Forensics @ NIST
December 7, 2010 – Gaithersburg, MD

DNA Interlaboratory Studies

David L. Duewer
Margaret C. Kline

Overview

- Interlaboratory Studies
  - Types and goals
- Focus Areas
  - RFLP, PCR, Quantitation, Mixtures
- Accomplishments
  - Tangibles, accomplishments
- Lessons Learned

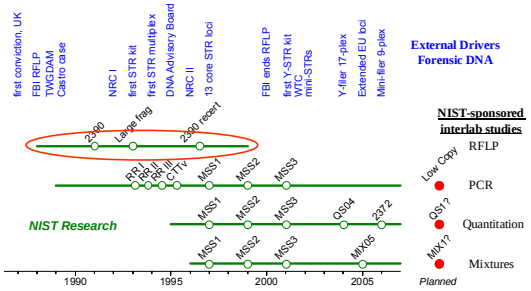
2

Interlaboratory Studies

- Multiple participants evaluate same materials
- NIST DNA interlabs are typically designed to…
  - ☺ Certify: Characterize material properties
  - ☺ Survey: Define state-of-the-measurement art
  - ☺ Gap-fill: Explore specific issues
  - 😐 Method development: Performance characteristics
    - mostly used with standardized, prescriptive methods
  - ☠ Proficiency Test (PT)
    - role of commercial providers
    - must be conducted "by the book"

3

NIST-Sponsored Studies

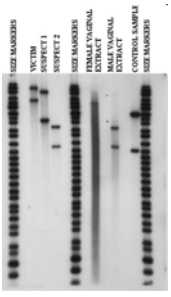


4

RFLP

Restriction Fragment Length Polymorphism

- Variable Number Tandem Repeats (VNTRs)
  - 15^{+} nucleotides / repeat
  - continuum of bands from <500 bp to >24,000 bp
- Intrinsic Limitations
  - 0.5 to 1 μg DNA
  - high quality (not-degraded) DNA
  - 1 to 2 weeks for each VNTR evaluated
- Band size measurements influenced by
  - restriction enzymes (*Hae III*, *Hinf I*, *Pst I*)
  - agarose gel composition
  - electrophoretic conditions
  - size standards
  - imaging method, equipment



5

RFLP: NIST Involvement

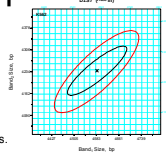
•TWGDAM Studies
- –1988: TWGDAM Phase 1 autoradiogram sizing (22 analysts)
- –1990: TWGDAM Phase 2 autoradiogram sizing (23 analysts)
- –1990: TWGDAM Phase 3 RFLP process (22 laboratories)

•Casework and Proficiency Test Compilations
- –1989 – 1996: US casework (30 laboratories)
- –1991 – 1997: proficiency tests (51 studies)
- –1993 – 1996: RCMP casework (6 regional laboratories)

•NIST-sponsored studies
- –1991: SRM 2390 material certification (29 laboratories)
- –1993: large-fragment gap-fill (20 laboratories)
- –1997: SRM 2390 recertification (20 laboratories)

6

RFLP: Tangible Outputs



- SRM 2390 DNA Profiling Standard
- CODIS K562 quality criteria
- Publications
  - Mudd *et al.* RFLP Interlaboratory Studies: Data and summary statistics. Anal Chem 1994;66:3303
  - Duewer *et al.* RFLP Interlaboratory Studies: Measurement uncertainty and its propagation. Anal Chem 1995;67:1220
  - Stolorow *et al.* RFLP Interlaboratory Studies: Repeatability and reproducibility of RFLP band sizing. Anal Chem 1996;68:1941
  - Duewer *et al.* RFLP Interlaboratory Studies: Protocol effects. Anal Chem 1997;69(10):1882
  - Duewer *et al.* RFLP Interlaboratory Studies: Precision and concordance. J Forensic Sci 1998;43:465
  - Gary *et al.* Graphical tools for RFLP DNA profiling. Laboratory Performance Charts. J Forensic Sci 1999;44:978
  - Duewer *et al.* Graphical tools for RFLP DNA profiling. Single-locus Charts. J Forensic Sci 1999;44:969
  - Duewer *et al.* RFLP band size standards: NIST Standard Reference Material® 2390. J Forensic Sci 2000;45:1093
  - Duewer *et al.* RFLP band size standards: Cell line K562 values from 1991 – 1997 PT studies. J Forensic Sci 2000;45:1106

