Amanda Fazi[1,2]*
Brianne Gobeski[1,3]*
David Foran[1,4]

[1]Forensic Science Program, School of Criminal Justice, Michigan State University, East Lansing, MI, USA
[2]Northville Crime Laboratory, Michigan State Police Forensic Science Division, Northville, MI, USA
[3]Laboratory Services Bureau, Phoenix Police Department, Phoenix, AZ, USA
[4]Department of Zoology, Michigan State University, East Lansing, MI, USA

Received February 28, 2014
Revised August 21, 2014
Accepted August 22, 2014

Research Article

Development of two highly sensitive forensic sex determination assays based on human DYZ1 and Alu repetitive DNA elements

Sex determination is a critical component of forensic identification, the standard genetic method for which is detection of the single copy amelogenin gene that has differing homologues on the X and Y chromosomes. However, this assay may not be sensitive enough when DNA samples are minute or highly compromised, thus other strategies for sex determination are needed. In the current research, two ultrasensitive sexing assays, based on real-time PCR and pyrosequencing, were developed targeting the highly repetitive elements DYZ1 on the Y chromosome and Alu on the autosomes. The DYZ1/Alu strategy was compared to amelogenin for overall sensitivity based on high molecular weight and degraded DNA, followed by assaying the sex of 34 touch DNA samples and DNA from 30 hair shafts. The real-time DYZ1/Alu assay proved to be approximately 1500 times more sensitive than its amelogenin counterpart based on high molecular weight DNA, and even more sensitive when sexing degraded DNA. The pyrosequencing DYZ1/Alu assay correctly sexed 26 of the touch DNAs, compared to six using amelogenin. Hair shaft DNAs showed equally improved sexing results using the DYZ1/Alu assays. Overall, both DYZ1/Alu assays were far more sensitive and accurate than was the amelogenin assay, and thus show great utility for sexing poor quality and low quantity DNA evidence.

Keywords:
Alu / DYZ1 / Pyrosequencing / Real-time PCR / Sex determination
DOI 10.1002/elps.201400103

1 Introduction

DNA analysis of biological evidence submitted to US crime laboratories is based on the amplification of thirteen or more autosomal STR loci, using a recommended 0.5–2.0 ng of input DNA. The amplicons are then separated by CE and visualized utilizing an optical sensor within a genetic analyzer. Current DNA typing techniques are extremely effective for high quality samples, but they often fail when applied to forensic specimens that harbor little or no nuclear DNA, such as highly degraded material, aged bone, teeth, hair shafts, and touch DNA [1]. In cases where partial or no DNA profiles are generated due to low quantity or quality of DNA, it may still be possible to obtain genetic information from the contributor using more sensitive molecular techniques.

Sex is also assessed using commercial forensic STR typing kits, via the amplification of amelogenin, a single copy gene with homologues on the X and Y chromosomes [2]. Both homologues amplify with male DNA, while only the X homologue amplifies with female DNA. This assay is generally robust, however problems can arise if mutations occur in a primer binding site such that one of the amplicons is lost [3,4]. Further, because amelogenin is a single copy locus, it is susceptible to dropout if very small amounts of DNA are analyzed or if the DNA is highly degraded. Because of this, a sex determination assay that focuses on multicopy loci, including ones specific to the Y chromosome, can enhance accurate sexing results.

Long repetitive sequences comprise over 60% of the human Y chromosome [5], thus there are several loci that could be useful for a highly sensitive sex determination assay. These elements, designated DYZ, are numbered in order of abundance. DYZ1 and DYZ2 are located on the long arm of the Y chromosome, the repeat units of which are 3.4 and 2.4 kb, respectively [5, 6]. DYZ1 can have up to 4000 copies [6], and DYZ2 has about 2000 copies [7]. The remaining three elements (DYZ3, DYZ4, and DYZ5) have 50 copies or less and range in size from 137 to 170 bp [5, 8]. DYZ3 is a block of alphoid satellite DNA positioned on the centromeric region of the Y chromosome, while DYZ4 and DYZ5 are located on the short arm of the Y chromosome.

Correspondence: Dr. David Foran, Forensic Science Program, School of Criminal Justice, Michigan State University, 560 Baker Hall, East Lansing, MI 48824, USA
E-mail: foran@msu.edu
Fax: +1-517-432-1787

*These authors have contributed equally to this work.

Colour Online: See the article online to view Figs. 1–3 in colour.

