Testing Candidate DNA Quantitation Standards with Several Real-Time Quantitative PCR Methods*

Margaret C. Kline, Peter M. Vallone, Amy E. Decker, Janette W. Redman,
David L. Duewer, and John M. Butler
Chemical Science and Technology Laboratory,
National Institute of Standards and Technology, Gaithersburg MD 20899, USA

Sources of Support: National Institute of Justice (NIJ) funded this work in part through an interagency agreement 2003-IJ-R-029 with the NIST Office of Law Enforcement Standards.

Abstract:

Numerous real-time quantitative PCR (qPCR) methods have been developed in the last several years for use with forensic DNA samples. Ten different qPCR methods were used to evaluate DNA samples distributed in the NIST Interlaboratory DNA Quantitation Study 2004 (QS04) [1]. The target DNA concentrations of the QS04 samples were from 1.5 ng/µL to 50 pg/µL. About one-fifth of all QS04 results came from qPCR methods. These data show differences among the qPCR methods, both with regard to precision and bias. It is unclear from these data whether the observed differences are inherent to the methods or reflect differences in the standards used in their calibration. We here present our evaluation of several qPCR methods using six different human DNA calibration materials. All of the qPCR methods evaluated are either commercially available or have been published recently. Three of the calibration materials used in the evaluation are commercially available; three are derived from in-house purified single-donor blood samples. This study was designed to help direct development of candidate Standard Reference Material 2372, Human Genomic DNA Quantitation Standard.

Introduction:

The National Institute of Standards and Technology (NIST) is the National Metrology Institute for the United States of America. As such, NIST is responsible for the production of Standard Reference Materials (SRMs) for many different measurement systems. SRMs allow laboratories to establish traceability of a measurement through NIST to internationally recognized scales and units.

Many questions must be answered before production of an SRM can begin. First, is the SRM truly needed? The need for a particular SRM is generally established from the emphatic requests of a measurement community, although regulatory agencies may request SRMs. Once the need for a specific SRM is established, a material that is fit for the intended purpose must be identified that can be produced in a manner that assures homogeneity, stability, recoverability, and traceability. Values assigned to SRM components are traceable to the designated certification method.

The Human Identity Project group at NIST has been involved with the forensic DNA measurement community since 1990. Over the subsequent years, members of this community have often requested that NIST provide a DNA quantitation standard. Although we have been exploring production and delivery methods for some time via NIST-sponsored interlaboratory challenge studies [1-3] , until recently, we were unable to establish production and certification protocols that could produce an SRM fit for this purpose.

In DNA Quantitation Study 2004 (QS04), laboratories evaluated 8 different samples using qPCR methods. Figure 1 displays composite results from this study that suggest qPCR methods differ both with regard to precision and bias. However, it is unclear from these data whether the observed differences are inherent to the methods or reflect differences in the standards used in their calibration.

To better understand the qPCR processes and determine how to best design a DNA quantitation standard suitable for use with qPCR techniques, we have evaluated several different qPCR methods using six different human DNA calibration materials. We have also evaluated the stability of genomic DNA in five different tubes claimed as suitable for delivery and storage of extracted DNA, and explored the impact of four different extraction methods on DNA quantification, and the utility of quantifying DNA from 15 year old bloodstains stored at ambient temperature.

Materials and Methods

DNA samples for qPCR methods evaluation

Three human genomic DNA standards were obtained from two commercial suppliers (Applied Biosystems, Foster City, CA and Promega Corp., Madison, WI). The DNA concentration ([DNA]) of these commercial materials was used as assigned by their vendor. Three single-source samples were purified at NIST from freshly obtained blood collected in EDTA blood tubes, using a modified salting out procedure [4]. The extracted samples were analyzed by UV absorbance, scanned from 320 nm to 230 nm, using a Cary UV/Visible Spectrophotometer (Varian, Walnut Creek, CA). The absorbance260 nm at was used to assign the [DNA] in ng/µL for the extracted samples based on the absorbance of 1 being equivalent to 50 ng/µL of double-stranded DNA in TE buffer [5]. We designate these standards as "S1" to "S6".

