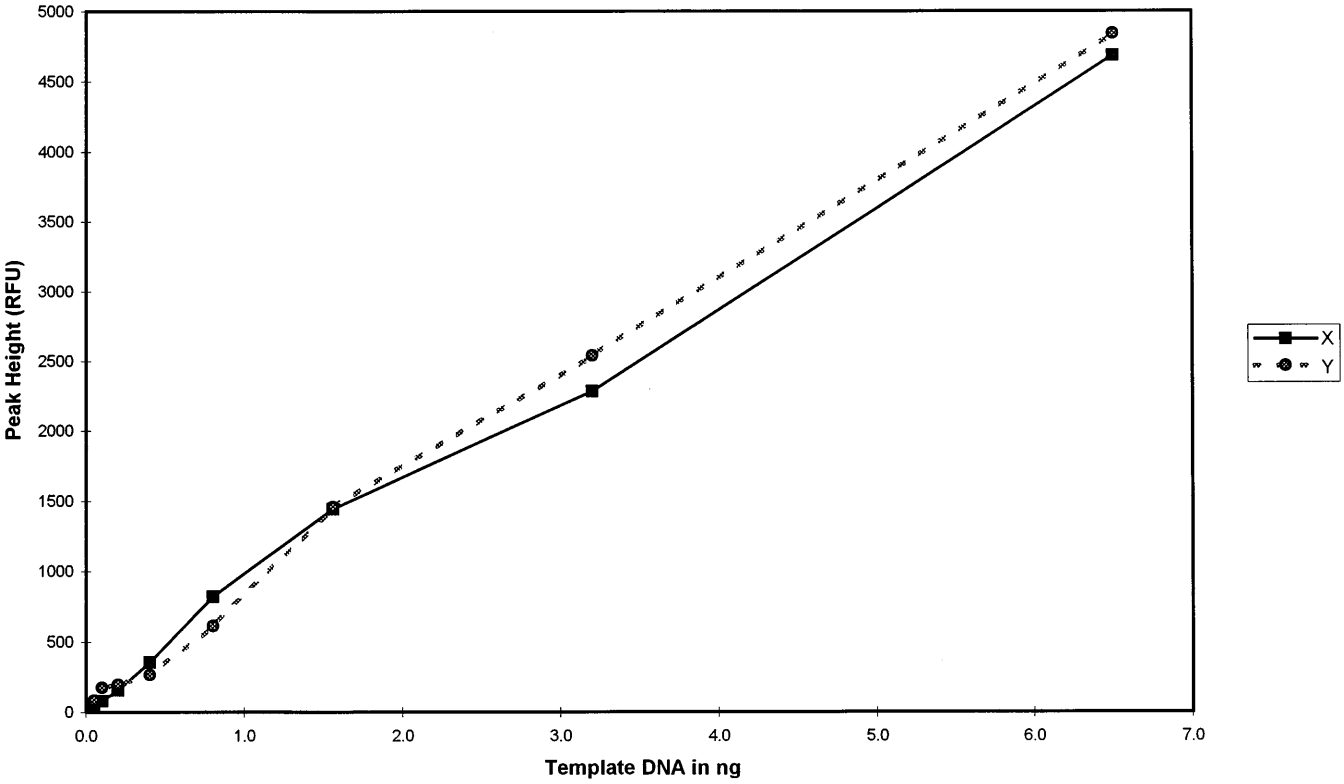


FIG. 1*b*—*Template DNA vs. peak height.*

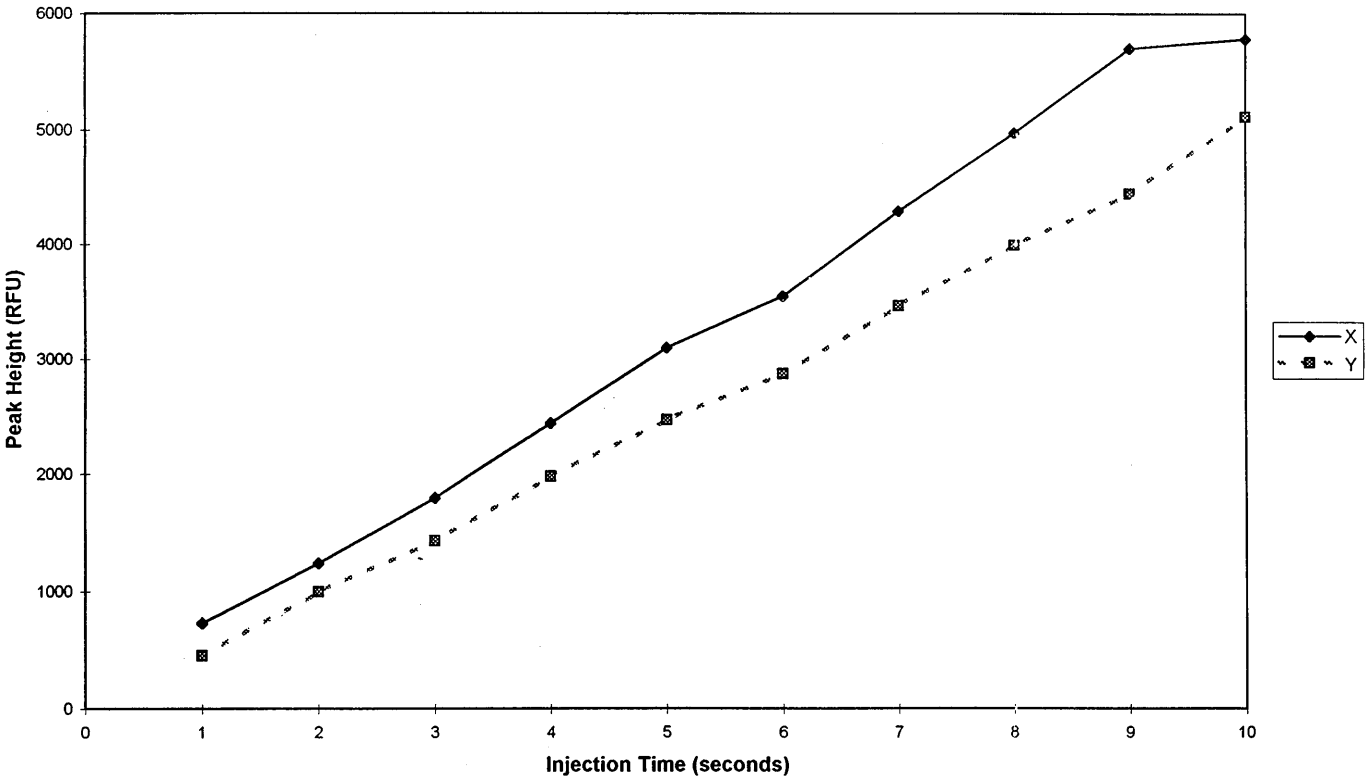


FIG. 2*a*—*Injection time vs. X and Y peak heights.*

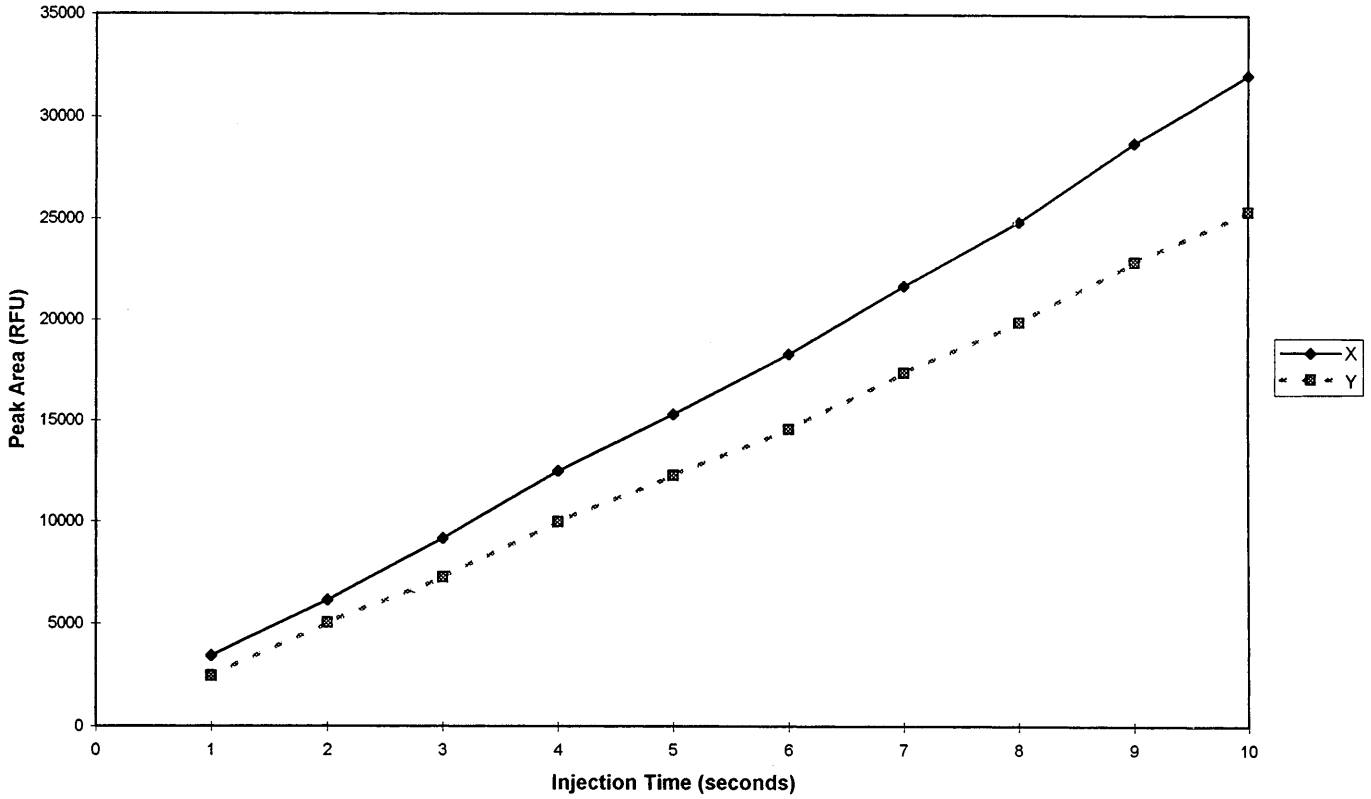


FIG. 2*b*—*Injection time vs. X and Y peak areas.*

amplified PCR products processed for CE analysis using inferior deionized formamide displayed extremely poor resolution. Reprocessing of these samples using formamide