Table 1. Real-time and pyrosequencing primers

Assay	Target (bp)	Forward (5′–3′)	Reverse (5′–3′)
Amel	X (69)	TCCCAGATGTTTCTCAAGTGG[a)]	ATCAGAGCTTAAACTGGGAAGCTG[b)]
	Y (77)	CATCCCAAATAAAGTGGTTTCTC[a)]	ATCAGAGCTTAAACTGGGAAGCTG[b)]
RT	Alu (113)	GAGATCGAGACCATCCCGGCTAAA[c)] probe: DHEX-GGGCGTAGTGGCGGG-DBH1[c)]	CTCAGCCTCCCAAGTAGCTG[c)]
	DYZ1 (143)	GGCCTGTCCATTACACTACATTCC[d)] probe: 6FAM-ATTCCAATCCATTCCTTT-MGBNFQ[d)]	GAATTGAATGGAATGGGAACGA[d)]
PYS	Alu (113)	GAGATCGAGACCATCCCGGCTAAA[c)]	CTCAGCCTCCCAAGTAGCTG[c)]
	DYZ1 (60)	ATTCCATTCCAATCCATTCCTTT	GCAGTAGAAGAGAATAGAATGGAAT

Primers used in the real-time amelogenin (Amel), real-time DYZ1/Alu (RT), and pyrosequencing DYZ1/Alu (PYS) assays are listed with their respective target and amplicon size (bp). The Alu TaqMan probe included the 5′-reporter dye Hex (DHEX) and the nonfluorescent 3′-quencher Black Hole-1 (DBH1). DYZ1 real-time primers targeted a 143 bp amplicon and a probe with 5′-reporter dye 6-FAM and a 3′-minor groove binding nonfluorescent quencher (MGBNFQ).
a) Primers derived using Primer3 v.0.4.0 (http://frodo.wi.mit.edu).
b) Primers derived from [13].
c) Primers derived from [14].
d) Primers derived from [15].

Detection of a repetitive element unique to the human Y chromosome shows that a forensic sample originated from a male, while its absence denotes a female origin. However, an assay based solely on the presence or absence of a Y chromosome marker may incorrectly sex a sample as female if amplification of the Y target fails [9, 10]. In this regard, simultaneously assaying a region on autosomes or the X chromosome can serve as a control and thus represents a more accurate sexing assay. Alu is the most abundant repetitive element in human DNA, with an estimated 1 million copies, accounting for more than 10% of the genome [11]. Nicklas and Buel [12] developed a sensitive sexing assay to quantify total and male DNA that targeted Ya5, an Alu subfamily, along with DYZ5. Although this assay is specific and reproducible, assaying DYZ5 still has sensitivity limitations, and amplification of more abundant DYZ repetitive elements could substantially increase the sensitivity of this type of assay. In this regard, diagnostic studies utilizing the DYZ1 repetitive element [6] suggest that it may be well suited for highly sensitive testing of forensic samples.

In the research presented here, two ultrasensitive assays for DNA-based sex determination of forensic samples were developed, based on real-time PCR and pyrosequencing of Alu and DYZ1. These assays were optimized with control male and female DNAs, and then compared to a standard amelogenin assay and to each other. Finally, DNAs from human hair shafts were sexed in a blind study using the amelogenin and DYZ1/Alu assays to compare the accuracy and sensitivity of each sexing technique.

2 Materials and methods

Assays were developed in bleached and/or UV-sterilized hoods, while wearing coats, gloves, masks, sleeves, etc. Reagents and equipment were UV-irradiated in a Spectrolinker XL-1500 UV Crosslinker (Spectronics, Westbury, NY, USA) for minimally 5 min (~2.5 J/cm^2). Primers were reconstituted in purified sterile water (Fisher Scientific, Pittsburg, PA, USA), and filtered through either Microcon YM-30 or YM-100 columns (Millipore, Billerica, MA, USA).

2.1 Real-time amelogenin assay

A segment of the amelogenin locus was amplified using different forward primers and a common reverse primer (Table 1). The forward X chromosome primer spanned the 6 bp deletion in the amelogenin gene detected in standard forensic STR kits [13], while the forward Y chromosome primer spanned the same region on the Y chromosome, resulting in amplicon sizes of 69 and 77 bp, respectively. Amelogenin real-time amplification parameters consisted of 15 μL reactions with 1× Fast SYBR Green Master Mix (Roche Applied Science, Indianapolis, IN, USA) or Power SYBR Green Master Mix (Life Technologies, Carlsbad, CA, USA), 2 μM each primer, 1 μL DNA, 10 μg BSA, and sterile water. Amplification was conducted with an initial denaturation of 95°C for 10 min, followed by 40 cycles of 95°C denaturation for 15 s, and 58.5°C annealing/extension for 1 min. PCR assays were performed on an iQ5 Multicolor Real-Time Detection System and iCycler (Bio-Rad Laboratories, Hercules, CA, USA) and analyzed with Bio-Rad iQ5—Standard Edition software. PCR products were also separated on a 4% agarose gel to determine if amplicons were the expected size.

