Anal. Chem. **2003,** *75,* 2463−2469

NIST Mixed Stain Study 3: DNA Quantitation Accuracy and Its Influence on Short Tandem Repeat Multiplex Signal Intensity

Margaret C. Kline, David L. Duewer,* Janette W. Redman, and John M. Butler

Chemical Science and Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland 20899

The Mixed Stain Study 3 (MSS3) interlaboratory challenge exercise evaluated the 2001 performance of STR multiplex DNA typing systems using a set of seven DNA extracts of designed concentration and composition. This initial report focuses on the linkages connecting the measurement of the concentration of DNA ([DNA]) to the observed STR multiplex signal intensities. There is a causal relationship between [DNA] measurement accuracy and the efficiency of STR multiplex analysis. There are no intrinsic measurement performance differences among the [DNA] measurement technologies reported. However, there are large differences in the efficiencies of amplification, separation, and detection among participants using the same nominal measurement systems.

Short tandem repeat (STR) multiplex assays are now the dominant forensic human identification technology.[1] Although multistep and chemically complex, current commercial STR multiplex assays provide results that are robust to typical laboratory preferences in sample preparation, polymerase chain reaction (PCR) equipment and protocols, and separation and visualization systems.[2−7] The National Institute of Standards and Technology (NIST) has coordinated a series of interlaboratory examinations of multiplexed STR systems.[4,8] In addition to documenting the evolution of STR assays and of the forensic community using them (see Table 1), these studies search for latent analytical difficulties by challenging analysts and assay systems with difficult samples presented in atypical contexts. No problem has been encountered that is intrinsic to properly performed STR multiplex analyses. However, the 1999 Mixed Stain Study 2 (MSS2) observed linkages between certain STR measurement anomalies and inaccurate DNA quantitation.[4]

The 2001 Mixed Stain Study 3 (MSS3) was designed to further explore the performance of STR multiplex systems and to resolve the DNA quantitation issues raised in the earlier interlaboratory challenges. Included in this report are the experimental details, the current state-of-practice for autosomal STR multiplexes, and evidence that DNA concentration ([DNA]) inaccuracy contributes to among-sample variability in STR multiplex signal intensity. Future reports will discuss other observations from the MSS3, including factors influencing allele identification, sources of DNA quantitation inaccuracy, the current state-of-practice for Y-chromosome STR multiplexes, and suggestions for the reporting and interpretation of mixed-source profiles.

MATERIALS AND METHODS

Participants. Participation in MSS3 was open to all human identity laboratories utilizing multiplex STR systems of five loci or more. Solicitation for participation was initiated in October 2000 and was completed in May 2001. In addition to directed invitations to institutions participating in previous NIST challenges, the study was advertised at the Scientific Working Group for DNA Analysis Methods (SWGDAM) meeting held during the October 2000 11th International Symposium on Human Identification, Biloxi, MS, and at the February 2001 53rd Annual Scientific Meeting of the American Academy of Forensic Sciences, Seattle, WA. Table 2 lists the 74 institutions that returned partial or complete results for the study. MSS3 was the first NIST interlaboratory challenge exercise for 38 of these participants.

Analyst experience with STR multiplex assays varied widely, from novice to expert. A number of participants reported using the MSS3 to help assess the utility of one or more multiplex systems not routinely used in their laboratories.

Samples. One control (labeled "R") and six study samples (labeled "S" to "X") were distributed in MSS3. These materials were prepared at NIST from DNA obtained from five female and seven male anonymous sources. Using standard procedures, DNA was extracted from commercially obtained whole blood.[9] After purification, the DNA was dissolved in Tris−EDTA (TE; 10

(1) Butler, J. M. *Forensic DNA Typing*; Academic Press: London, 2001.
(2) Kimpton, C. P.; Oldroyd, N. J.; Watson, S. K.; Frazier, R. R.; Johnson, P. E.; Millican, E. S.; Urquhart, A.; Sparkes, B. L.; Gill, P. *Electrophoresis* **1996**, *17* (8), 1283−93.
(3) Gill, P.; Sparkes, R.; Fereday, L.; Werrett, D. J. *Forensic Sci. Int.* **2000**, *108* (*1*), 1−29.
(4) Duewer, D. L.; Kline, M. C.; Redman, J. W.; Newall, P. J.; Reeder, D. J. J. *Forensic Sci.* **2001**, *46* (5), 1199−1210.
(5) Wallin, J. M.; Holt, C. L.; Lazaruk, K. D.; Nguyen, T. H.; Walsh, P. S. J. *Forensic Sci.* **2002**, *47* (1), 52−65.
(6) Holt, C. L.; Buoncristiani, M.; Wallin, J. M.; Nguyen, T.; Lazaruk, K. D.; Walsh, P. S. J. *Forensic Sci.* **2002**, *47* (1), 66−96.
(7) Krenke, B. E.; Tereba, A.; Anderson, S. J.; Buel, E.; Culhane, S.; Finis, C. J.; Tomsey, C. S.; Zachetti, J. M.; Masibay, A.; Rabbach, D. R.; Amiott, E. A.; Sprecher, C. J. J. *Forensic Sci.* **2002**, *47* (4), 773−785.
(8) Kline, M. C.; Duewer, D. L.; Newall, P.; Redman, J. W.; Reeder, D. J. J. *Forensic Sci.* **1997**, *42* (5), 897−906.
(9) Signer, E.; Kuenzle, C. C.; Thomann, P. E.; Hubscher, U. *Nucleic Acids Res.* **1988**, *16* (15), 7738.

