NIST Standards for **Genetic Testing**:

Past, Present, and Future

John M. Butler, Ph.D.
National Institute of Standards and Technology
Biochemical Science Division/Applied Genetics Group

SACHHS
December 1, 2008

Presentation Overview

- Past
  - Extensive experience with developing forensic DNA reference materials and genotyping assays and technologies
- Present
  - Applied Genetics Group to consolidate forensic DNA with clinical genetics and agricultural biotech efforts
  - Work with genetic genealogy
- Future
  - Planned genetic testing standards

Congress Passed **the DNA Identification Act of 1994** (Public Law 103 322)

Formalized the FBI's authority to establish a national DNA index for law enforcement purposes.

FBI's DNA Advisory Board

Quality Assurance Standards for Forensic DNA Testing Laboratories

(October 1, 1998)

STANDARD 9.5

The laboratory shall check its DNA procedures annually or whenever substantial changes are made to the protocol(s) against an appropriate and available NIST standard reference material or standard traceable to a NIST standard.

The Tools of DNA Typing and SRM Needs

- RFLP Testing (Late 1980's) ~~SRM 2390~~
  - Radioactive Based
  - Chemiluminescent Based

Technology no longer used

- PCR-Based Testing (Mid 1990's)
  - Dot-Blot
  - VNTR
  - STR (Fluorescent markers used today)

SRM 2391..a..b

Growth area

- DNA Sequencing (Late 1990's) **SRM 2392, 2392-I**
  - Mitochondrial DNA

- Y-Chromosome Testing (early 2000's) **SRM 2395**

Growth area

2003: NIST SRM 2391b

Driven primarily by commercial kit loci…

National Institute of Standards & Technology

Certificate of Analysis

Standard Reference Material® 2391b

PCR-based DNA Profiling Standard

This Standard Reference Material (SRM) is intended primarily for use in the standardization of forensic and paternity quality assurance procedures for Polymerase Chain Reaction (PCR)-based genetic testing and for instructional law enforcement or non-clinical research purposes. This SRM can also be used for quality assurance when assigning values to in-house control materials. It is not intended for any human or animal clinical diagnostic use. Note that SRM 2391b is slightly modified from SRM 2391, in that there is more emphasis on Short Tandem Repeats (STRs) and less emphasis on D1S80 [1,2] reflecting the growing interest and utility of STRs [3 to 14]. Additional information on each STR locus can be found at a NIST-sponsored database on the internet: http://www.cstl.nist.gov/biotech/strbase [14].

This SRM is composed of well-characterized human deoxyribonucleic acid (DNA) in two forms: genomic DNA and DNA to be extracted from cells spotted onto filter paper. A unit of the SRM is composed of 12 frozen components packaged in one box. See the section in this certificate entitled *Description of Components* for a complete listing of the components.

Certified Values: The SRM is certified for genetic loci of forensic interest that were commercially available at the time of production. Genetic types for these loci can be found in Tables 1, 2, and 3. The tables are organized as follows: Table 1 lists the genetic types for the Federal Bureau of Investigation's (FBI's) CODIS (COmbined DNA Index System) core STR loci; Table 2 lists additional STR loci of interest; and Table 3 lists the genetic types for D1S80, AmpliType® PM + HLADQA1, and Amelogenin.

Expiration of Certification: The certification of this SRM is valid until **31 December 2008**, provided the SRM is handled and stored in accordance with the instructions given in this certificate. However, the certification is invalid if the SRM is contaminated or otherwise modified.

Maintenance of SRM Certification: NIST will monitor this SRM over the period of its certification. If

SRM 2391b Page 1 of 3

2. Certified Values for Additional STR Loci

F13B	FES/FPS	LPL	Penta D	Penta E	D2S1338	D19S433
10,10	12,12	10,11	10,15	7,12	17,23	13,16.2
8,10	10,11					16
9,10	11,12					4
6,9	10,13					3
8,9	11,13					14
9,10	11,11	10,12	9,12	12,14	25,25	12,14
6,8	11,11*	11,12	3.2,11	12,16	17,22	13,15.2

22 autosomal STRs characterized across 12 DNA samples

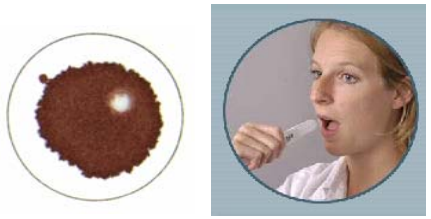
Consumption of SRM 2391b has slowed because we have encouraged labs to create NIST-traceable materials or only use portions of the SRM's 12 components each time when the annual calibrations are performed (i.e., to stretch out the use of one unit of SRM 2391b)

Steps in Forensic DNA Analysis

Usually 1-2 day process (a minimum of ~5 hours)

Steps Involved

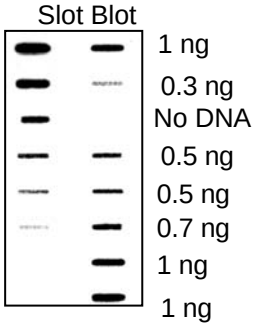
- Collection
- Specimen Storage
- Extraction
- Quantitation
- Multiplex PCR
- STR Typing
- Interpretation of Results
- Database
  Storage & Searching
- Calculation of Match Probability