2.2 Real-time DYZ1/Alu assay

A 113 bp region within the Alu repeat element [14] was targeted as a positive human control for the real-time assay. The Alu TaqMan probe (Table 1) included the 5′-reporter

dye Hex and the nonfluorescent 3′-quencher Black Hole-1. DYZ1 primers targeted a 143 bp amplicon and a probe with 5′-reporter dye 6-FAM along with a 3′-minor groove binding nonfluorescent quencher. DYZ1 and Alu primers and probes were combined in a TaqMan real-time PCR assay. Optimized 15 μL reactions included 1× iQ Supermix (Bio-Rad Laboratories), 500 nM each forward primer, 900 nM each reverse primer, 250 nM each probe, 1 μL DNA, 10 μg BSA, and sterile water. Cycling parameters included an initial 95°C hold for 10 min, and 50 cycles of a 95°C denaturation for 15 s and 60°C annealing/extension for 1 min.

The sensitivity of the real-time amelogenin and DYZ1/Alu assays was compared based on 15-fold dilutions of high molecular weight male DNA (made in house) and DNA digested with DNase I (Roche Applied Science). Target quantities of high molecular weight DNA ranged from 20 000 to 0.00176 pg, and from 47 000 to 0.00413 pg for DNase digested DNA. Each DNA quantity was tested in duplicate for both the real-time amelogenin and DYZ1/Alu assays.

2.3 Pyrosequencing DYZ1/Alu assay

Amplification of Alu was performed using the primers shown in Table 1, while DYZ1 primers were designed using Beacon Designer 7 Software (PREMIER BioSoft International, Palo Alto, CA, USA), yielding a 60 bp product (Table 1). 5′-biotinylated pyrosequencing primers were HPLC-purified during production (Integrated DNA Technologies, Coralville, IA, USA and Sigma-Aldrich, St. Louis, MO, USA). DYZ1 and Alu primers were combined in 20 μL reactions, including 0.75 U AmpliTaq Gold DNA Polymerase (Life Technologies), 2 μM each primer, 1 ng male or female control DNA or 2 μL DNA from degraded or low quality samples, GeneAmp 10× PCR Buffer II (Life Technologies), 0.2 mM each dNTP, and sterile water (Fisher Scientific) to volume. PCR parameters consisted of an initial 94°C denaturation for 10 min, and 35 cycles of 94°C for 30 s, 56°C for 1 min, 72°C for 20 s, followed by a final extension of 72°C for 5 min.

DNA extracts were prepared for pyrosequencing on a PyroMark Q24 vacuum workstation (Qiagen, Germantown, MD, USA) in SQA mode according to the manufacturer's instructions, using Streptavidin Sepharose High Performance Beads (GE Healthcare, South San Francisco, CA, USA). Sequencing reactions were conducted on an automated microtiter plate based PyroMark Q24 Pyrosequencer (Qiagen) with method parameter code 001 Rev. A, correlating to the cartridge used. Primers for multiplex sequencing were at 0.4 μM, as suggested by the manufacturer. Data were analyzed using PyroMark Q24 Software 2.0.6 (Qiagen). Dispensation order was generated specific to the target DYZ1 and Alu sequences immediately adjacent to the sequencing primers (TACGTCGACTGTCATC(AAGGTTCC)$_{2–3}$). Additionally, erroneous bases were dispensed at the beginning and scattered throughout, and the end dispensation consisted of a simple repeat pattern for monitoring purposes.

The pyrosequencing DYZ1/Alu assay was applied to touch DNAs obtained in house from T-shirts worn by 20 individuals. These touch DNAs, ranging from 0 to 101 pg/μL, had been previously quantified with Quantifiler Human DNA Quantification Kit (Life Technologies), and amelogenin sexing results were previously obtained using an AmpFlSTR Identifiler PCR Amplification Kit (Life Technologies) [16].

2.4 Application of assays to hair shaft DNAs

The real-time amelogenin assay, real-time DYZ1/Alu assay, and pyrosequencing DYZ1/Alu assay were each used to blindly assess sex from 71 different DNA extracts of 30 hair shafts of known sex, which had previously been collected in house and extracted using either an organic or alkaline method [17]. In any instance where both male and female results were obtained from different extracts of the same hair, the result was deemed "inconclusive." In cases where no results were obtained from any extracts of a given hair, the hair was designated as "no result".

2.5 Binary comparisons

The true positive rate, true negative rate, positive predictive value, and negative predictive value of the amelogenin and the two DYZ1/Alu assays were determined according to [18]. Assays with no or inconclusive results were considered incorrect for these calculations.

3 Results

3.1 Real-time amelogenin and DYZ1/Alu assays

Exemplary amelogenin real-time assay results based on control DNA are displayed in Fig. 1, wherein male DNA produced two curves while female DNA produced a single curve. Similarly, the DYZ1/Alu assay produced two curves for control male DNA and one for female DNA (Fig. 2).