qPCR methods

The [DNA] of all six of the DNA standards was evaluated using the following five methods: Quantifiler Human DNA Quantification Kit (Applied Biosystems) [6], Quantifiler Y Human Male Quantification Kit (Applied Biosystems) [7], an Alu-based assay [8], the Centre of Forensic Sciences (CFS) assay [9], and the California Department of Justice (CA DOJ) assay [10]. All PCR primers needed for the published assays were purchased from Qiagen Operon (Alameda, CA). The CFS assay developed a TaqMan® probe labeled with a VIC reporter dye and Minor Groove Binder (MGB) quencher dye (Applied Biosystems). The nuclear DNA (nDNA) primer and probe sequences for the CA DOJ assay were used as described in Ref 10; these target the TH01 gene on chromosome 11p15.5. The nuclear probe for this assay is a FAM labeled TaqMan® Black Hole Quencher (BHQ) purchased from Qiagen Operon.

Working solutions of each of the six standard DNAs were prepared fresh daily, by serially diluting the stock to the following concentrations of ng/µL in sterile deionized water: 10, 4.0, 1.6, 0.64, 0.26, 0.10, and 0.04. Two microliters of each dilution was added per reaction for all assays, with duplicate reactions for each of the six quantification standards. All assay reactions were carried out in 20 µL volumes and dispensed into a 96-well reaction plate. Each assay plate was prepared and run in duplicate on the ABI Prism 7500 Real-Time PCR System (Applied Biosystems). Information regarding qPCR cycling conditions and reagents is listed in Table 1. All data were analyzed using 7500 System SDS Software (Applied Biosystems). Standard curve information and quantification results for all assays were generated using the parameters provided in Table 2. The parameters listed for the CA DOJ assay are those described in the CA DOJ DNA Laboratory protocol.

DNA bloodstain extraction and analysis methods

Four sets of 10 mm^2 bloodstains from four individuals, prepared in 1995 and stored at –20 ºC on S&S 903 paper as described in [11] were processed along with four sets of 10 mm^2 bloodstains freshly prepared on S&S 903 paper from two individuals. The extraction methods included: Chelex 100 [12], DNA IQ, (Promega Corp), organic extraction [13], and inorganic "salt-out" extraction [4]. Extracts were quantified using the Quantifiler Human DNA quantification kit and diluted to a nominal [DNA] of 1 ng/µL. Dilutions were amplified with AmpF*l*STR Profiler Plus (Applied Biosystems). PCR products were analyzed with an ABI 3100 (Applied Biosystems). Additionally, 51 DNA chelex extracts from 15-year-old bloodstains

stored at ambient temperature [14] were amplified with AmpF*l*STR Identifiler PCR Amplification Kit (Applied Biosystems).

STR profiles of genomic DNA standards were generated using the AmpF*l*STR Identifiler PCR Amplification kit. Amplification was performed according to the manufacturer instructions with varying amounts of DNA template. A sensitivity study of the genomic DNAs included a seven-fold serial dilution ranging from 2 ng/µL to 31 pg/µL. PCR products were separated and detected using the ABI Prism® 3100 Genetic Analyzer (Applied Biosystems) following the manufacturer's standard protocol. Data was analyzed with GeneScan® v3.7 and Genotyper® v3.7 software (Applied Biosystems).

Tube Study: recovery/stability of extracted DNA

Five different DNA storage tubes were used in the DNA recovery/stability study. The tubes were obtained from: Savillex Corp (Minnetonka, MN), Axygen Scientific (Union City CA), Sarsedt Inc. (Newton, NC), and Simport Plastics Ltd (Beloeil, Canada). The different tube types were labeled "A" to "E". A stock solution of a 10 ng/µL DNA extract was diluted to 5.0 ng/µL, 1.0 ng/µL, and 0.2 ng/µL. The stock DNA solution and the dilutions were stored in Teflon containers. The three diluted DNAs were distributed as 100 µL aliquots into six tubes of each type. Two sets of each type at each dilution were stored at laboratory ambient temperature, 4 ºC, and –80 ºC. The stock solutions of DNA were evaluated using Alu-based assay [8] on day one. The dilutions held at each temperature were evaluated one set per quantitation plate in duplicate, two plates per day. Analysis time points were one day, one week, one month, four months, and seven months.