Blood Stain Buccal swab

Sample Collection & Storage

DNA Extraction



DNA Quantitation

Biology

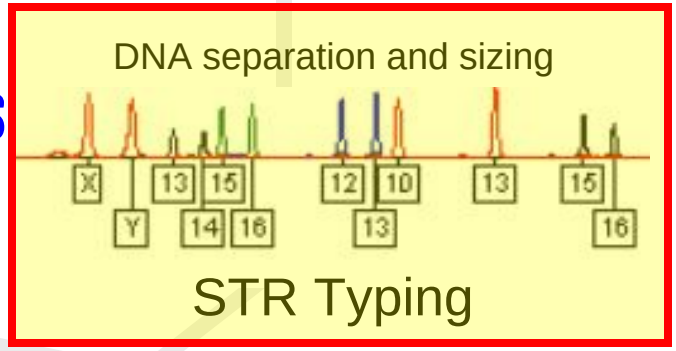
Multiplex PCR Amplification

Genetics

If a match occurs, comparison of DNA profile to population allele frequencies to generate a case report with probability of a random match to an unrelated individual

DNA Database Search

Technology

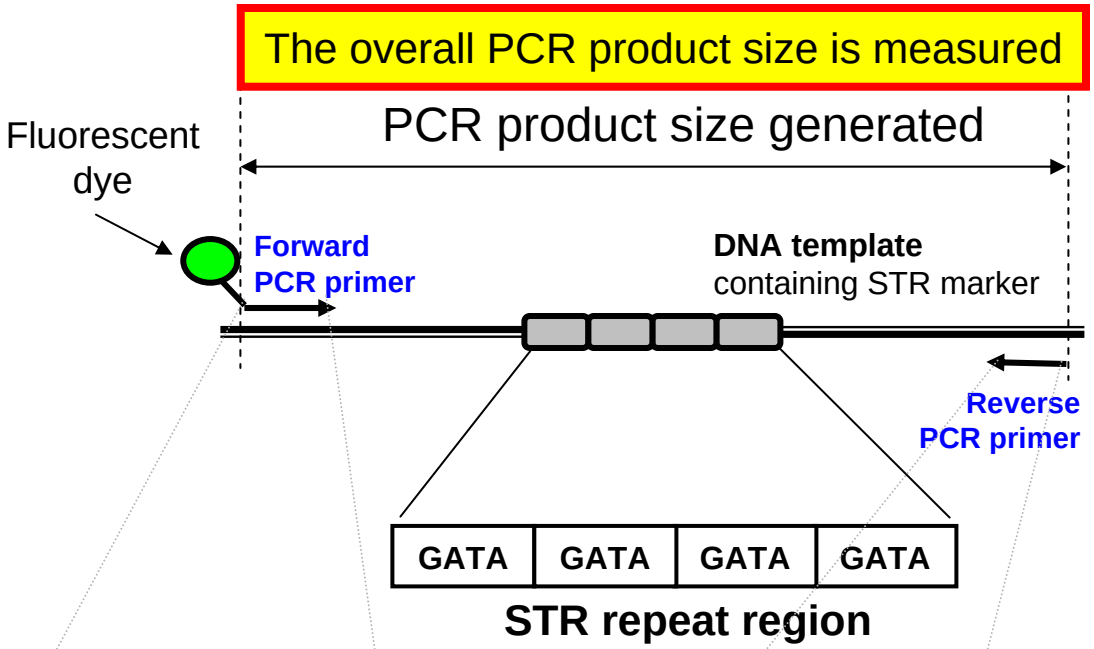


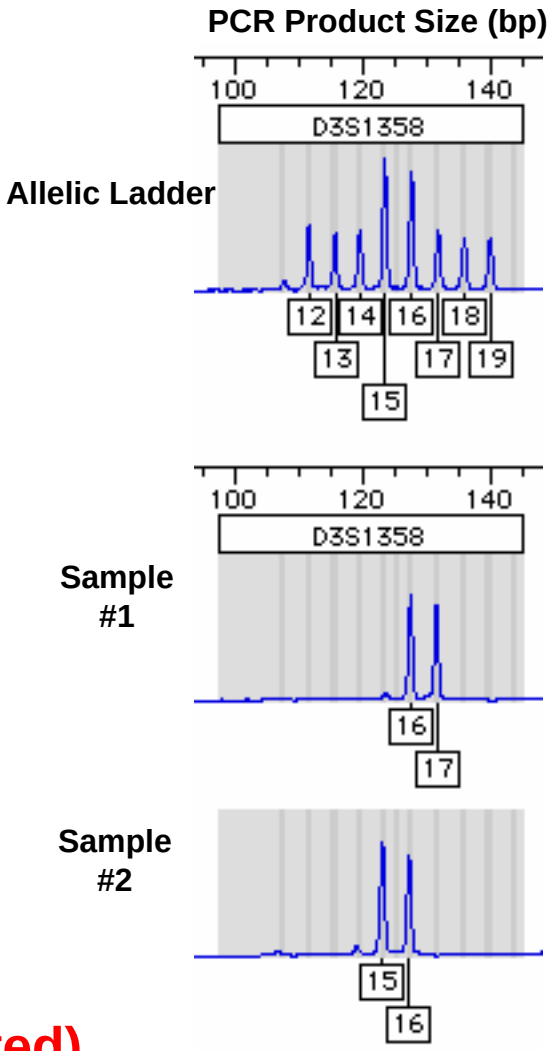
Male: 13,14-15,16-12,13-10,13-15,16

Interpretation of Results

Short Tandem Repeat (STR) Markers

PCR primers anneal to unique sequences bracketing the variable STR repeat region

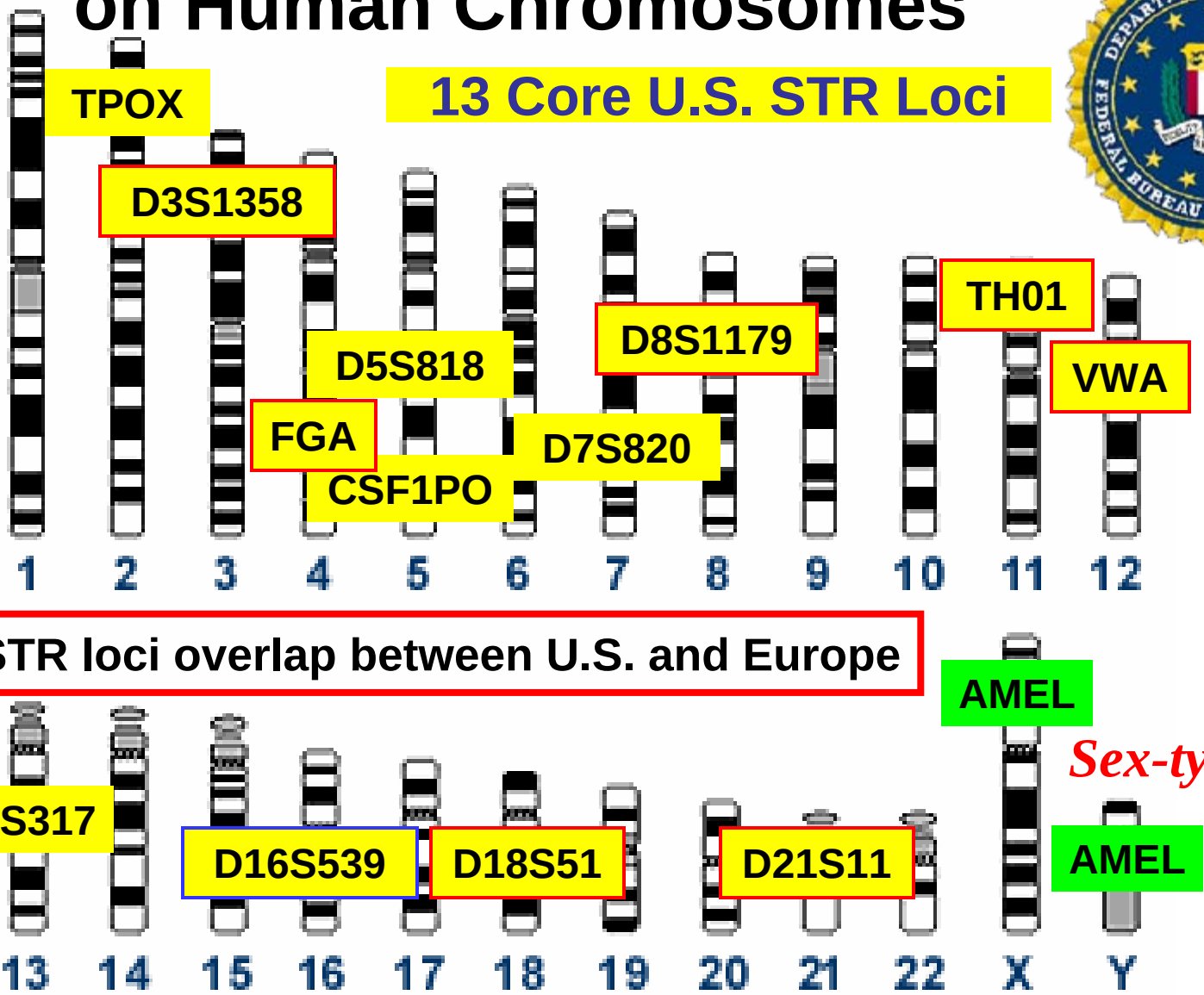




TCCCAAGCTCTTCCTCTTCCCTAGATCAATACAGACAGA
AGACAGGTG**GATAGATAGATAGATAGATAGATAGATA**
GATAGATAGATAGATATCATTGAAAGACAAAACAGAGA
TGGATGATAGATACATGCTTACAGATGCACAC

= 11 GATA repeats ("11" is all that is reported)