The relative sensitivity of the amelogenin and DYZ1/Alu assays, based on 15-fold dilutions of high molecular weight or DNase treated DNA, is shown in Table 2. The amelogenin assay generated a positive male result in at least one replicate down to 88.9 pg of high molecular weight DNA, while the DYZ1/Alu assay was approximately 1500 times more sensitive, with positive male results obtained from at least one replicate with 0.0623 pg of high molecular weight DNA, below which DYZ1 failed to amplify. The lowest tested degraded DNA concentration from which amelogenin amplified was 208.9 pg, while the DYZ1/Alu assay was sensitive down to 0.00413 pg. This represents a 50 000-fold improvement over amelogenin, although these data are approximations that are bound by the dilution factor. As the quantity and quality of DNA decreased, the real-time curves often moved away from one another, with the DYZ1 curve reaching the C_t value later. Owing to this, results were considered male positive if both

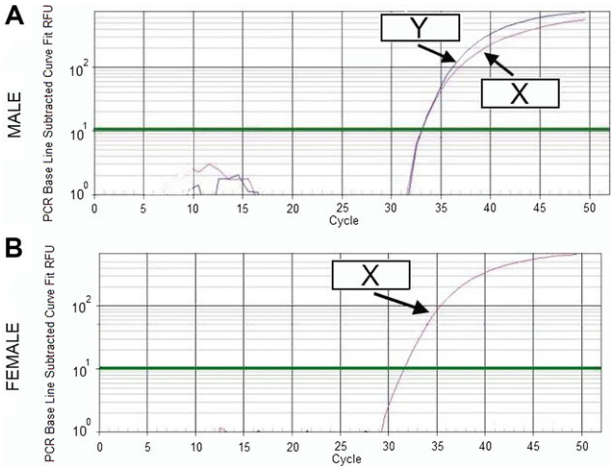

Figure 1. Exemplary amplification curves for the amelogenin assay using male and female DNA (*x*-axis = cycle number, *y*-axis = RFU value). The bold line represents the threshold value. (A) An amplification plot for male DNA. Both the X and Y homologues amplified with a C_t value of 33 cycles. (B) An amplification plot for female DNA. Only the X homologue amplified with a C_t value of 31 cycles.

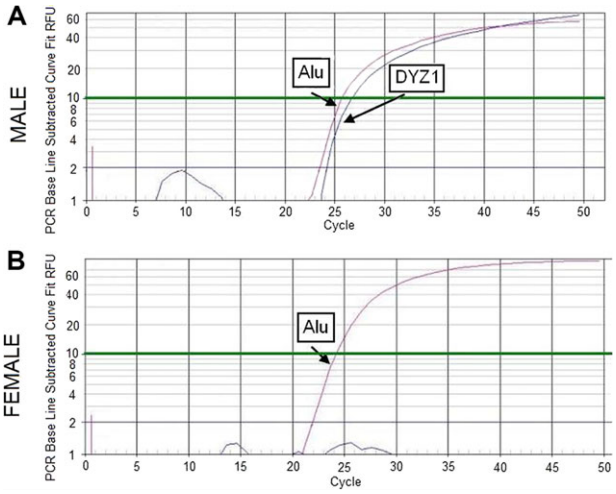

Figure 2. Exemplary amplification curves for the real-time DYZ1/Alu assay using male and female DNA (*x*-axis = cycle number, *y*-axis = RFU value). The bold line represents the threshold value. (A) An amplification plot for male DNA. Both the DYZ1 and Alu loci amplified with similar C_t values of 27 and 26 cycles, respectively. (B) An amplification plot for female DNA. Only the Alu locus amplified with a C_t value of 24 cycles.

curves crossed the threshold before 40 cycles, and were within seven cycles of one another.

3.2 Pyrosequencing DYZ1/Alu assay

Pyrosequencing of control DNAs resulted in pyrograms with a clean combination of DYZ1 and Alu sequences for male DNA, and a pure Alu sequence for female DNA (Fig. 3). Negative control pyrograms resulted in no called bases. Comparison of the sexing data from touch samples using a standard STR kit and pyrosequencing DYZ1/Alu assay is displayed in Table 3. Amelogenin amplification from the kit produced correct results six times, incorrect results five times, inconclusive results one time, and no results 22 times. In contrast, the pyrosequencing DYZ1/Alu assay produced correct results 26 times, incorrect results seven times, and no result once.