10.1021/ac026410i CCC: $25.00
Published on Web 04/17/2003

Table 1: NIST-Sponsored STR Multiplex Interlaboratory Challenges

study	period[a]	labs[b]	sets[c]	plexity[d]	samples	focus
CTT	12/1995 5/1996	34/41	46	3.3	4 single-source extracts, 4 single-source stains	factors affecting sizing variability.
MSS1	4/1997 11/1997	22/28	37	5.9	6 single-source stains, 4 two-source stains, 1 three-source stain	donor types given a complete set of reference sources
MSS2	1/1999 5/1999	45/52	70	7.5	part A: 4 single-source stains, 1 two-source stain, 1 three-source stain	donor types given incomplete set of reference sources
					part B: 5 vials of a four-level concentration series	performance of DNA quantitation assays
MSS3	12/2000 10/2001	74/83	117	9.1	1 single-source extract, 5 two-source extracts, 1 three-source extract	effect of DNA quantitation on STR typing performance

[a] Time from start of sample distribution to last accepted data set. [b] Number of participants (i.e., laboratories returning at least one data set)/number of laboratories receiving samples. [c] Total number of STR multiplex analyses sets reported. [d] Average number of genetic loci assayed per independent data set.

mmol/L Tris−HCl, 0.1 mmol/L EDTA, pH 8.0) buffer. The purity and approximate total [DNA] of all 12 resulting master solutions were assayed using UV/visible absorbance spectrophotometry and yield gels.[10] The control and study materials were prepared by quantitative volumetric combination of the master solutions and TE buffer. We will present elsewhere our quantitative analyses of the master solutions and the control and study materials.

Control R is a single-source (male) material, S to W are two-source (male and female) materials, and X is a three-source (two males, one female) material. With the exception of samples T and V, no source was used in the preparation of more than one material. Samples T and V were prepared from the same two sources to have identical total [DNA], but with reciprocal female/male source composition ratios. The control material was designed to have a total [DNA] of 1 ng/μL; the samples were designed to have total [DNA] in the range of 1−4 ng/μL.

Ninety-six complete sets of MSS3 samples were prepared. Each set consisted of one 50-μL aliquot of each of the seven stock solutions. Each sample aliquot was packaged in a tightly sealed 0.75-mL BioStor vial (United Laboratory Plastics, St. Louis, MO). The fitness-for-purpose of these inert LN_2-grade polypropylene vials was confirmed prior to beginning sample shipment. All samples were stored at −80 °C until shipped. Samples were shipped on dry ice. Participants were asked to store all samples at −20 °C until analysis.

Protocol. The MSS3 consisted of two major activities: (1) quantifying the DNA (as ng/μL) in the control and study samples and (2) analyzing all of the samples using one or more STR multiplex. From the first activity, participants were asked to report the [DNA] in each sample and to specify the quantification protocol used. From the second activity, participants were asked to report the volume of each sample used in each PCR amplification, to report the type and intensity of all observed alleles in each sample, and to assign ,where possible, alleles to major and minor contributor sources. Participants were requested to analyze the control sample as the first and last sample in every set of analyses performed and to report the intensity of all alleles observed in each analysis. Participants were also requested to provide hardcopies of all gel image or electropherogram results. No sample handling, analysis, data analysis, or result reporting procedures or formats were specified. All results were required to be submitted to NIST no later than October 10, 2001, when preliminary results of the study were first discussed publically.[11]

All quantification, typing, and donor assignment data were transcribed at NIST from the participants' reports. Beginning in December 2001, participants were provided with a copy of their own values as recorded in a standardized format. They were requested to confirm the recorded values and correct any errors, oversights, and misinterpretations. Any clerical errors that could be confirmed as such against provided hardcopy were flagged and corrected. Participants that had not specified the PCR reaction volumes or sample loading parameters were requested to provide these values. Data confirmation was completed in June 2002.