7

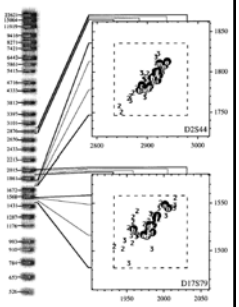
RFLP: Accomplishments

- Adequately complete metrological description of RFLP measurement process
- Measurement science established as integral contributor to interpretation and presentation of DNA evidence
- NIST established as an active participant in US and international forensic human identification communities

8

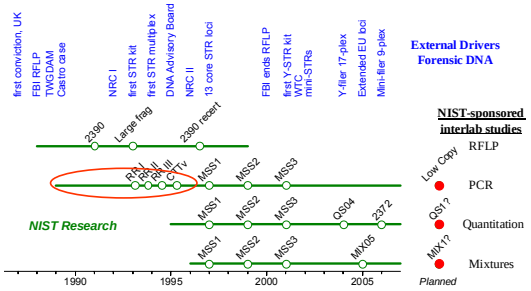
RFLP: Lessons Learned

- Small changes in protocol significantly effect sizing – need larger match windows
- DNA quantitation is an issue
- Graphical feedback more effective than just words
- Active involvement in the community essential for trust
- Building trust takes time



9

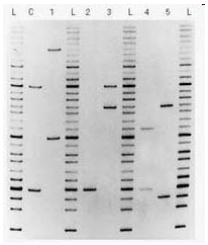
PCR: Evolving Technologies



10

PCR: D1S80

- VNTR
  - 16 bp repeat
  - Size 200 bp to 800 bp
  - Discrete alleles identified
- Intrinsic limitations
  - 2 to 5 ng DNA required
  - tolerates some degradation
- Allele size measurements influenced by
  - Gel Systems



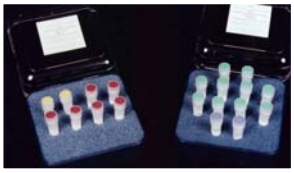
11

PCR D1S80: Early Studies

- 3 interlaboratory studies
  - RR I 16 reports
  - RR II 15 reports
  - RR III 21 reports
- Evaluated available:
  - Methods (2 gel systems, horizontal & vertical)
  - Allelic ladders (FBI/Roche)
  - Sizing standards
    - not as critical with inclusion of allelic ladder

12

PCR D1S80: Output

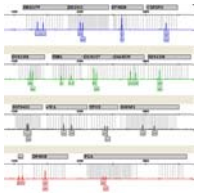


- SRM 2391
  - certified alleles
  - material/method commutability
- Harmonized allelic ladders

13

PCR: STR

Short Tandem Repeat

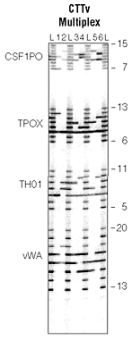


- STR
  - Relatively small number of 3 to 5 bp repeats
  - Size 60 bp to 450 bp
  - Discrete alleles identified
- Intrinsic limitations
  - 100 pg to 2 ng DNA required
  - tolerates some degradation
  - Simultaneous amplification/analysis at 16+ loci
- Allele size measurements influenced by
  - manufacturer's of kits
  - primers
  - allelic ladders
  - analysis platforms

14

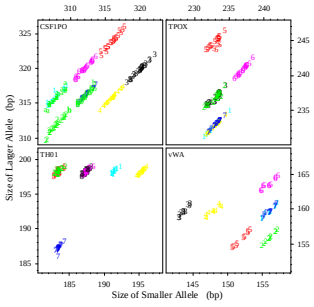
PCR: Early STRs

- CTTv
  - First commercial multiplex
    - 4 simultaneously amplified loci
- Interlab (34 participants) evaluated:
  - Static Imaging
    - MD bioimager
    - FMBIO imagers
  - Dynamic imaging
    - 373 Sequencer
    - 377 Sequencer
  - Polyacrylamide gels
    - Urea concentration 6.0 M, 7.0 M, 8.3 M
  - Precision & Trueness



15

PCR CTTv: Sizing vs Calling

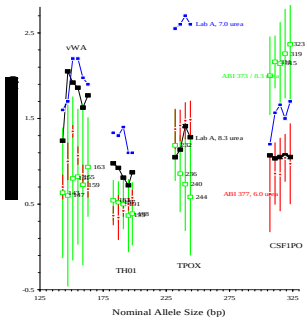


Some community leaders had argued that sizing precision was good enough that data could be exchanged as bp sizes.