Results and Discussion

Interlaboratory performance-reproducibility

Based on the results reported by all participants in the NIST QS04 challenge study, the forensic community as a whole quantifies DNA with an expected one-standard deviation variation of a factor of 1.7 for both single- and multiple-donor samples. Against expectations, the results for ten different qPCR methods used in QS04 appeared to be biased by at least this amount relative to each other. We evaluated five of the qPCR methods used in QS04 to sort out whether the observed biases are method- or standard-based. Figure 2 displays the observed crossing threshold (C_Ts) of the different samples at a single [DNA] in the Quantifiler Human assay. The C_Ts of samples S2 and S6 differ by 1.1, or a factor of 2.2. There is also a grouping of four samples whose C_Ts are within 0.5 of each other, or within a factor of 1.4.

Figure 3 takes the same [DNA] and plots the absolute C_Ts within each method. All methods had a trend for the same four samples being more similar than the other two samples tested. There appears to be slight differences in relative sample performance that are consistent among the methods.

Figure 4 determines how the consistency of a standard used in a qPCR method would affect quantitation results. The modeled "calibration curves" were generated based on the assumption that at 100% PCR efficiency there is a doubling of the PCR product. Arbitrary C_Ts were used to plot the "nominal" truth standard curve. Then ± 0.3 C_Ts were plotted based on the "truth" curve. Depending on the relative position on the calibration curve the effect of variability changes. A nominal [DNA] of 1.8 ng/µL would range from 1.4 ng/µL to 2.2 ng/µL. If however the nominal [DNA] is 14.3 ng/µL the range is now 11.2 ng/µL to 18.0 ng/µL.

Tube study

There was no obvious difference between storage temperatures, Table 3 lists the data as the difference of the average C_T for a given [DNA] across the three storage temperatures for each tube type with data taken at 4 time points across 7 months. The data have been normalized to the Teflon tube in log ^{2} space. The more concentrated the DNA solution, in general, the better recovery of the material from any tube type.

Extraction methods

The peak heights seen from analyzing PCR products that had "1 ng" input DNA varied depending on the extraction method used. The peak heights of the inorganic and organic extracts were similar. Peak heights from the DNA IQ extracts were approximately 40% of the inorganic, organic extracts, while the Chelex extracts were 60% of the inorganic, organic extracts, Figure 5. In Figure 6 Internal PCR Control (IPC) of these extracts would indicate that there was some inhibition associated with Chelex and DNA IQ that could explain peak differences. The Chelex extracts were more concentrated than the DNA IQ extracts and therefore had to be diluted more.

Aged stains

The IPC's of the aged-bloodstain Chelex extracts do not correlate with the resulting STR peak heights, or completeness of the typing results obtained. Figures 7 and 8 compare the uppermost and lowermost bounds of the IPC seen with the resulting Identifiler profiles. While uppermost samples should be the most "inhibited", those samples had more complete profiles than the lowermost IPC samples that would be expected to have the least amount of inhibition. The "1 ng" of template, which produced a PCR product with an IPC in the middle of the calibration curve IPCs, only had 3 peaks above a 150 rfu threshold (Figure 9). Also shown in Figure 9 is the complete profile of the same sample amplified at "5 ng"

Conclusions

Laboratories working with new methodologies such as qPCR typically run their own internal validation studies to establish the "ng" quantity that produces a PCR product that produces an acceptable range of signal with their instrumentation. The results from this study may assist those laboratories still "looking" at qPCR methods to the differences seen with sample extraction and age studies presented here. The qPCR methods appear to yield similar results when the "same" standards are used. Problems could arise if a method is validated with a – 0.3 C_Ts calibration material and then a +0.3 C_Ts calibration material is used. Whatever calibration standard is used, it must be stable over time in order to provide consistent results.

Requirements for NIST SRM 2372 Human DNA Quantitation Standard include that it must be, homogenous, stable, will withstand shipping and normal storage, recoverable in terms of the [DNA] that went in the tube is the same that the end-user recovers. The values assigned are traceable to the designated certification method. The SRM with these characteristics can then be used to transfer traceability of the end-users measurements. All the above studies have influenced the decisions made in the production of a human DNA quantitation standard.