Nonstandard Notation. All of the quantitative results are better represented as log-normal than as normal distributions. These data were, therefore, logarithmically transformed, $y' = \log(y)$, before analysis. Any given value of the log-transform data is easily inverse-transformed back to the original measurements, $y = 10^{y'}$. However, an additively symmetrical interval for the log-transformed data, $y' \pm ku' = y' - ku' \le y' \le y' + ku'$, becomes multiplicatively symmetrical when inverse-transformed back to the more familiar measurement domain, $y \times / u^k = (y/u^k) \le y \le y \times u^k$.[12] The notation "$\times /$" is employed as the multiplicative analogue of "±".

RESULTS AND DISCUSSION

[DNA] Measurements. Figure 1 is a "box and whisker" plot detailing the distributions of reported MSS3 control and test sample [DNA].[13] The result distributions are nearly identical for samples T and V, the two samples produced to have the same total [DNA] using the same two donors but at reciprocal female/male ratios. For these randomly chosen sources, none of the [DNA] measurement methods utilized by MSS3 participants are sensitive to the particular source of the DNA.

The among-participant variability in measuring [DNA] can be estimated from the average interquartile range of the individual distributions; that is, from the average height of the individual-boxes.[14] This robust estimate of the among-participant [DNA] standard deviation (SD), expressed as a multiplicative factor, is $\times /1.6$. Since the similarly defined estimate of among-participant [DNA] variation in MSS2 was $\times /$ 1.8,[4] the [DNA] measurement

(10) Maniatis, T.; Fritsch, E. F.; Sanbrook, J. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1982; Appendix A, pp 468−469.

(11) Kline, M. C.; Redman, J. W.; Duewer, D. L.; Butler, J. M. *Preliminary Results from the 2001 NIST Mixed-Stain Study 3: DNA Quantitation And Identification Of Minor Contributor(s)*. Poster 4, 12th International Symposium on Human Identification, Biloxi, MS, Oct. 10−12, 2001; http://www.promega.com/geneticidproc/ussymp12proc/abstracts/kline.pdf.

(12) Duewer, D. L. *J.* Chemometrics **1990**, *4*, 299−321.

(13) Tukey, J. W. *Exploratory Data Analysis*. Addison-Wesley: Reading, MA, 1977.

Table 2. MSS3 Participants

participant	location	participant	location
Genetic Technologies Corporation Pty. Ltd.	Fitzroy, Australia	Armed Forces DNA Identification Laboratory	Rockville, MD
The Centre of Forensic Sciences, Biology Section	Toronto, Canada	Baltimore County Police Department, Forensic Services Section	Towson, MD
Royal Canadian Mounted Police, Forensic Laboratory Services	Ottawa, Canada	Detroit Police Department, Forensic Services Division	Detroit, MI
Forensic Science Service, Department of Research & Development	Birmingham, England	Michigan State Police, Grand Rapids Laboratory	Grand Rapids, MI
Italian National Police, DCPC − Servizio Polizia Scientifica − Div. III	Rome, Italy	Bureau of Criminal Apprehension, Forensic Science Laboratory	St. Paul, MN
Institute of Environmental Science & Research Ltd., Forensic Biology	Auckland, New Zealand	Kansas City Missouri Police Department, Crime Laboratory	Kansas City, MO
University of Granada, Department of Legal Medicine	Granada, Spain	Charlotte Mecklenburg Police Department, Crime Laboratory	Charlotte, NC
Alaska Department of Public Safety, Scientific Crime Detection Laboratory	Anchorage, AK	Laboratory Corporation of America	RTP, NC
Alabama Department of Forensic Sciences, Mobile Laboratory	Mobile, AL	Albuquerque Police Department, Criminaliztics Laboratory	Albuquerque, NM
Arkansas State Crime Laboratory	Little Rock, AR	New Mexico Department of Public Safety	Santa Fe, NM
Arizona Department of Public Safety, Northern Regional Crime Laboratory	Flagstaff, AZ	Las Vegas Metropolitan Police Department − Forensic Lab	Las Vegas, NV
Mesa Police Department, Crime Laboratory	Mesa, AZ	Washoe County Sheriff's Office, Forensic Science Division	Reno, NV
Arizona Department of Public Safety, Central Regional Crime Laboratory	Phoenix, AZ	Suffolk County Crime Laboratory	Hauppauge, NY
Tucson Police Department, City−County Crime Laboratory	Tucson, AZ	New York City Office of Chief Medical Examiner, Forensic Biology	New York, NY
MiraiBio Inc. (Hitachi Genetic Systems)	Alameda, CA	Ohio Bureau of Criminal Identification & Investigation, DNA Laboratory	London, OH
Kern County Regional Crime Laboratory	Bakersfield, CA	Oklahoma State Bureau of Investigation, Criminaliztics Services Division	Oklahoma City, OK
Applied Biosystems, Technical Training Department	Foster City, CA	Oregon State Police Forensic Laboratory, DNA Analysis Unit	Portland, OR
Los Angeles Police Department, Scientific Investigation Division	Los Angeles, CA	Pennsylvania State Police DNA Laboratory	Greensburg, PA
San Diego County Sheriff's Office, Forensic Biology Section	San Diego, CA	City of Phoenix Police Department, Laboratory Services Bureau	Phoenix, AZ
Orange County Sheriff-Coroner Department, Forensic Science Services	Santa Ana, CA	Rhode Island Department of Health − Forensic Sciences	Providence, RI
Colorado Bureau of Investigation, Montrose Regional Facility	Montrose, CO	South Dakota State Forensic Laboratory	Pierre, SD
Connecticut Department of Public Safety, Forensic Science Laboratory	Meriden, CT	Tennessee Bureau of Investigation, Jackson Crime Laboratory	Jackson, TN
FBI Laboratory, DNA Analysis Unit I	Washington, DC	Tennessee Bureau of Investigation, Forensic Services Division	Nashville,TN
Miami-Dade Police Department, Crime Laboratory	Miami, FL	Austin Police Department, Forensic DNA Section	Austin, TX
Palm Beach County Sheriff's Office Crime Laboratory	West Palm Beach, FL	Texas Department of Public Safety, Headquarters Crime Laboratory	Austin, TX
Georgia Bureau of Investigation, Forensic Biology	Decatur, GA	Orchid Cellmark Dallas	Dallas, TX
United States Army, Criminal Investigation Laboratory	Fort Gillem, GA	Houston Police Department, Crime Laboratory	Houston, TX
Idaho State Police Forensic Services	Meridian, ID	Harris County Medical Examiner's Office, DNA Laboratory	Houston,TX
Indianapolis − Marion County Forensic Services Agency	Indianapolis, IN	Texas Department of Public Safety, Houston Crime Laboratory	Houston,TX
Sedgwick County, Regional Forensic Science Center	Wichita, KS	Texas Department of Public Safety, Lubbock Crime Laboratory	Lubbock,TX
Kentucky State Police, Forensic Biology Unit	Frankfort, KY	Bexar County, Criminal Investigation Laboratory	San Antonio, TX
North Louisiana Criminaliztics Laboratory	Shreveport, LA	Utah Bureau of Forensic Services	Salt Lake City, UT
Massachusetts Department of State Police Crime Laboratory	Sudbury, MA	Virginia Department of Criminal Justice Services, Central Laboratory	Richmond, VA
Baltimore City Police Department, Crime Laboratory	Baltimore, MD	Vermont Department of Public Safety, Forensic Laboratory	Waterbury, VT
National Institute of Standards and Technology, Biotechnology Division	Gaithersburg, MD	Washington State Patrol, Crime Laboratory Division	Seattle, WA
Maryland State Police Crime Laboratory	Pikesville, MD	Wisconsin State Crime Lab − Madison	Madison, WI