Position of Forensic STR Markers on Human Chromosomes



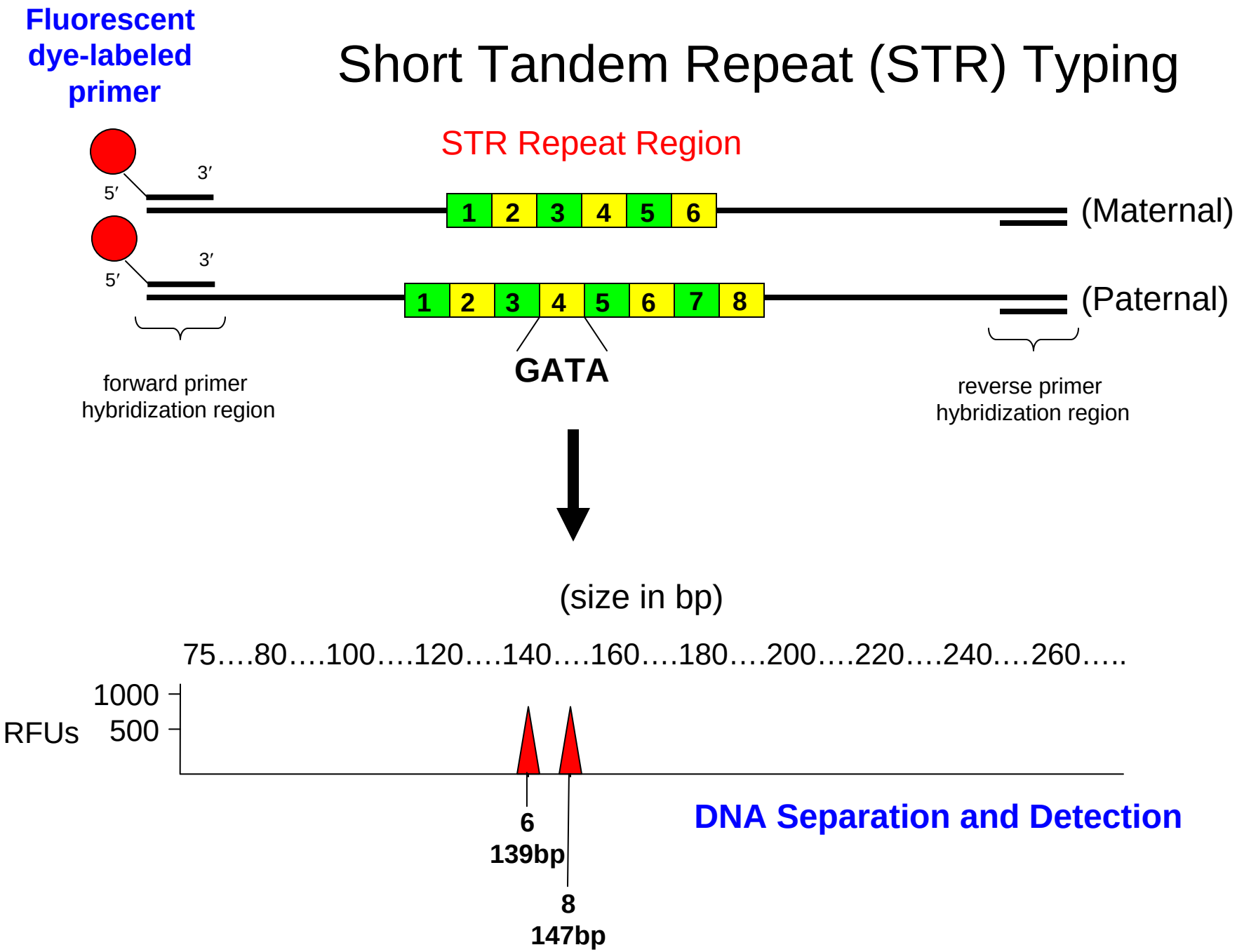
13 Core U.S. STR Loci

1997

8 STR loci overlap between U.S. and Europe

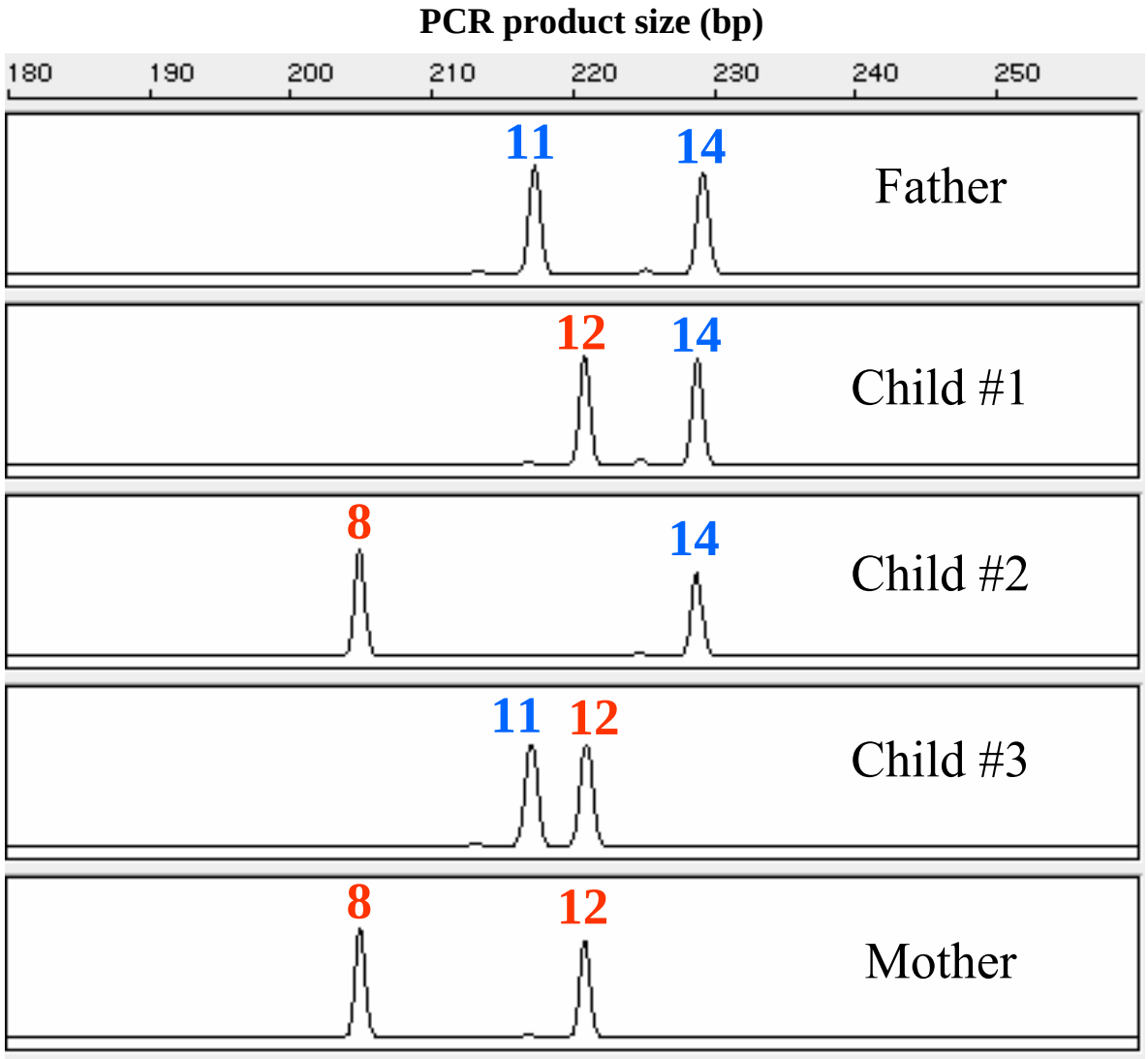
Sex-typing