Table 1: Assay Conditions for Real-Time qPCR Methods

Assay	Kit Master Mix	Chemistry	Final Volume (ul)	Cycle #	50°C UNG activation	95°C activation	95°C/ 60°C annealing	72°C 3-step
Quantifiler Human	Included[1]	TaqMan® MGB	20	40	-	10 min	15 s/1 min	-
Quantifiler Y Male	included[1]	TaqMan® MGB	20	40	-	10 min	15 s/1 min	-
Alu	Invitrogen[2]	SYBR	20	30	2 min	2 min	15 s/ 35 s	30 s
CA DOJ nDNA	Universal no UNG[1]	TaqMan® BHQ	20	45	-	10 min	15 s/1 min	-
CFS	Universal[1]	TaqMan® MGB	20	40	2 min	10 min	15 s/1 min	-

[1] Applied Biosystems, Foster City,CA
[2] Platinum SYBR Green qPCR Super Mix UDG (Invitrogen Corp., Carlsbad, CA).

Table 2: Analysis Parameters using the ABI 7500 SDS Software

Assay	Data Collection	Threshold Value	Baseline Value (cycles)
Quantifiler Human	60°C	0.200	3-15
Quantifiler Y Male	60°C	0.200	3-15
Alu	68°C	0.001	3-4
CA DOJ nDNA	60°C	0.150	6-18
CFS	60°C	0.100	3-15

Table 3 Tube study

[DNA]	A	B	C	D	E
0.2 ng/µL	1.0	0.74	1.14	0.72	0.69
1.0 ng/µL	1.0	0.88	0.98	0.86	0.88
5.0 ng/µL	1.0	0.99	0.91	0.94	0.72

[DNA] is given in ng/µL

The data are presented as the difference of the average C_T for a given [DNA] across the three storage temperatures for each tube type with data taken at 4 time points across 7 months. The data have been normalized to the Teflon tube in log 2 space. That is the difference of the average C_T for a tube type at a [DNA]. The average C_T of the Teflon tube at the same [DNA] in $\log_2$ space.

References

1 Kline MC, Duewer DL, Redman JW, Butler JM. Results from the NIST 2004 DNA Quantitation Study. J Forensic Sci 2005;50(3):571-578.

2 Kline MC, Duewer DL, Redman JW, Butler JM. NIST Mixed Stain Study #3: DNA Quantification Practice and its Influence on Short Tandem Repeat Multiplex Performance. Anal Chem 2003;75(10):2463-2469.

3 Duewer DL, Kline MC, Redman JW, Newall PJ, Reeder DJ. NIST Mixed Stain Studies #1 and #2: Interlaboratory Comparison of DNA Quantification Practice and Short Tandem Repeat Multiplex Performance with Multiple-Source Samples. J Forensic Sci 2001;46(5):1199-1210.

4 Miller SA, Dykes DD, Polesky HF.A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1988; 16(3):1215.

5 Sambrook and Russell 3rd edition Molecular Cloning a Laboratory Manual ,2001, Cold Spring Harbor Laboratory Press, Cold Spring, NY. Vol 3, p A8.20.

6 Quantifiler™ Human DNA Quantification Kit PN4343895

7 Quantifiler™ Y Human Male Quantification Kit PN4343906

8 Nicklas J, Buel E. Development of an Alu-based, Real-Time PCR Method for Quantitation of Human DNA in Forensic Samples. J Forensic Sci 2003; 48:936-944.

9 Richard ML, Frappier RH, Newman JC. Developmental Validation of a Real-Time Quantitative PCR Assay for Automated Quantification of Human DNA. J Forensic Sci 2003;48:1041-1046.

10 Timken M, Swango K, Orrego C, Buoncristiani M. A Duplex Real-Time qPCR Assay for the Quantification of Human Nuclear and Mitochondrial DNA in Forensic Samples: Implications for Quantifying DNA in Degraded Samples. J Forensic Sci, September 2005;50:1044-1060.

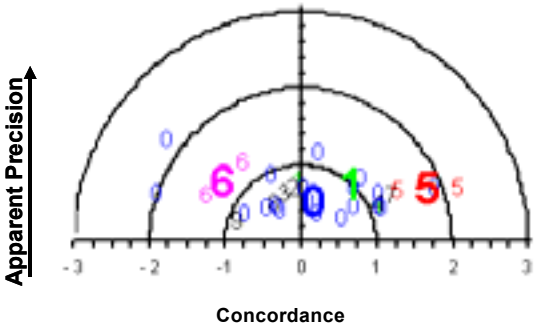
11 Kline MC, Duewer DL, Redman JW, Butler JM. PCR Amplification of DNA from Aged Blood Stains: Quantitative Evaluation of the "Suitability for Purpose" of Four Filter-Papers as Archival Media. Anal Chem 2002;74(8):1863-1869.