deionized with a different resin bed markedly improved resolution. (Data not shown.) Resolution measurements also allow system performance comparisons between laboratories.

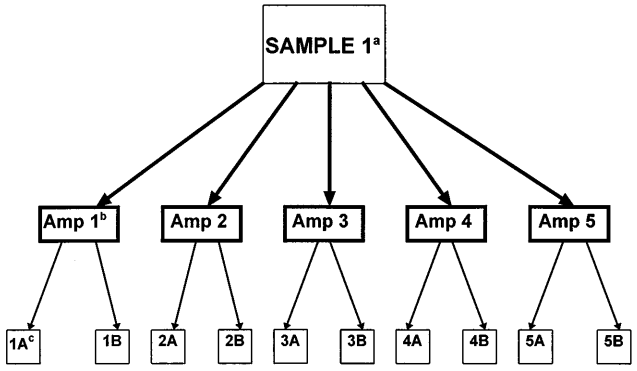
From sequencing data, the expected size of the X and Y amplicons are 107 and 113 basepairs respectively, which includes the addition of a terminal nucleotide to the PCR product (Kathy Lazaruk, Applied Biosystems, Foster City, CA). The average basepair size for the X and Y amplicons seen using CE is 103.15 and 108.80 basepairs. Table 1 summarizes the results of the basepair size study. Amelogenin also has a six basepair size difference in the sequences copied from X and Y chromosomes. Though close to the expected six basepair difference, the 5.65 average basepair difference reflects the uncertainty that may be seen in the actual values obtained by CE analysis.