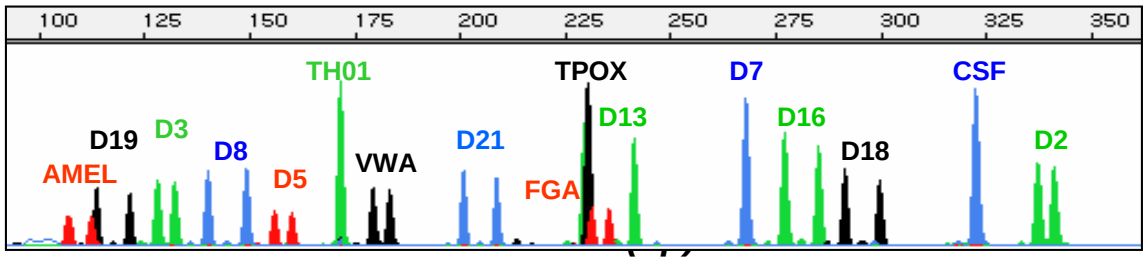
Short Tandem Repeat (STR) Typing

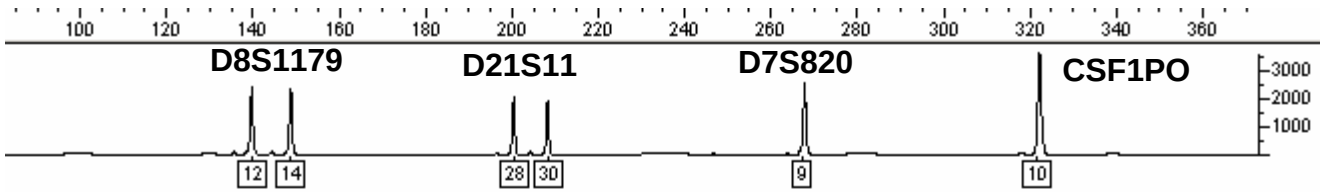


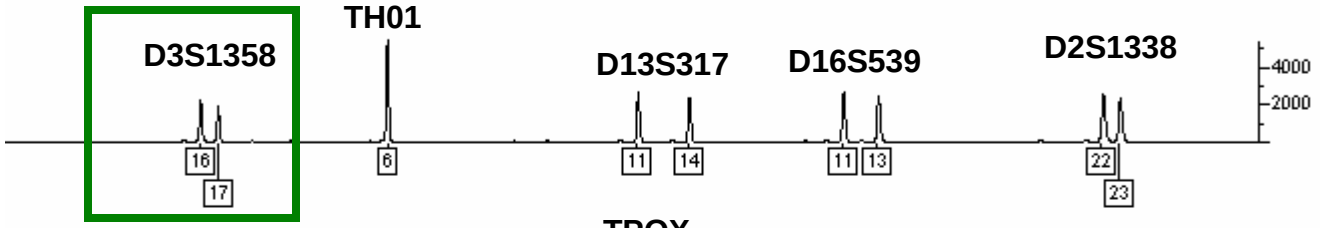
PATERNITY TESTING

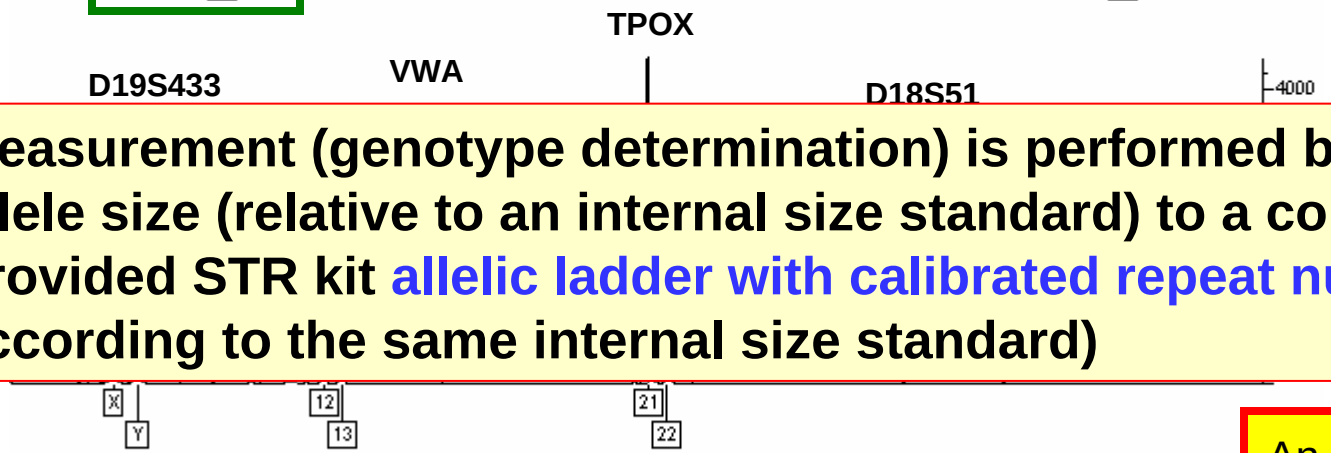
Family Inheritance of STR Alleles (D13S317)