Table 2. Sexing results of the dilution series for the amelogenin and real-time DYZ1/Alu assays

Dilution series	Amelogenin assay		DYZ1/Alu assay	
	X	Y	Alu	DYZ1
HMW DNA				
20 000 pg	+	+	+	+
1330 pg	+	+	+	+
88.89 pg	+	+	+	+
5.93 pg	−	−	+	+
0.393 pg	−	−	+	+
0.0623 pg	−	−	+	+
0.00176 pg	−	−	+	−
Digested DNA				
47 000 pg	+	+	+	+
3130 pg	+	+	+	+
208.9 pg	+	+	+	+
13.93 pg	−	+	+	+
0.928 pg	−	−	+	+
0.0619 pg	−	−	+	+
0.00413 pg	−	−	+	+

The amplification success for the dilution series of high molecular weight (HMW) and digested male DNA for the amelogenin and real-time DYZ1/Alu assay. A locus that correctly amplified in at least one replicate is indicated with a plus sign (+) and a locus without amplification is indicated with a minus sign (−).

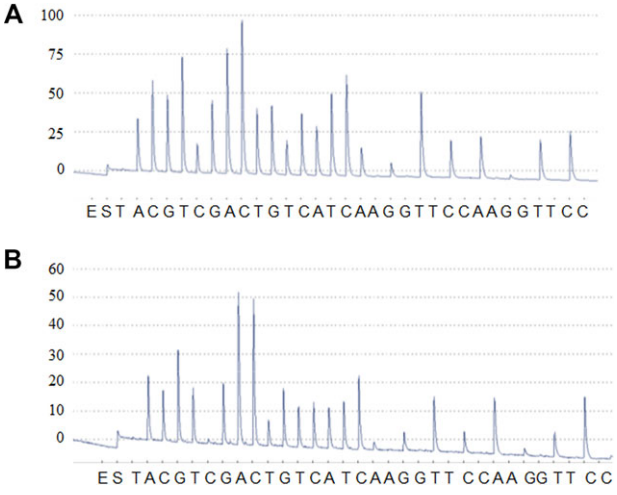

Figure 3. Representative pyrograms generated from male (A) and female (B) control DNA. Sex of DNA extracts were assigned based on which pyrogram they matched. The *x*-axis corresponds to dispensation order of nucleotides and *y*-axis to luminescence intensity. Enzyme (E) and substrate (S) peaks correspond to internal controls.

Table 3. Amelogenin and pyrosequencing DYZ1/Alu sexing results of DNA from 34 touch samples

ID	Known	Amel	DYZ1/Alu
1	M	N	M
2	F	F	M
3	F	N	N
4	M	N	F
5	M	N	M
6	F	N	F
7	F	N	F
8	F	N	F
9	F	N	F
10	M	M[a)]	M
11	F	N	F
12	M	N	M
13	M	N	M
14	M	N	F
15	F	N	F
16	F	N	F
17	M	F	M
18	M	F	M
19	F	N	F
20	M	N	M
21	M	N	M
22	F	F	M
23	M	F	M
24	M	F	M
25	F	N	M
26	M	F	M
27	M	N	M
28	M	M	M
29	F	N	F
30	M	N	M
31	F	F	F
32	F	F	M
33	F	F	M
34	M	N	M

Previously collected touch DNAs with known sex and sexing results based on amelogenin (Amel) [16] were compared to the pyrosequencing DYZ1/Alu assay in this study. Sexing results are listed as male (M), female (F), or no result (N). Sexing results inconsistent with knowns are underlined.
a) No X allele was detected in the sample, thus it was inconclusive.

3.3 Hair shaft sexing using all assays

Known sex of each of the 30 hair donors and the sexing results from the real-time amelogenin, real-time DYZ1/Alu, and pyrosequencing DYZ1/Alu assays are displayed in Table 4. The amelogenin real-time assay produced 10 of 30 hair results consistent with the known sex, nine of which were from females while one was from a male. The remaining 20 hairs yielded no results. The real-time DYZ1/Alu assay yielded 29 of 30 sexing results consistent with the known sex (22 females and 7 males), while one hair from a male donor typed as female. The pyrosequencing DYZ1/Alu assay yielded sexing results from 14 hairs that matched the known sex: nine female and five male. Of the remaining 16 hairs, ten yielded no pyrosequencing results and six yielded inconclusive results.

Table 4. Sexing results of DNA from 30 hair shafts

ID	Known	Amel	RT	PYS
1	F	N	F	F
2	F	F	F	F
3	F	N	F	I
4	M	N	F	M
5	F	N	F	N
6	F	N	F	N
7	F	F	F	N
8	F	F	F	N
9	M	N	M	M
10	F	N	F	F
11	F	F	F	N
12	M	N	M	M
13	F	N	F	F
14	M	N	M	M
15	M	N	M	M
16	F	N	F	I
17	M	N	M	I
18	F	F	F	F
19	F	N	F	F
20	F	N	F	F
21	F	F	F	N
22	F	N	F	N
23	F	F	F	N
24	M	M	M	I
25	F	F	F	I
26	M	N	M	I
27	F	N	F	F
28	F	N	F	N
29	F	N	F	F
30	F	F	F	N

Thirty hair shaft DNAs collected and extracted with known sex and sexing results from the real-time amelogenin (Amel), real-time DYZ1/Alu (RT), and pyrosequencing DYZ1/Alu (PYS) assays are classified as male (M), female (F), inconclusive (I), or no result (N). Sexing results inconsistent with the known are underlined.