comparability among the forensic community appears to have improved from 1999 to 2001.

Figure 1 also displays the reported [DNA] values for participants who reported the three lowest and the three highest [DNA] values for control sample R. Although these six participants are not always the most discordant for every sample, they are very low or very high for nearly all samples. The between-sample bivariate correlations of the log([DNA]) among all participants are all fairly strong, ranging from a low of 0.45 to a high of 0.83.

Table 3 lists the various measurement methods and detection systems that participants used to estimate [DNA] in the MSS3 samples. Figure 2 summarizes their relative performance as

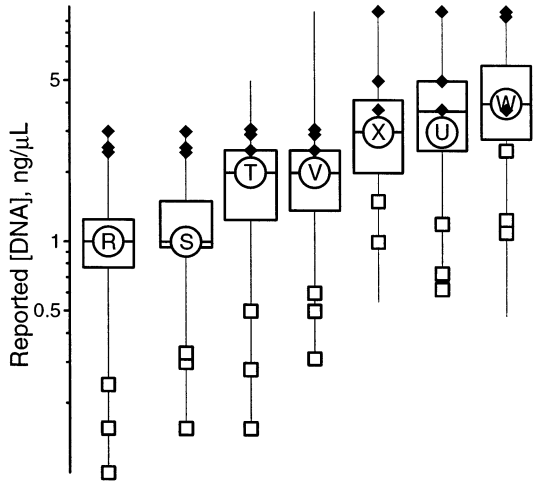

Figure 1. Quantitation results. The "box-and-whisker" structure for each sample displays the distribution of reported [DNA] values. Each box encloses the central 50% of the reported values, with the horizontal line within the box denoting the median value. The vertical lines connected to the box span from the largest to the smallest reported values. The open squares denote the reported [DNA] values for the three participants reporting the smallest [DNA] for sample R; the solid diamonds, likewise, denote the values for the three participants reporting the largest [DNA] for sample R. The labeled circles denote the design [DNA] for each sample.