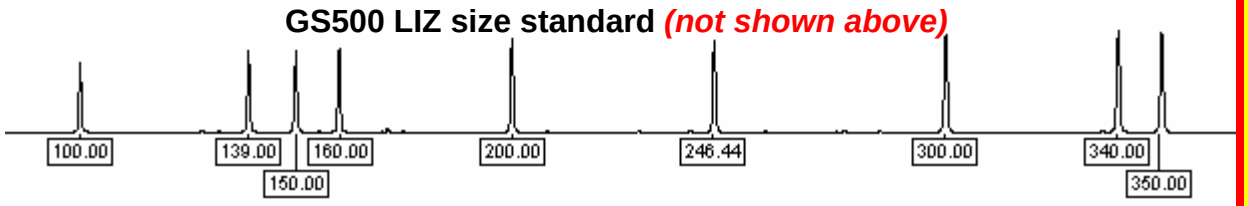


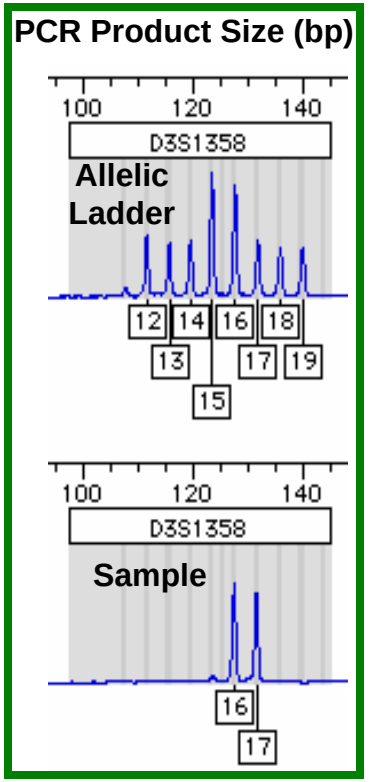










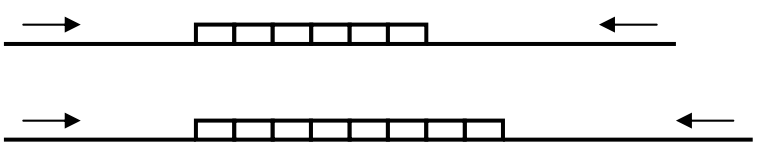


Measurement (genotype determination) is performed by comparing allele size (relative to an internal size standard) to a commercially provided STR kit allelic ladder with calibrated repeat numbers (sized according to the same internal size standard)

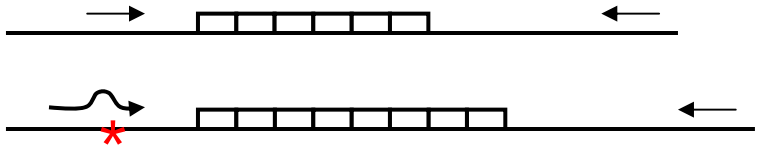
An **internal size standard** is run with each sample and external standard to correlate sizes.

Different Genetic Tests Can Give Different Results Based on PCR Primer Positions

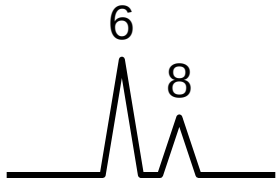
PCR Primers in Different Positions around the STR repeat region



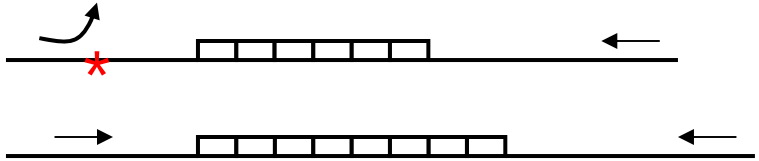
Heterozygous alleles are well balanced

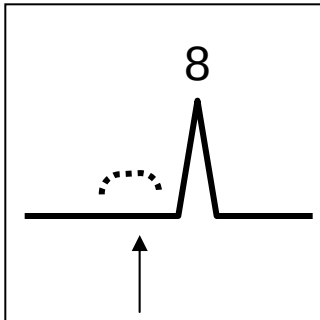


Imbalance in allele peak heights



Mutations in the DNA Sequence (impact PCR primer annealing)