3.4 Binary comparison of sexing assays

Relative comparison of the three sexing techniques are shown in Table 5. In all instances, the DYZ1/Alu assays outperformed the standard amelogenin assay in identifying both males and females, and not misidentifying the sexes. The sole exception was positive identification of female hairs using amelogenin and pyrosequencing, which were equal, although the latter assay still outperformed in the other measures (see discussion below).

4 Discussion

The DYZ1/Alu real-time and pyrosequencing strategies, based on highly repetitive autosomal and Y chromosome elements, were designed as more sensitive forensic sexing

Table 5. True positive rate, true negative rate, positive predictive value, and negative predictive value of the three assays

Sample	True pos rate	True neg rate	Pos pre value	Neg pre value
Amelo shirts male	0.06	0.31	0.08	0.23
Amelo shirts female	0.31	0.06	0.23	0.08
Amelo hairs male	0.13	0.41	0.07	0.56
Amelo hairs female	0.41	0.13	0.56	0.07
Pyro shirts male	0.89	0.63	0.73	0.83
Pyro shirts female	0.63	0.89	0.83	0.73
Pyro hairs male	0.63	0.41	0.28	0.75
Pyro hairs female	0.41	0.63	0.75	0.28
RT hairs male	0.88	1	1	0.96
RT hairs female	1	0.88	0.96	1

Data are from Tables 3 and 4. Assays are as follows: Amelo = amelogenin; Pyro = DYZ1/Alu pyrosequencing; RT = DYZ1/Alu real time PCR. 'True positive rate" (True pos rate) is correctly identifying a sex, and 'True negative rate" (True neg rate) is not identifying the wrong sex. "Positive predictive value" (Pos pre value) is the likelihood of a positive result being accurate, while "negative predictive value" (Neg pre value) is the same for a negative result. Tests with no results (N) were considered incorrect, as all samples should have a sex.

assays relative to the traditional one based on amelogenin. Targeting high-copy loci produced real-time results that were up to 1500 times more sensitive than the amelogenin assay when testing high molecular weight DNA, and even more sensitive when testing degraded DNA (discussed below). The assay was successful down to 0.004 pg of starting DNA, which is orders of magnitude lower than that tested in [12], which was based on assaying DYZ5. Similarly, DYZ1/Alu pyrosequencing produced results far superior to amelogenin when testing touch samples, many of which had previously quantified at or near zero using Quantifiler. Finally, both assays generated greatly improved results from hair shaft DNAs, where nuclear DNA typing is rarely successful.

Comparison of the real-time and pyrosequencing assays led to several interesting observations beyond the drastic improvement in sensitivity relative to amelogenin. An occasional difficulty with the real-time assay was determining a clear distinction between male and female results. This was never a problem with reasonably high quality DNA (a very relative term, as even hair shafts produced good results), wherein the DYZ1 and Alu curves crossed the threshold at similar cycles, however if the DYZ1 curve crossed the threshold several cycles after Alu it was unclear if a sex call of male or female should be made. In the end, a difference within seven cycles was selected as the cutoff, after which the result was denoted female. The cycle cutoff was somewhat arbitrary, however for this research it was necessary to establish a defined limit so that results could be categorized as correct or incorrect. In a forensic setting, the "correct" results are not known, thus obtaining an inconclusive result may be acceptable. Further, male/female mixtures are often encountered in forensic samples, and large differences in male/female signals are not unexpected. While mixtures were not tested in the current study, the effectiveness of their interpretation using either of the extremely sensitive DYZ1/Alu assays would be worth examining.

The nature of the pyrosequencing DYZ1/Alu assay was such that it did not generate ambiguous results: the data were readily interpretable as either male or female, a definite advantage as no arbitrary cutoff was required. On the other hand, because the assay is not quantitative, any combination of X and Y chromosomes will necessarily type as male. Still, the pyrosequencing assay proved to be extremely sensitive. Results from the touch samples showed a substantial increase in sexing success when compared to amelogenin. This included numerous correct calls for samples where amelogenin was negative, and correct male calls when amelogenin signified only an X chromosome existed. There were some instances where DYZ1/Alu pyrosequencing resulted in male calls for female samples, indicating that a small amount of male DNA was being detected that was presumably missed during the amelogenin assay. This is not necessarily surprising as shirts were worn for an entire day without restriction, and some male DNA could easily have been deposited on them. In the few instances where DYZ1/Alu pyrosequencing resulted in female calls of male touch DNA, amelogenin produced no results at all, thus it seems likely that DNA levels were so low that dropout of DYZ1 occurred (discussed in more detail below). For hair shaft DNAs, the pyrosequencing assay produced more negative results (no sex call) than did real-time, however it should be noted that the pyrosequencing assay was developed and utilized well after the real-time assay, thus the hair DNAs had been stored years longer and were likely even more degraded by the time the pyrosequencing assay was applied to them.