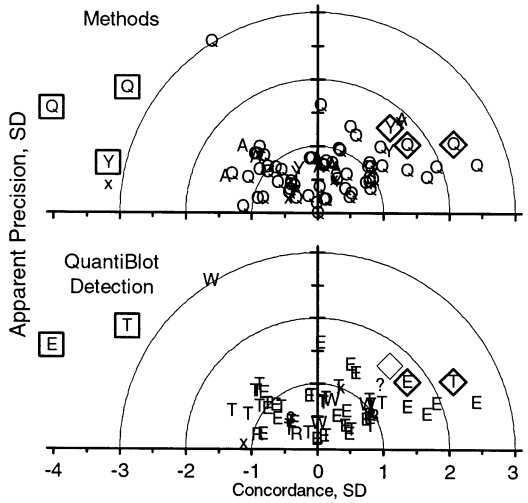

Figure 2. DNA quantitation measurements by method and detection systems. The upper target plot displays the measurement concordance (horizontal axis) and apparent precision (vertical axis) for all participants, with the different quantitation methods denoted with the codes lists in the second column of Table 3. The inner ring of the target encloses a combined concordance and apparent precision of a 1 standard deviation (SD) factor about the consensus medians, the middle ring encloses a 2 SD factor, and the outer ring encloses a 3 SD factor. Approximately 95% of all participants with measurement characteristics qualitatively similar to the consensus should plot within the middle ring. The open squares denote measurement characteristics of the three participants reporting the smallest [DNA] for sample R; the open diamonds, likewise, denote the characteristics for the three participants reporting the largest [DNA] for sample R. The lower target plot redisplays the measurement characteristics for all participants that used some form of Quantiblot assay, with the different visualization methods denoted with the codes listed in the fourth column of Table 3.

Table 3. Quantitation Systems

system	code	detection	code	sets
ABI 5700 Sequence Detection	x	Taqman probe		1
Aces 2.0	A	Lumi-Phos Plus chemiluminescence		8
AluQuant	x	Bioluminescence		1
Picogreen	x	Fluorescence		1
QuantiBlot	Q	Chemiluminescence, ECL	E	25
		Chemiluminescence, North2South	x	1
		Chemiluminescence, Super Signal	x	1
		Chemiluminescence, West Dura	W	3
		Chemiluminescence, West Femto	x	1
		Chemiluminescence, Renaissance Plus	R	4
		Colorimetric, TMB	T	27
		unspecified	?	2
Yield gel	Y	ethidium bromide		9
		total		84

"target" plots.[15] The distributions of both methods and results are similar to those observed in MSS2.[4] While the results from a few individual participants differ significantly from the consensus values, none of the methods used by multiple participants is consistently less concordant or precise than any other.[16] The same is true for the different detection systems used with QuantiBlot assays. While one of the single-participant methods produced consistently quite low results, individual QuantiBlot and yield gel results were similarly discordant. We will separately report investigations into the origins of the observed within-method variation in measurement performance.

Quantity of Sample Amplified. Figure 3 displays the inverse proportionality between measured [DNA] and the volume of each sample participants elected to amplify per STR multiplex reaction. The quantity of DNA that participants intended to amplify is given by the product, (estimated [DNA] in ng/μL) $\times$ (volume sample amplified in μL). On average, participants intended to amplify 1.25 ng of control R and 1.45 ng of all test samples. However, the average amount targeted varied considerably among the participants, ranging from 0.5 ng to nearly 7 ng. The among-participant SD in the intended amplification quantity, expressed as a multiplicative factor, is $\times$ 1.7 for both control and test samples.

The upper segment of Figure 4 uses the "target plot" format to display the among-participant distribution of the intended amplification quantities. Each symbol denotes both the average relative quantity targeted and the relative uniformity of that target quantity across the seven samples for a given participant. The clustering of the symbols along the bottom (perfect uniformity) of the plot indicates that most participants intended to amplify about the same quantity of DNA for all seven samples. The few participants who targeted different amplification quantities for the different samples often reported reamplifying one or more of the samples after evaluating their initial results. A few of these

(14) Analytical Methods Committee, Royal Society of Chemistry. *Analyst* **1989**, *114*, 1693−1697.
(15) Duewer, D. L.; Kline, M. C.; Sharpless, K. E.; Brown Thomas, J.; Gary, K. T.; Sowell, A. L. *Anal.Chem.* **1999**, *71* (9), 1870−1878.
(16) Duewer, D.L.; Lalonde, S. A.; Aubin, R. A.; Fourney, R. M.; Reeder, D. J. *J. Forensic Sci.* **1998**, *43* (3), 465−71.