Study revealed strong between-lab biases: exchanging results required "calling" to allelic ladders.

16

PCR CTTv: Urea Effect



Lab A used a 7.0 M urea gel (blue) that set their data apart from labs using 8.3 M (green) or 6.0 M (red) urea gels. The most dramatic effect was at the TPOX locus.

While bp sizing varied, calling to the allele was correct in all cases.

17

PCR CTTv: Outcomes

- Excellent within-laboratory sizing precision
- Static method as good as dynamic methods
- Between-laboratory sizing too variable (even within-kit) to exchange data as bp sizes
- Between-laboratory trueness requires "calling alleles" using allelic ladders

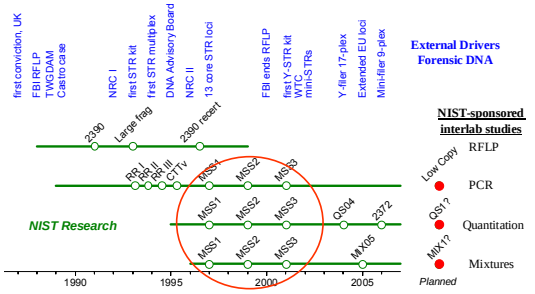
18

PCR CTTv: Tangibles

- **Community-wide agreement to use allelic ladders**
- Kline *et al*. Interlaboratory evaluation of STR triplex CTT. J Forensic Sci 1997;42:897

19

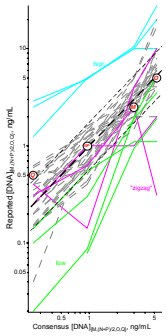
Mixed Stain Studies



20

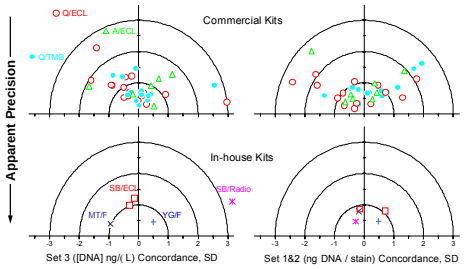
MSS: Mix Stain Studies 1 & 2

- Two studies on same materials
  - MSS1 – 22 labs
  - MSS2 – 45 labs
- Five prepared stains
- Laboratories asked to:
  - Extract
  - Quantify
  - Amplify
  - Interpret profiles
- Participants used different:
  - Extraction methods
  - Quantification methods
  - Amplification kits
  - Analysis instrumentation



21

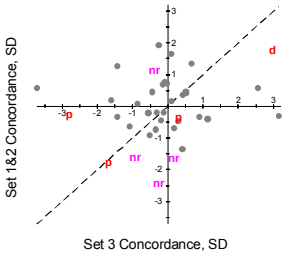
MSS: Quantitation



Concordance – multi-material analogue of bias
Apparent precision – multi-material analogue of precision
Semi-circles delimit 1, 2 and 3 comparability SDs

22

MSS: Calling Mixtures



p Partial H_u profile
d Allelic dropout
nr No result for one or more samples

Labs were learning STR analysis as well as mixture interpretation.

Many different opinions on mixture interpretation and what should be called.

Some participants had no protocols/would not generally interpret mixtures but took the study as an opportunity to explore.

23

MSS: Mixed Stain Study #3

- 74 participating laboratories
- Sent liquid DNA samples
- Sample R was sent as a single source sample to assist in understanding difference in analysis methods

24

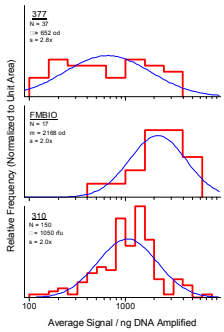
MSS: Tangible Outputs

- Duewer *et al*. 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;46:1199
- Kline *et al*. NIST Mixed Stain Study #3: DNA quantitation accuracy and its influence on short tandem repeat multiplex signal intensity. Anal Chem 2003;75:2463
- Duewer *et al*. NIST Mixed Stain Study #3: Signal Intensity Balance in Commercial Short Tandem Repeat Multiplexes. Anal Chem 2004;76:6928

25

MSS: Lessons Learned

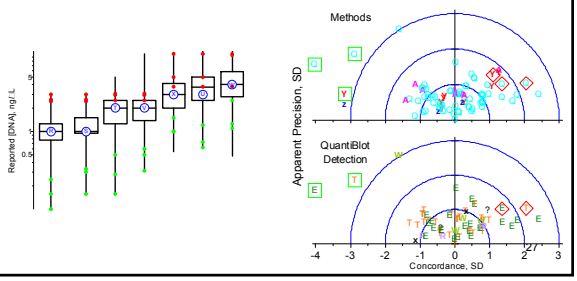
- Analysis instruments have a wide signal response range for the same input DNA
- Threshold settings need to be lab/instrument specific