The strength of the two assays described here—highly increased sensitivity even on low quality samples—is clear, but the limitations of either assay are also worth noting. First among these is the effect of stochastic sampling. Problems resulting from stochastic sampling are well known when working with low copy DNA testing in a forensic setting, typically being considered worrisome at levels below approximately 100 pg of input human DNA (e.g. [19]). However, with thousands of DYZ1 and Alu copies present in a single human cell, stochastic sampling does not manifest itself until much lower DNA levels are encountered, a clear advantage of the DYZ1/Alu assays. Nevertheless, this does not mean that stochastic levels cannot be reached, as they undoubtedly were in the current study. Such levels were probably realized in the drift or loss of the DYZ1 curve in some of the real-time assays, as well as when a male touch sample tested as female using the pyrosequencing assay. Even with the extreme sensitivity of the assays detailed here, stochastic sampling will continue to be an issue, albeit at orders of magnitude lower than with single copy markers.

Ambiguous sexing results can also originate from contamination, and the extreme sensitivity of the DYZ1/Alu assay means that contamination is a critical consideration. Extensive safeguards were incorporated to decrease the chance

of contamination, including both standard ones such as UV irradiating all possible supplies and reagents, and nonstandard ones such as filtering commercially supplied primers to remove contaminating human DNA. It was noted early on in this research that purchased primers were often contaminated with human DNA. These were likely at trace levels, since they passed manufacturer quality control standards, but the sensitivity of the DYZ1/Alu assays detected them in negative controls. In our hands, UV-irradiating primers and probes is sometimes possible while retaining a viable assay, but clearly it is not desirable, requiring an extra level of optimization so that the contaminating genomic DNA is destroyed while the primers/probes are not. To circumvent this, primers were passed through Microcon filters that trapped higher molecular weight DNA, and the filtrate was then used for PCR. It is possible that the entire PCR master mix (less the Taq) could be filtered, aliquoted, and stored in a similar manner, which would save time and effort, although this was not attempted in the current study. Regardless, these extremely sensitive assays mean that not only must great care be taken when utilizing them, but that results at the far end of their sensitivity (e. g., after 40 PCR cycles in the real-time assay) must always be viewed with extreme caution (e.g. [12]).

An interesting finding in this research was that real-time PCR results from amelogenin and DYZ1/Alu differed when testing high molecular weight or degraded DNA. Amelogenin acted as expected, in that it became less sensitive when DNA was artificially degraded using DNase I. In contrast, the DYZ1 portion of the high copy assay clearly, and counterintuitively, gained sensitivity following DNase treatment of high molecular weight DNA. The reason for this likely lies in the genomic distribution of Alu and DYZ1. The former is abundant and distributed among all human chromosomes, thus even if a small fraction of a cell's DNA is assayed, Alu members are likely to amplify well. On the other hand, DYZ1 exists as a linked family of repeats on the Y chromosome, and sampling less than a cell's worth of high molecular weight DNA means that those linked markers may or may not be present. However, when male DNA is degraded, DYZ1 members become functionally unlinked, and can be distributed throughout the DNA sample, resulting in their detection even when quantities of DNA below one human cell are assayed.

As new methods and strategies are introduced into the forensic sciences, the strengths and weaknesses of each must be examined. Strategies for extracting ever more information from smaller and smaller samples are a boon to the forensic sciences, but as can be seen in the research presented here, they also necessitate some caveats. With regard to genetic methods for identification, today's state-of-the-art technique is next generation (Next-Gen) sequencing, wherein whole genomes can be deciphered in a single experiment, at a cost of several hundred to a few thousand dollars. The amount of data recovered is immense, and has been used to determine the sex and other traits of ancient remains (e.g. [20, 21]), in which DNA is also highly degraded. However, Next-Gen sequencing requires new and costly equipment, software, and expertise, which may not currently be available to crime laboratories, and the glut of data produced is likely beyond the needs of forensic laboratories outside of very specific instances (such as differentiating identical twins). The two assays described here were designed to answer a specific question quickly and easily—the sex of a sample—and both do so for a few dollars per sample. Further, quantitative PCR is already well established in crime laboratories, so the real-time DYZ1/Alu assay could be easily implemented, especially until Next-Gen sequencing or alternatives are available or deemed beneficial to the forensic community.