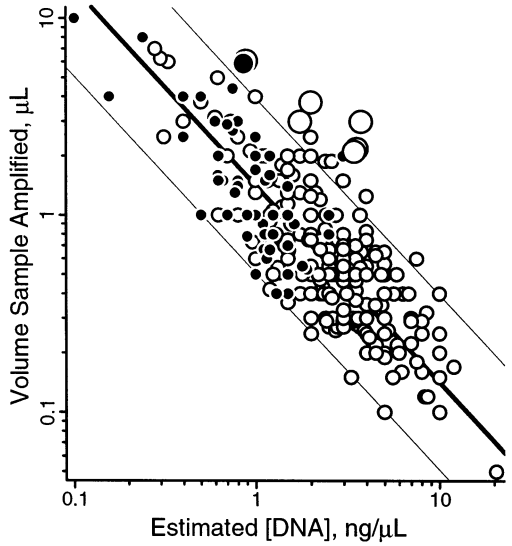

Figure 3. Sample volume amplified as a function of DNA quantification results. Each solid circle denotes the volume of the single-source sample R amplified in a given PCR multiplex as a function of the participant's estimated [DNA]. Likewise, each open circle denotes the volume of one of the multiple-source samples S−X amplified as a function of the participant's estimated [DNA] for the particular sample. The larger circles mark all values reported by one participant who selected to amplify unusually large amounts of DNA. The dark line denotes the median (1.3 ng) of DNA targeted for amplification. The light lines denote the 95% confidence interval about the median (the robust SD, expressed as a multiplicative factor, is × 1.7).

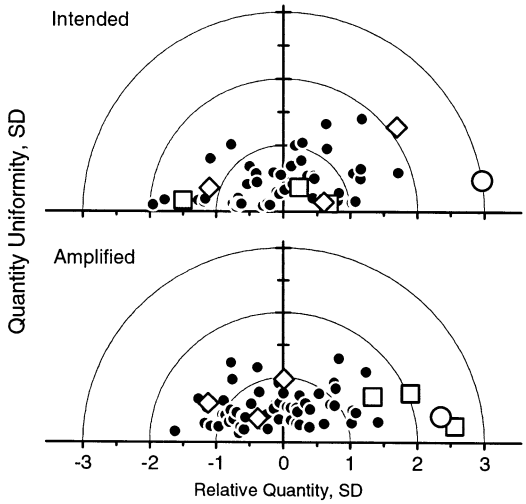

Figure 4. Intended and actual quantities of sample DNA amplified. The upper target plot characterizes the quantities of template DNA participants intended to amplify; the lower plot characterizes the quantities they actually amplified. The "Relative Quantity" and "Relative Uniformity" axes are analogous to the "Concordance" and "Apparent Precision" of Figure 2. The open circle marks the participant who chose to amplify unusually large quantities of template DNA. See the legend of Figure 2 for a description of the other graphical elements.

participants specifically noted that it is their institution's policy to reamplify samples until the STR multiplex analytical signal intensity falls within a specified range.

The lower segment of Figure 4 uses the same graphical format to display the among-participant distribution of the actual amounts of DNA amplified, (nominal [DNA] in ng/μL) × (volume sample amplified in μL). Since the sample volume amplified is inversely proportional to the measured [DNA], participants reporting high [DNA] tended to actually amplify small quantities of template DNA and those reporting low [DNA] results tended to actually amplify large quantities.

Two of the three participants who reported the highest [DNA] for sample R (see Figure 1) intended to amplify larger than average quantities—in one case, after evaluating the products of their initial PCR reactions—and (eventually) amplified fairly average quantities. The other of the three chose a quite low amplification target; they achieved this goal, although a few other participants amplified somewhat smaller quantities.

All three of the participants who reported the lowest [DNA] for sample R (see Figure 1), in fact, amplified very large quantities of DNA for all samples. The participant who reported the very lowest [DNA] for sample R actually amplified more template than the participant using the largest intended target quantity (Figure 3).

Allele Intensity Measurements. Participants in MSS3 were requested to report both the identity and the associated signal intensity for all STR alleles. Most participants who reported identities for one or more STR analysis also provided intensities, either tabulated or as part of their detailed hardcopy documentation. A few participants reported only allele identities or provided hardcopy that did not explicitly state the signal intensities. Although several participants reported allele intensities as both height and area, the vast majority reported only heights. Allele heights were reported in optical density or relative fluorescence units (RFUs).

Table 4 lists the STR multiplexes used in MSS3. Many participants amplified and analyzed one or more of the samples more than once, generally using somewhat different experimental conditions. Many participants analyzed samples with two or more multiplexes, generally in order to identify the allelic types at all

Table 4. STR Multiplexes

multiplex	code	no. loci total	no. loci CODIS[b]	sets[c]
AmpF/STR COfiler	Cof	7	6	29
AmpF/STR Profiler	Pro	10	9	1
AmpF/STR Profiler Plus	Pro+	10	9	35
AmpF/STR Identifiler	Idf	16	13	1
AmpF/STR SGM Plus	SGM+	11	8	3
PowerPlex 1.1	PP1.1	8	8	1
PowerPlex 1.1 + Amelogenin	PP1.1+	9	8	6
PowerPlex 2.1	PP2.1	9	8	4
PowerPlex 16	PP16	16	13	5
FSS Quad	Quad	4	2	1
combined: Cof and Pro+[a]	Cof, Pro+	14	13	28
combined: PP1.1 and PP2.1[a]	PP1.1, 2.1	14	13	1
combined: Cof, Pro+, Idf, SGM+, and PP16[a]	FM	18	13	1
			total	115

[a] One set of allele types reported per unique locus. [b] Number of CODIS core loci included in multiplex [c] Number of complete seven-sample sets of typing results reported. The number of individual one-sample analyses is 7 × sets.