Table 1 also shows the variability of the Y/X ratio found in male samples. Data from the 82 male samples with peak heights below 5000 RFU, the observed upper limit of fluorescence linearity, were examined. The Y/X peak height ratio of these samples ranged from 0.65 to 1.10, with an average value of 0.88, showing that the X amplicon typically predominates. For these samples showing peak heights within the linear range of fluorescence, Y/X ratio variation appeared independent of relative fluorescence units.

The aim of the work detailed in Fig. 3 was to determine which step, the amplification, the denaturation process, or the electrophoresis, most influenced Y/X ratio variability. First, approximately 10 ng of DNA extracted from a male subject were evenly distributed to five tubes and amplified. Each amplification product was then split in half and processed separately. Finally, each processed product was run three times on the capillary column.

TABLE 1—*Basepair sizes and ratios for X an Y amplicons.*

	Average	Standard Deviation	Minimum Value	Maximum Value
X basepair size	103.14	0.19	102.78	103.92
Y basepair size	108.80	0.17	108.47	109.42
Y/X ratio	0.88	0.08	0.65	1.10



a. The sample was amplified five times.
b. Each amplification product (Amp 1-5) was split in half and processed separately.
c. In addition, each processing (1A-5B) was run through the column three times.

FIG. 3—*Design of the ''controlling variable'' experiment.*

TABLE 2—*Effects of amplification, processing, and electrophoresis on the Y/X height ratio of a sample.*

Sample	Variables	Average Y/X Height Ratio	Minimum Ratio	Maximum Ratio	Range
Sample 1	Amplification, Processing, & Electrophoresis	0.870	0.500	1.115	0.615
Amp 1	Processing & Electrophoresis	0.839	0.791	0.884	0.093
Amp 2	''	1.039	1.022	1.071	0.049
Amp 3	''	0.608	0.500	0.651	0.151
Amp 4	''	0.787	0.628	0.865	0.237
Amp 5	''	1.077	1.038	1.115	0.077
1A	Electrophoresis	0.842	0.834	0.849	0.016
1B	''	0.836	0.791	0.884	0.093
2A	''	1.032	1.022	1.038	0.016
2B	''	1.047	1.033	1.071	0.038
3A	''	0.586	0.500	0.642	0.142
3B	''	0.629	0.617	0.651	0.034
4A	''	0.773	0.628	0.865	0.237
4B	''	0.801	0.780	0.835	0.054
5A	''	1.068	1.038	1.115	0.077
5B	"	1.085	1.075	1.096	0.021

The findings of this experiment, summarized in Table 2, indicate that most of the variability demonstrated by these samples comes from the amplification step. Little change was seen between samples when electrophoresis was the only variable. For example, Sample 1A had an average Y/X height ratio of 0.842, and showed a difference of only 0.016 for separate electrophoretic runs. With separate processings included in the variables, this difference rose to 0.093. Finally for Sample 1, the difference in ratios after all of the amplifications, processings and runs on the CE, was 0.615. This change was significantly greater than those observed when processing and electrophoresis, or electrophoresis alone, were the variables. Amplification, therefore, introduces more of the variability seen in the Y/X height ratio of these samples, than does any of the other steps.

During the analysis of the data, monomorphic peaks were identified. These were sized as approximately 91 and 96 basepairs. These were found regardless of the sex of the DNA donor, but were typically less than 20% of the X or Y amplicon. The 91 peak was usually larger in height than the 96 peak. Since these peaks are not included in the size range examined for amelogenin, further studies were not undertaken.