26

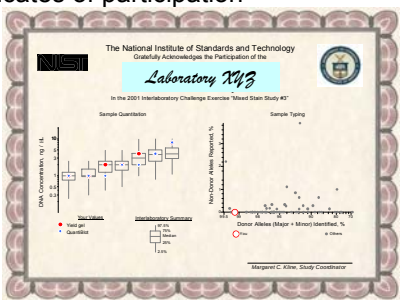
Lessons Learned

- Quantifications issues may be related to the different "standards" used.



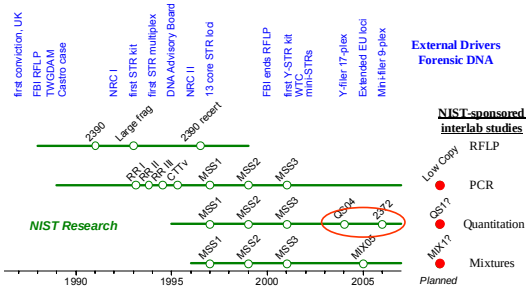
Lessons Learned

- Participants who did well framed their certificates of participation



28

Quantitation Studies



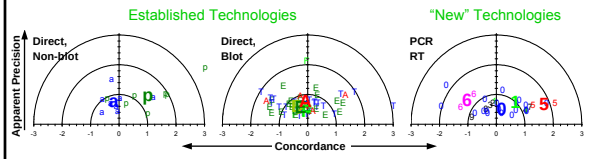
29

QS04: Quantitation Study 2004

- Eight DNA Liquid samples of [DNA]
  - 50 pg/µL to 1.5 ng/µL
- Explored:
  - concentration effects and performance characteristics at the lower DNA concentration levels frequently seen in forensic casework
  - consistency with various methodologies across multiple laboratories
  - single versus multiple source samples
  - DNA stability over time and shipping in two types of storage tubes
- 287 data sets returned from 80 participants

30

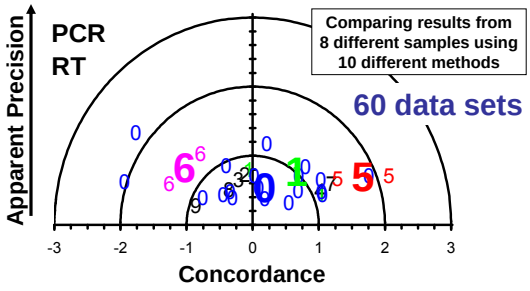
QS04: Among-Participant Results



"Bold" characters represent the median performance of all results submitted for a particular method

31

QS04: "New" qPCR methods

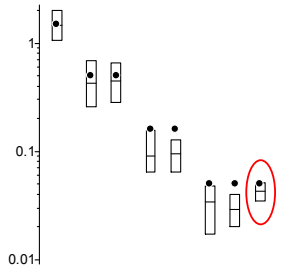


0 = Quantifiler 1 = Alu RT-PCR 5 = BRCA1 6 = CFS-HUMRT

Kline, *et al.* (2005) *J. Forensic Sci.* 50(3):571-578

32

QS04: Qfiler Results from NIST



Results for the sample in PFA (*ie*, Teflon) tubes were consistently close to the nominal DNA concentration of 50 pg/µL .

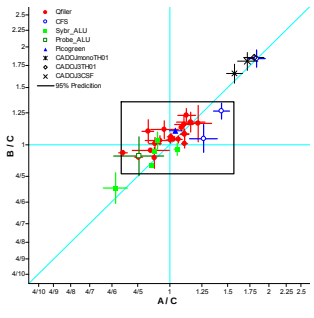
33

SRM 2372 Quantitation Standard

- Study limited to 32 forensic DNA testing facilities
- One unit of each of three components: A, B, C
- Participants were asked to:
  - use the same dilution scheme for all three components: 50, 10, 5, 2.5, 1.25, 0.62, 0.31, and 0.16 ng/µL.
  - use component C dilutions as the calibration standard for their quantitation assay(s)
  - assume that the "true" [DNA] of component C was exactly 50 ng/µL
  - calculate the apparent [DNA] for all of the dilutions made for components A and B
  - report all associated cycle threshold (Ct) values if qPCR method(s) used

34

SRM 2372: Data



While the Interlaboratory data for Component B looks good, it failed homogeneity testing resulting in all 1700 vials being **disposed** of and a **new lot** of Component B being produced.