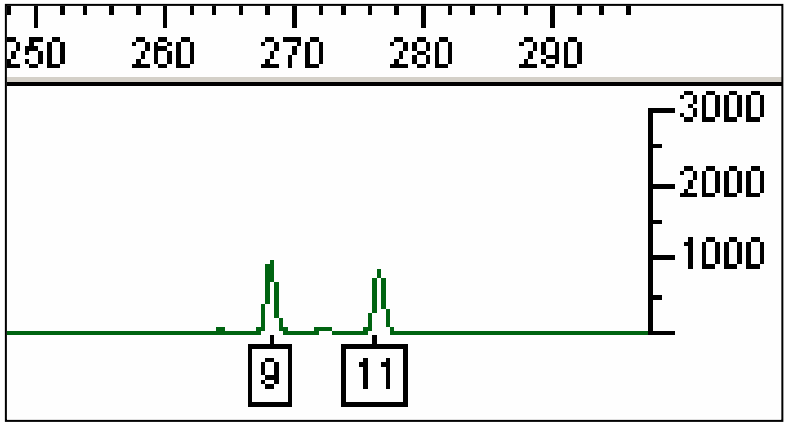


"Null" Allele from Allele Dropout

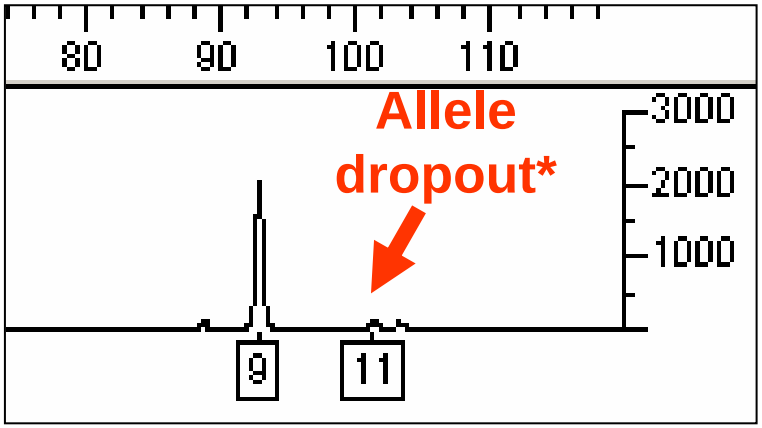
SRM 2391b Genomic 8 with D16S539

All allele calls with MiniFiler for CSF1PO, D7S820, D13S317, D18S51, D21S11, FGA, and D16S539 (with the exception noted below) ***match previously certified values.***