Crime scene materials are often exposed to degradative environmental conditions or to contaminating substances. These may interfere with DNA amplification and subsequent analysis either by degrading DNA or by inhibiting amplification. To evaluate the ability of the CE to produce reliable results from potentially compromised samples, three sets of samples were set up to simulate a variety of conditions and treatments. Blood was patched to cotton fabric and exposed to differing conditions involving light, temperature and exposure times. Common contaminants were applied directly to bloodstained cotton. Blood was also allowed to dry on typically encountered substrates. Extracted DNA from these stains was amplified and analyzed by CE and compared to untreated reference standards. As seen in Table 3, the results of all samples tested matched those obtained from the reference samples.

Blood, hair, skin and saliva sample sets were taken from two individuals. Amplification of the hair samples failed, but all other

TABLE 3—*Samples subjected to chemical contaminants or deposited on substrates followed by amplification with amelogenin primers and CE analysis.*

Treatment	Amelogenin
CONTROL (uncontaminated)	XY
5% Bleach	XY
Cardboard (substrate)	XY
$CuCl_2$ 0.02M	XY
Denim (substrate)	XY
5% Detergent	XY
10% EDTA	XY
Glass (substrate)	XY
3% H_2O_2	XY
Metal (substrate)	XY
$MgCl_2$ 0.02M	XY
$MnCl_2$ 0.02M	XY
$NiCl_2$ 0.02M	XY

TABLE 4—*Animal samples amplified with amelogenin primers and analyzed by CE.*

Animal	Peak size
Bear	99.04
Bobcat	99.29
Cat	no amplified product
Chicken	no amplified product
Cow	98.59
Dog	98.94
Deer	98.71
Goat	99.06
Horse	98.71
Human (male)	102.92/108.61
Moose	98.60
Porcupine	no amplified product
Rabbit	no amplified product
Raccoon	99.29
Sheep	98.94
Swine	99.64
Turkey	no amplified product

materials amplified and typed accurately. Subsequently, other hair samples have been successfully amplified. Correct typing of these samples was obtained by CE analysis.

A panel of non-primate animal samples representing both common domestic and wild species was amplified using the amelogenin primers. Coding for an integral protein of mammalian tooth enamel development, the amelogenin gene has been studied in a number of species and is found to have large areas in both the coding and non-coding regions which have been conserved throughout evolution. Five of the 16 samples tested yielded no amplified product. The successful amplifications each exhibited a peak in the range of 98.59 to 99.64 basepairs. This distinctive peak could not be confused with either the X or the Y peak or with the monomorphic peaks sometimes seen in human samples. (See Table 4.) No significant differences in basepair size were noted between the sexes when amplicons derived from pairs of animals were analyzed. However, gender and peak height correlations in these samples were not explored.

Data derived from the mixed male and female sample studies indicate that a sample within the linear fluorescence range, that is, up to 5000 RFUs, and containing a 108 basepair peak above the background fluorescence, possesses a male component. Male : female mixtures of 1 : 20 to 1 : 30 yielded a 108 base pair peak with fluorescence three times background. As would be expected in mixes such as this in which the female fraction is much larger than the male, the X peak is significantly larger than the Y peak.

More difficult to evaluate were mixes in which the male and female fractions approached equal amounts or the female component was the lesser. As mentioned earlier, the Y/X ratio for standard male samples ranged from 0.65 to 1.10 with an average of 0.88. While a Y/X ratio less than 0.65 may be suggestive of a mix, greater deviations may be encountered (Stephen G. LaBonne, New York State Police Investigation Center, Albany, NY), and care should be used in interpreting these mixtures.

Conclusions

Tandem use of fluorescently tagged amelogenin primers and CE technology is a convenient and reliable methodology for gender determination of forensic samples. Amplified samples from simulated field conditions and from standards were shown to type correctly. Samples derived from low concentrations of template material could be satisfactorily typed. Since amelogenin products are very small, even highly degraded samples should be amenable to typing. Additionally, this method allows unambiguous distinction between human and non-primate animal samples.