Table 5. Electrophoretic Instrumentation

instrument	description	users
ABI 310	single-column capillary	52
ABI 377	slab gel	12
ABI 3100	16-column capillary array	1
Hitachi FMBio	slab gel	7
	total	72

13 of the U.S. Federal Bureau of Investigation's CODIS core loci.[17]

Table 5 lists the electrophoretic separation systems used in the MSS3. Most participants used only one system. The few using two or more types of separation systems generally analyzed different multiplexes on their different systems.

In principle, allele intensity should be a function of the amount of product DNA loaded into the slab gel or capillary column. In practice, there are at best modest relationships connecting the actual amount of template DNA amplified and the average signal per genetic locus[18] for any of the multiplex−instrument combinations used in MSS3. Regression models that included amplification reaction volume, product load volume, and (for capillary column systems) the injection duration, as well as template quantity, are little better related to the average signal than the template quantity alone (data not shown).

Although the average signal per ng DNA amplified by one participant does not well predict the signal observed by other participants, Figure 5 reveals that, for those participants who analyzed two different multiplexes under the same nominal amplification and separation conditions, the average signal for a sample in one multiplex system is predictive of the signal in the other. This suggests that most of the observed among-participant differences are intrinsic to the measurement systems involved and cannot be attributed to lot-to-lot variability in the STR multiplexes.

Further, for most participants, the average signal per ng DNA amplified of one sample does predict the signals for the other samples. The upper right "target plot" segment of Figure 6 displays the among-participant distribution of the average signal per locus for several STR multiplexes. Each symbol denotes both the average signal intensity and the relative uniformity of the signal intensities for a given participant. Although the distributions of relative intensities (the horizontal axis) for slab gel systems are broader then are those from capillary systems, the distributions of relative uniformities (the vertical axis) are about the same for all systems. This suggests that although there is greater among-participant operational diversity in slab gel then in capillary systems, the within-participant relationships between quantity amplified and multiplex signal intensity are equally consistent, regardless of system.

These suggestions are quantitatively explored in the scattergram segments of Figure 6. The lower-right segment compares the relative signal intensities with the relative quantities amplified (see Figure 4). The average quantities amplified are clearly correlated (≈0.5) with the average signal intensities only for the capillary systems. The left segment compares the among-sample uniformities in amount of sample amplified with those for signal intensity. These uniformities are correlated equally well (>0.6) in all systems. We infer from these relationships that (1) the absolute efficiencies of the overall STR multiplex measurement process (including the amplification, injection, separation, and detection subprocesses) can be quite variable among participants, even when the processes are nominally identical and (2) the relative efficiency of each participant's measurement process is quite stable, at least over the days-to-months required by the MSS3 study.

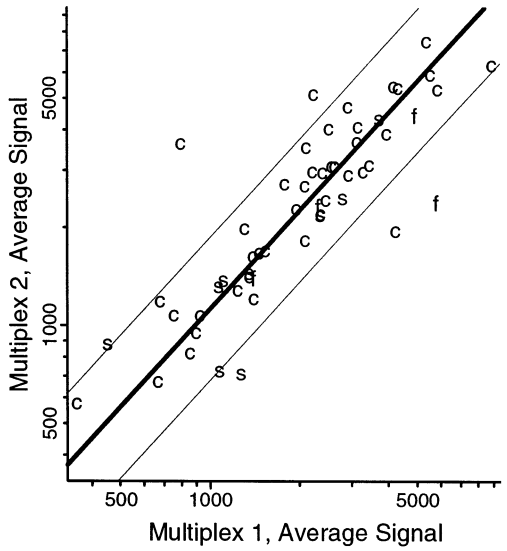

Figure 5. Within-participant, between-multiplex average signal intensities for control sample R. Each symbol denotes the average signal intensities for control sample R reported for two different STR multiplex systems analyzed on the same instrument using the same nominal analysis conditions. Comparisons of Cof to Pro+ on ABI 310 capillary instruments are denoted "c", Cof to Pro+ on ABI 377 slab gel instruments are denoted "s", and PP1.1 or PP1.1+ to PP2.1 on FMBIO slab gel instruments are denoted "f". The dark line denotes the median ratio (1.1) of Pro+/Cof average intensities. The light lines denotes the 95% confidence interval about the median (the robust SD, expressed as a multiplicative factor, is ×1.3).