35

SRM 2372: Lessons

- Participants did as we asked but:
  - There were differences in the linear dynamic range of the assays used
    - Could have used a different dilution series
  - There were apparent method dependent bias for the different components
    - Not totally unexpected
- Faster isn't always quicker
  - Homogeneity evaluation first, interlab second

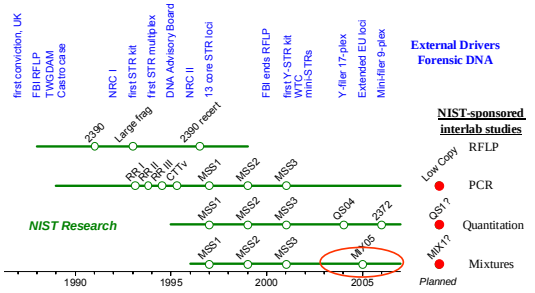
36

SRM 2372: Tangibles

- SRM 2372 Human DNA Quantitation Standard
- Kline *et al*. Production and certification of NIST SRM 2372 Human DNA Quantitation Standard. Anal Bioanal Chem 2009;394:1183

37

Mixture Study



38

MIX05: Data Interpretation

- Designed to:
  - evaluate STR mixture interpretation in the forensic DNA typing community
  - aid development of training tools for mixture interpretation and reporting
- DNA mixtures for 4 mock sexual assault case scenarios
  - six kits: Profiler Plus, COfiler, SGM Plus, Identifiler, PP16, PP16 BIO
- In each case, we provided the "evidence" sample result,
  - a mixture of at least one perpetrator and a victim
  - the "victim" reference sample
  - electrophoretic data (ABI 3100 .fsa files made available at http://www.cstl.nist.gov/biotech/strbase/interlab/MIX05.htm).
  - Labs, including Macintosh-based users, that could not download data from the MIX05 website were shipped CD-ROMs or zip disks.
- **94** laboratories enrolled

39

MIX05: What We Requested

- Report the results as though they were from a real case including whether a statistical value would be attached to the results
- Summarize the perpetrator(s) alleles in each "case" as they might be presented in court—along with an appropriate statistic (if warranted by laboratory SOP) and the source of the allele frequencies used to make the calculation
- State which kit(s) were used to solve each case
- Estimate the ratio for samples present in the evidence mixture and describe how this estimate was determined
- Copy of laboratory's mixture interpretation guidelines and a brief explanation as to why conclusions were reached in each scenario

40

MIX05: Data received

- 69 labs returned results
- 50 labs made allele calls
  - *ie*, 19 labs did *not* make allele calls
- 39 labs estimated ratios
- 29 labs provided stats
- Remember - ***94*** *labs signed up*

41

MIX05: Lessons Learned

- Wording of the scenario is important!
  - We did not say these were ***intimate*** samples so many labs would not continue
- Labs are not comfortable with analyzing data that was not collected by their own protocols

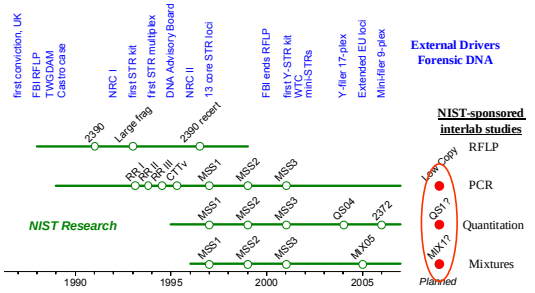
42

MIX05: Outputs

- Poster at: 16th International Symposium on Human Identification, Grapevine, TX, Sept 26-28, 2005
- SWGDAM Interpretation Guidelines have recently become available
- AAFS 2008 DNA Mixture Workshop DNA Mixture Interpretation: Principles and Practice in Component Deconvolution and Statistical Analysis
- Workshop at the 21st International Symposium on Human Identification (San Antonio, TX), October 11, 2010, "Mixture Interpretation: Principles, Protocols, and Practice"

43

Future



44

Thank you for your Attention!!

- Margaret Kline
  - 301-975-3134
  - margaret.kline@nist.gov
- Dave Duewer
  - 301-975-3935
  - david.duewer@nist.gov

Thanks to:
John Butler
Dennis Reeder
Janette Redman
Kristy Richie

45