12 Walsh PS, Metzger DA, Higuchi R. Chelex 100 as a Medium for Simple Extraction of DNA for PCR-Based Typing from Forensic Material. 1991; BioTechniques, 10(4), 506-513.

13 Maniatis T, Fritsch EF, Sambrook J. 1989; Molecular Cloning: A Laboratory Manual, 2nd Edition, Vol. 2. Cold Spring Harbor Press pp. E10 – E14

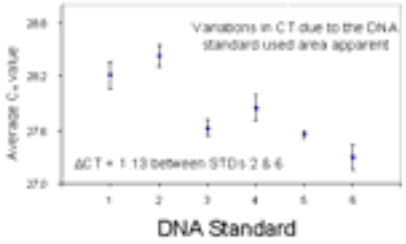
14 Kline MC, Redman JW, Parker B, Peter J,.Duewer.DL, Butler, JM, 2001; Recovery of DNA from aged-bloodstains on un-treated paper. Poster at 12th International Symposium on Human Identification

Figure 1. Laboratory Performances with qPCR Methods from NIST QS04



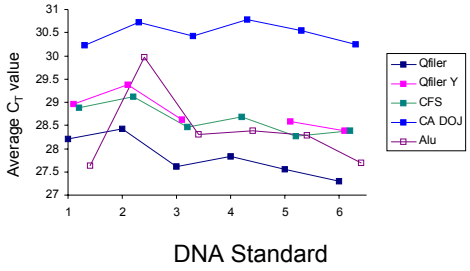
Comparing results from 8 different samples using 10 different qPCR methods, this "Target" plot summarizes among-laboratory measurement performance characteristics. Each small symbol represents a single set ([DNA] of eight samples) of quantitative among-laboratory results: concordance is displayed along the horizontal axis, apparent precision along the vertical axis, and total comparability is the distance from the target center. Method codes are: 0 = Quantifiler, 1 = Alu Q-PCR, 2 = Alu_Sifis, 3 = Alu_tqman, 4 = Aluprobe, 5 = BRCA 1, 6 = CFS-HUMRT, 7 = GB:L78833.1, 8 = RB1, 9 = RTALU. The large bold-face symbols represent the median performance of the among-laboratory results for methods reported by two-or-more laboratories. Three reference semi-circles are displayed: the inner-most semi-circle delimits a total comparability of one standard deviation from perfect agreement with the consensus medians for all samples, the middle delimits two standard deviations, and the outer delimits three standard deviations.

Figure 2 Comparison of Quantifiler Results for Samples S1 – S6 at Nominal 1.6 ng/µL [DNA]; n=4.



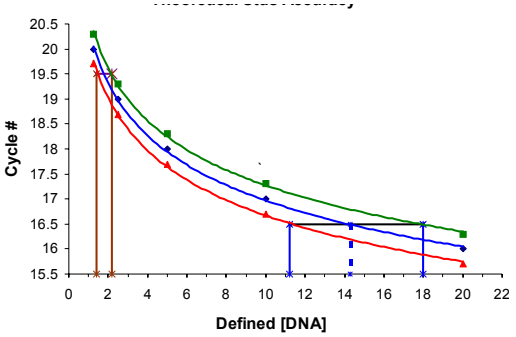
The observed crossing threshold (C_Ts) of the different samples at a single [DNA] in the Quantifiler Human assay. The C_Ts of samples S2 and S6 differ by 1.1, or a factor of 2.2. There is also a grouping of four samples whose C_Ts are within 0.5 of each other, or within a factor of 1.4. Error bars represent ± 1 SD.

Figure 3 The CT Trends Observed for Samples S1 – S6 for Five qPCR methods



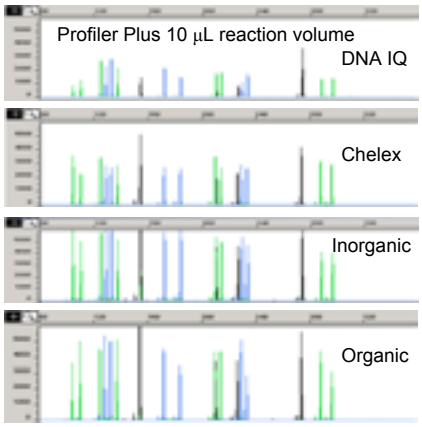
Absolute C_T values vary for each qPCR method. The same [DNA] is plotted as the absolute C_Ts within each method for a given sample. Absolute C_T will of course vary due to different analysis parameters but there appears to be little method based bias. All methods had a trend for the same four samples being more similar than the other two samples tested. There appears to be slight differences in relative sample performance that are consistent among the methods, indicating no or minimal qPCR method based bias.