CONCLUSIONS

The MSS3 results document the connections linking within-participant [DNA] measurement inaccuracy to STR multiplex signal variability. For any given participant, improving [DNA] measurement accuracy may reduce the need for repeated amplification and analysis to achieve a desired signal level.

The MSS3 results also suggest that there are 10-fold differences in amplification, separation, and detection efficiencies among similar STR multiplex systems. Measurement particulars for a given laboratory at a given time period for particular instrumentation do not adequately predict the performance of nominally identical systems. This among-participant variability cannot be attributed to generic methods or protocols, but rather, is associated with specific instruments, reagents, and analysts. This implies that STR multiplex DNA typing protocol and signal quality criteria should be performance-based and not prescriptive.

ACKNOWLEDGMENT

Interlaboratory studies are made possible by the cooperation of many analysts and laboratory supervisors. We thank them for

(17) Budowle, B. *Studies for selecting core STR loci for CODIS*. Presented at the DNA Forensics: Science, Evidence, and Future Prospects, Cambridge Healthtech Institute, McLean, VA, November 17−18, 1997.

(18) Kline, M. C.; Duewer, D. L.; Redman, J. W.; Butler, J. M.; Boyer, D. A. *Anal. Chem.* **2002**, *74* (8), 1863−1869.

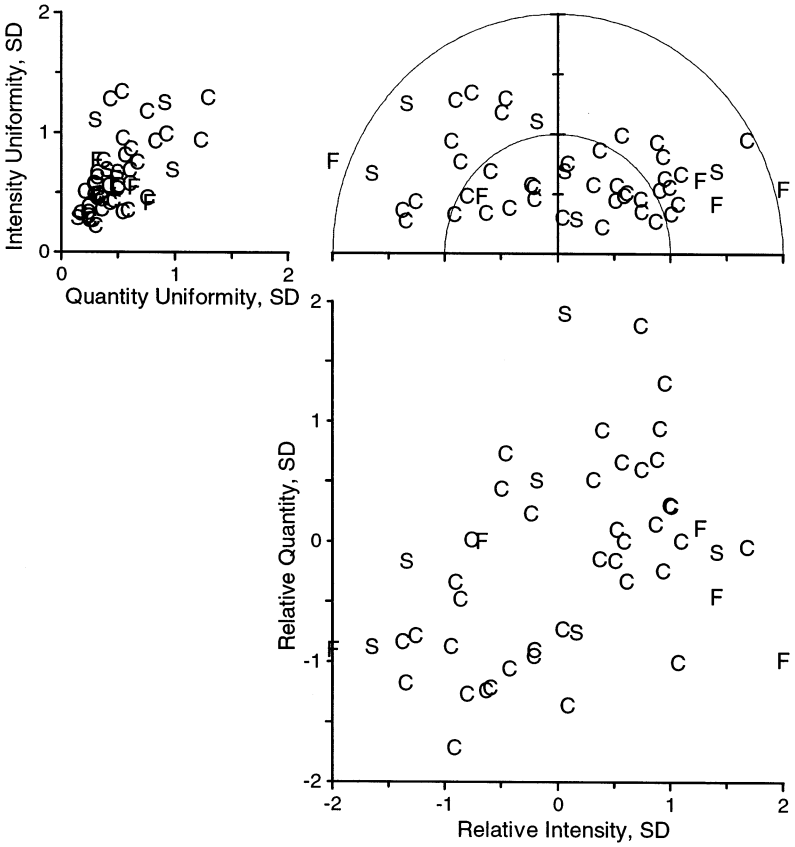

Figure 6. Signal intensities. The upper-right graphical segment is a target plot showing the among-participant distribution of average signal intensities reported per locus for three different STR multiplex systems. The lower-right segment contrasts the among-participant relative signal intensities with the relative amplified quantities (see Figure 4). The upper-left segment contrasts the within-participant uniformity of the signal intensities among the seven samples with the uniformity of the amplified quantities (see Figure 4). Results for the Cof multiplex analyzed on ABI 310 capillary instruments are denoted "C", for Cof analyzed on ABI 377 stab gel instruments are denoted "S", and PP2.1 analyzed on FMBIO slab gel instruments are denoted "F".

sharing with us their time and resources and for their willingness to tackle our contrived samples. Their voluntary and open participation in challenge exercises speaks as much to pursuit of analytical truth as to confidence in analytical systems. This work was supported in part by the National Institute of Justice through an interagency agreement with the NIST Office of Law Enforcement Standards.

Certain commercial equipment, instruments, or materials are identified in this report to specify adequately experimental conditions or reported results. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the equipment, instruments, or materials identified are necessarily the best available for the purpose.

Received for review December 10, 2002. Accepted March 10, 2003.

AC026410l