The validation work presented here established operating parameters for our CE instrumentation and criteria for linearity and resolution. Also established was a range for the Y/X ratio of standard male samples, which we expect will be useful for interpreting mixes of male and female samples.

References

1. Northrop D, McCord B, Butler J. Forensic applications of capillary electrophoresis. J Capillary Electrophoresis 1994;1:158–68.
2. Butler J, McCord B, Jung J, Lee J, Budlowe B, Allen R. Application of dual internal standards for precise sizing of polymerase chain reaction products using capillary electrophoresis. Electrophoresis 1995;16:974–80.
3. Buel E, Schwartz M, LaFountain M. Capillary electrophoresis STR analysis: comparison to gel-based systems. J Forensic Sci 1998; 43(1):164–70.
4. Nakahori Y, Takenka O, Nakagome Y. A human X-Y homologous region encodes 'Amelogenin.' Genomics 1991;9:264–9.
5. Bailey D, Affara N, Ferguson-Smith M. The X-Y homologous gene amelogenin maps to the short arms of both the X and Y chromosomes and is highly conserved in primates. Genomics 1992;14: 203–5.
6. Buel E, Wang G, Schwartz M. PCR amplification of animal DNA with human X-Y primers used in gender determination. J Forensic Sci 1995;40:641–4.
7. Gibson C, Golub E, Abrams W, Shen G, Ding W, Rosenbloom J. Bovine amelogenin message heterogeneity: alternative splicing and Y-chromosome gene transcription. Biochemistry 1992;31:8384–8.
8. Lyngstadaas S, Risnes S, Nordbo H, Flones A. Amelogenin gene similarity in vertebrates: DNA sequences encoding amelogenin seem to be conserved during evolution. J Comp Physiol Bio 1990; 160:469–72.
9. Watson J, Spencer J, Marshall Graves J, Snead M, Lau E. Autosomal localization of the amelogenin gene in monotremes and marsupials: Implications for mammalian sex chromosome evolution. Genomics 1992;14:785–9.
10. Akane A, Seki S, Shioni H, Nakamura H, Hasegawa M, Kagawa M, et al. Sex determination of forensic samples by dual PCR amplification of an X-Y homologous gene. Forensic Sci Int 1992;52: 143–8.
11. Akane A, Shiono H, Matsubara K, Nakahori Y, Seki S, Nagafuchi S, et al. Sex identification of forensic specimens by polymerase chain reaction (PCR): two alternative methods. Forensic Sci Int 1991;49:81–8.
12. Casarino L, De Stefano F, Mannucci A, Canale A. HLA-DQA1

and amelogenin coamplification: a handy tool for identification. J Forensic Sci 1995;40(3):456–8.
13. Mannucci A, Sullivan K, Ivanov P, Gill P. Forensic application of a rapid and quantitative DNA sex test by amplification of the X-Y homologous gene amelogenin. Int J Legal Med 1994;106:190–3.
14. Roy R, Steffens D. Infrared fluorescent detection of PCR amplified gender identifying alleles. J Forensic Sci 1997;42(3):452–60.
15. Nakahori Y, Hanamo K, Iwaya M, Nakagome Y. Sex identification by polymerase chain reaction using X-Y homologous primers. Am J Med Gen 1991;39:472–3.
16. Sullivan K, Mannucci A, Kimpton C, Gill P. A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of X-Y homologous gene amelogenin. Biotechniques 1993;15:636–41.
17. Buel E, Schwartz M. DAPI, a simple alternative to ethidium bromide staining of DNA in agarose gels. Appl Theor Electrophoresis 1993;3:253–5.
18. Luckey J, Norris T, Smith L. Analysis of resolution in DNA sequencing by capillary gel electrophoresis. J Phys Chem 1993; 97:3067–75.

Additional information and reprint requests:
Eric Buel, Ph.D.
Vermont Forensic Laboratory
103 South Main Street
Waterbury, VT 05676