Figure 4. Affect of a Change in Consistency of a Standard Used in a qPCR Method.



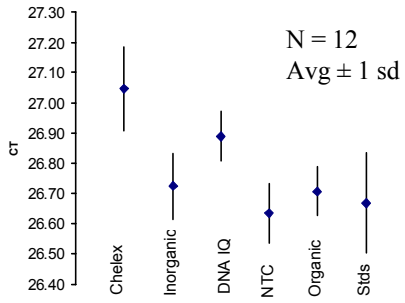
Theoretical standard curves generated based on the nominal and ±0.3 C_T (factor of 1.5). The center line is the "truth" curve generated based on the assumption that there is a two-fold difference in the [DNA] per cycle. The upper curve is + 0.3 C_Ts while the lower curve is – 0.3 C_Ts. The [DNA] of a sample would range from 1.4 ng/µL to 2.2 ng/µL at the lower concentrations while the upper end sample would range from 11.2 ng/µL to 18 ng/µL.

Figure 5. Electropherograms of "1 ng" template amplified.



Plots are fixed at the same RFU scale in all cases. Peaks heights from the DNA IQ extracts were approximately 40% of the inorganic, organic extracts, while the Chelex extracts were 60% of the inorganic, organic extracts

Figure 6. Average IPC ± 1 sd for the various extraction methods.



The average C_T of the IPC for 12 measurements per extraction method is plotted as the center point with the error bars representing ± 1 SD. The IPCs of these extracts would indicate that there was some inhibition associated with Chelex and DNA IQ that could explain peak differences. The chelex extracts were more concentrated than the DNA IQ extracts and therefore had to be diluted more. The DNA IQ extracts may have been inhibited dye to the relatively large size of the bloodstain extracted.

Figure 7. The Uppermost Bounds of the IPC seen with the Resulting Identifiler Profiles of Aged Bloodstain Extracts.

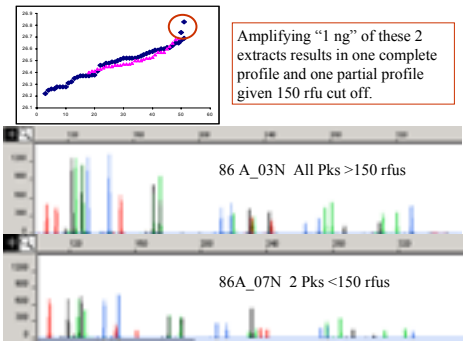


Figure 8. The Lowermost Bounds of the IPC seen with the Resulting Identifiler Profiles of Aged Bloodstain Extracts.

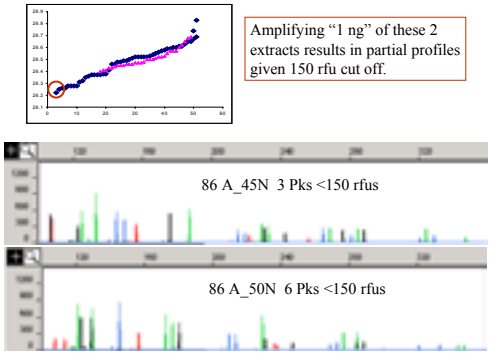
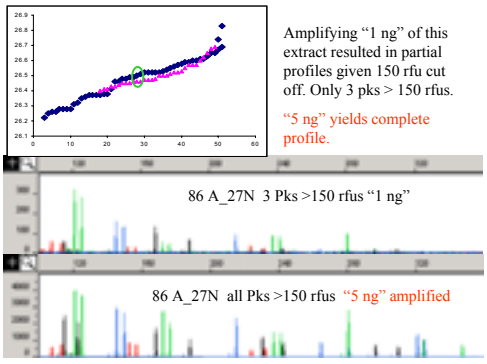


Figure 9. The Middle Areas of the IPC seen with the Resulting Identifiler Profiles of Aged Bloodstain Extracts.



Note: When "5 ng" of this sample is amplified a full profile is seen.