Identifiler

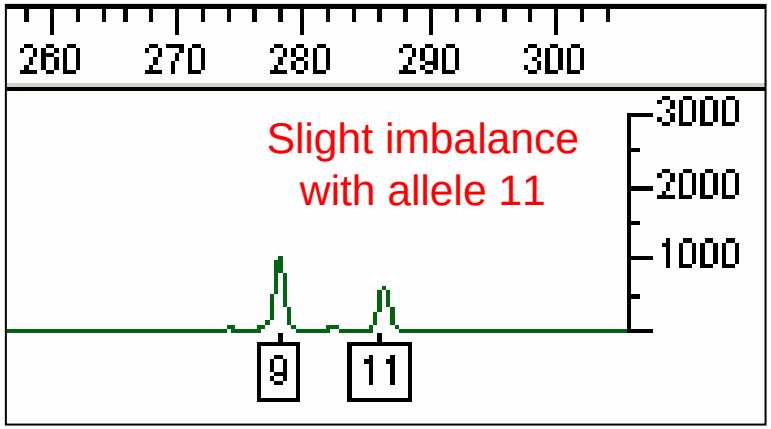


MiniFiler



***Due to primer binding site mutation**

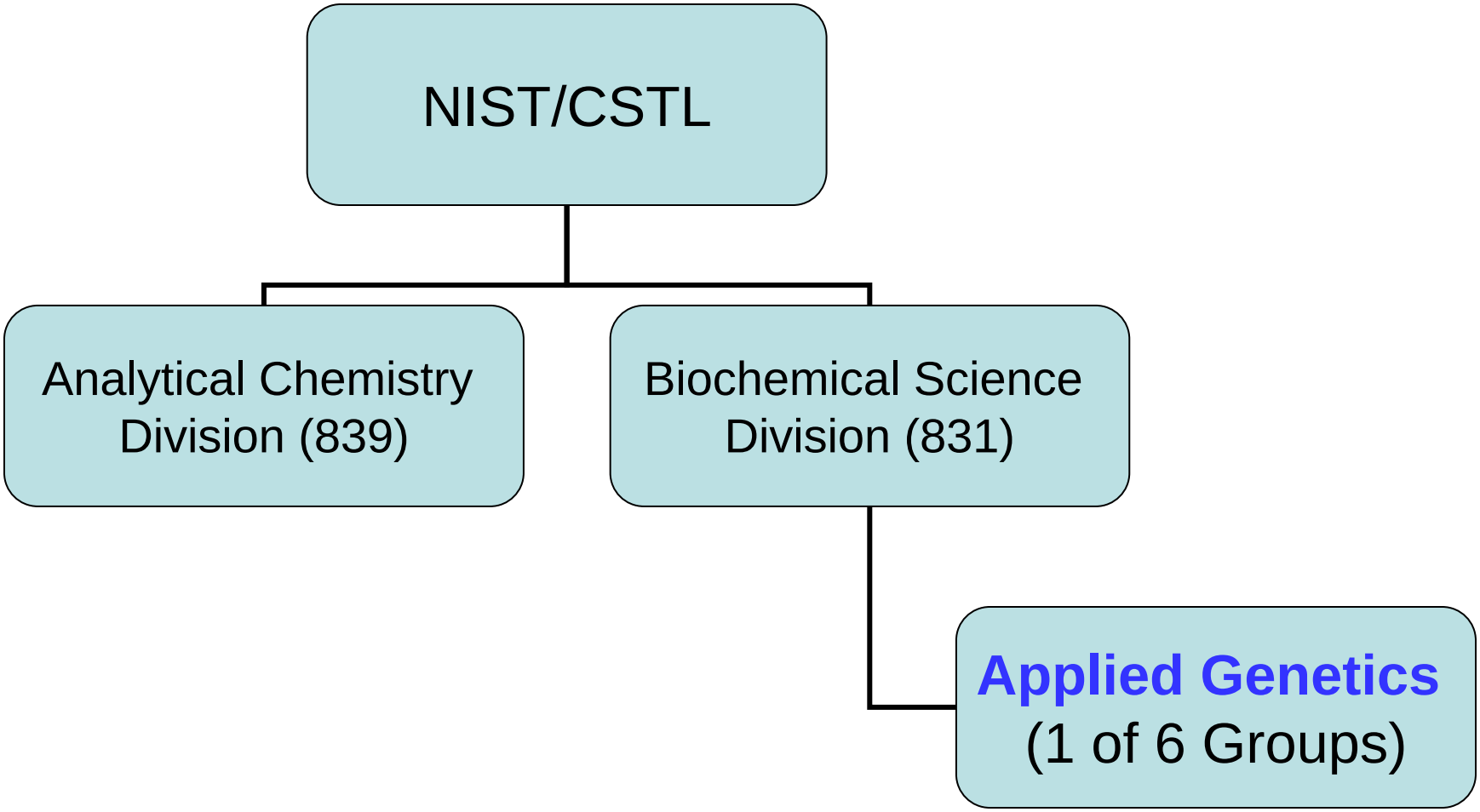
PowerPlex 16



Current Areas of NIST Effort with Forensic DNA

- **Standards** http://www.cstl.nist.gov/biotech/strbase/
  - **Standard Reference Materials**
  - Standard Information Resources (STRBase website)
  - Interlaboratory Studies
- **Technology**
  - Research programs in SNPs, miniSTRs, Y-STRs, mtDNA, qPCR
  - Assay and software development
- **Training Materials**
  - Review articles and workshops on STRs, CE, validation
  - PowerPoint and pdf files available for download

An Abbreviated Organizational Chart



Responsibilities for Forensic DNA Testing
and Clinical and Agricultural Diagnostics (GMOs)

NIST Applied Genetics Group

Applied Genetics

Formally organized October 2008

Group Leader

John Butler

Marcia Holden

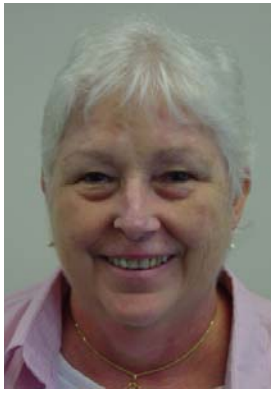
Margaret Kline

Pete Vallone

Amy Decker

Ross Haynes

Becky Hill



Jan Redman

NIST
Applied Genetics

Group Mission Statement

Advancing technology and traceability through quality genetic measurements to aid work in

- **forensic DNA testing,**
- **clinical genetics,**
- **agricultural biotechnology,** and
- **DNA biometrics.**

Applied Genetics

Group Expertise and Funding Sources

Group Expertise

- Reference Material Characterization
- Standard Information Resource Development
- Rapid Multiplex PCR Assay Construction
- Short Tandem Repeat (STR) Genotyping
- Single Nucleotide Polymorphism (SNP) Genotyping
- DNA Sequencing
- Training Materials and Workshops (validation info)

Current Funding Sources