Overall, the results presented show that both the real-time and pyrosequencing DYZ1/Alu assays are orders of magnitude more sensitive, and have higher true positive and true negative rates, as well as positive and negative predictive values, than does the assay generally used by forensic laboratories. Samples that gave no results or incorrect results based on amelogenin produced DYZ1/Alu results with a very high level of accuracy. While the extreme sensitivity of the DYZ1/Alu assays may create challenges, by taking the aforementioned precautions, these assays are excellent methods for sexing the most challenging samples. Ultimately, either of these assays can assist in determining the sex of an individual for forensic investigations when standard sex determining methods are insufficient.

The authors would like to thank Carrie Jackson for her extensive efforts in the early parts of this project, and Liz Kopitke and Sarah Rambadt, for providing many of the samples used in this research. Portions of this project were supported by the National Institute of Justice, Office of Justice Programs, U. S. Department of Justice award no. 2008-DN-BX-K216, via the American Academy of Forensic Sciences Forensic Science Foundation, awarded to A. F. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect those of the Department of Justice.

The authors have declared no conflict of interest.

5 References

[1] Andréasson, H., Allen, M., *J. Forensic Sci.* 2003, *48*, 1280–1287.

[2] Nakahori, Y., Takenaka, O., Nakagome, Y., *Genomics* 1991, *9*, 264–269.

[3] Maciejewska, A., Pawłowski, R., *Genetics* 2009, *3*, 265–267.

[4] Santos, R. F., Pandya, A., Tyler-Smith, C., *Nat. Genet.* 1998, *18*, 103.

[5] Roewer, L., Arnemann, J., Spurr, N. K., Grzeschik, K. H., Epplen, J. T., *Hum. Genet.* 1992, *89*, 389–394.

[6] Rahman, M. M., Bashamboo, A., Prasad, A., Pathak, D., Ali, S., *DNA Cell Biol.* 2004, *23*, 561–571.

[7] Schmid, M., Gettenback, M., Nanda, I., Studer, R., Epplen, J. T., *Genomics* 1990, *6*, 212–218.

[8] Tyler-Smith C., Taylor L., Müller U., *J. Mol. Biol.* 1988, *203*, 837–848.

[9] Honda, K., Harihara, S., Nakamura, T., Hirai, M., Misawa, S., *Japanese J. Legal Med.* 1990, *44*, 293–301.

[10] Hummel, S., Herrmann, B., in: Herrmann, B., Hummell, S. (Eds.), *Y-chromosomal DNA From Ancient Bones*, Springer-Verlag, New York 1994, pp. 205–210.

[11] Mighell, A. J., Markham, A. F., Robinson, P. A., *FEBS Lett.* 1997, *417*, 1–5.

[12] Nicklas, J. A., Buel, E., *J. Forensic Sci.* 2006, *51*, 1005–1015.

[13] Sulivan, K. M., Manucci, A., Kimpton, C. P., Gill, P., *BioTechniques* 1993, *15*, 636–641.

[14] Nicklas, M. A., Buel, E., *J. Forensic Sci.* 2005, *50*, 1081–1090.

[15] Jackson, C. B., *Thesis for Degree of M. S. School of Criminal Justice*, Michigan State University, MI, USA 2005.

[16] Thomasma, S. M., *Thesis for the Degree of M. S. School of Criminal Justice*, Michigan State University, MI, USA 2012.

[17] Graffy, E. A., Foran, D. R., *J. Forensic Sci.* 2005, *50*, 1110–1122.

[18] Kramer, M. S., *Clinical Epidemeology and Biostatistics*, Springer-Verlag, New York 1988.

[19] Gill, P., Whitaker, J., Flaxman, C., Brown, N., Buckleton, J., *Forensic Sci. Int.* 2000, *112*, 17–40.

[20] Green, R. E., Krause J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M. H-Y., Hansen, N. F., Durand, E. Y., Malaspinas, A-S., Jensen, J. D., Marques-Bonet, T., Alkan, C., Prüfer, K., Meyer, M., Burbano, H. A., Good, J. M., Schultz, R., Aximu-Petri, A., Butthof, A., Höber, B., Höffner, B., Siegemund, M., Weihmann, A., Nusbaum, C., Lander, E. S., Russ, C., Novod, N., Affourtit, J., Egholm, M., Verna, C., Rudan, P., Brajkovic, D., Kucan, Z., Gusic, I., Doronichev, V. B., Golovanova, L. V., Lalueza-Fox, C., de la Rasilla, M., Fortea, J., Rosas, A., Schmitz, R. W., Johnson, P. L. F., Eichler, E. E., Falush, D., Birney, E., Mullikin, J. C., Slatkin, M., Nielsen, R., Kelso, J., Lachmann, M., Reich, D., Pääbo, S., *Science* 2010, *328*, 710–722.

[21] Skoglund, P., Stora, J., Gotherstrom, A., Jaobsson, M., *J. Archaeol Sci.* 2013, *40*, 4477–4482.