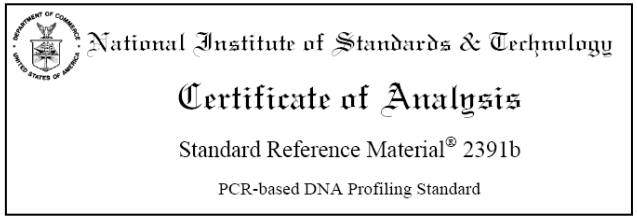
- **National Institute of Justice** (Forensic DNA)
- NIST (SRM development and production)

We are looking to strengthen our portfolio in clinical genetics and agricultural biotech

Standard Reference Materials (SRMs)

http://www.nist.gov/srm

Traceable standards to ensure accurate and comparable measurements between laboratories

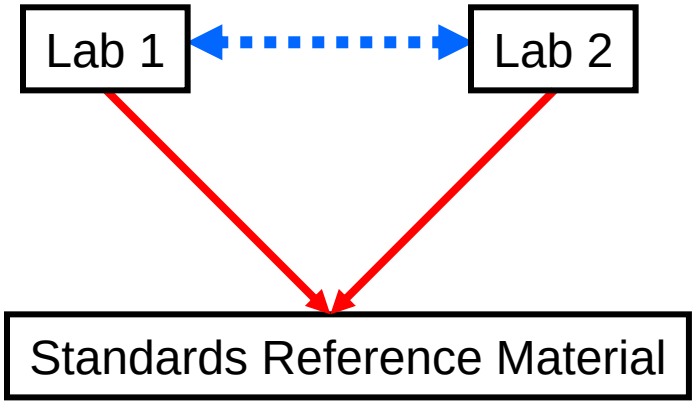

National Institute of Standards & Technology

Certificate of Analysis

Standard Reference Material® 2391b

PCR-based DNA Profiling Standard

SRM 2391b – autosomal STRs
SRM 2392 &-I – mtDNA sequencing
SRM 2395 – Y-STRs
SRM 2372 – DNA quantitation
SRM 2394 – mtDNA heteroplasmy
SRM 2399 – Fragile X



Calibration with SRMs enables confidence in comparisons of results between laboratories

Helps meet ISO 17025 needs

for traceability to a national metrology institute

NIST DNA Reference Materials

Applied Genetics

Date of release or certificate revision (r)

Forensic Applications

- **STR PCR DNA Profiling** (SRM 2391b) – 1995, r2008
- Mitochondrial DNA Sequencing (SRM 2392-I, 2392) – 1999, 2003
- **Human Y-Chromosome DNA Profiling** (SRM 2395) – 2003, r2008
- RFLP DNA Profiling (SRM 2390) – 1992, r2001, *now obsolete*

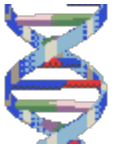
Clinical Applications

- Fragile X Human DNA Triplet Repeat (SRM 2399) – 2004, r 2007
- Huntington's Disease CAG Repeats (SRM 2393) – in process

Platform Testing

- **Human DNA Quantitation** (SRM 2372) - 2007
- Heteroplasmic mtDNA Mutation Detection (SRM 2394) - 2004
- DNA Sequence Library for External RNA Controls (SRM 2374)

A few others are in early stages of development

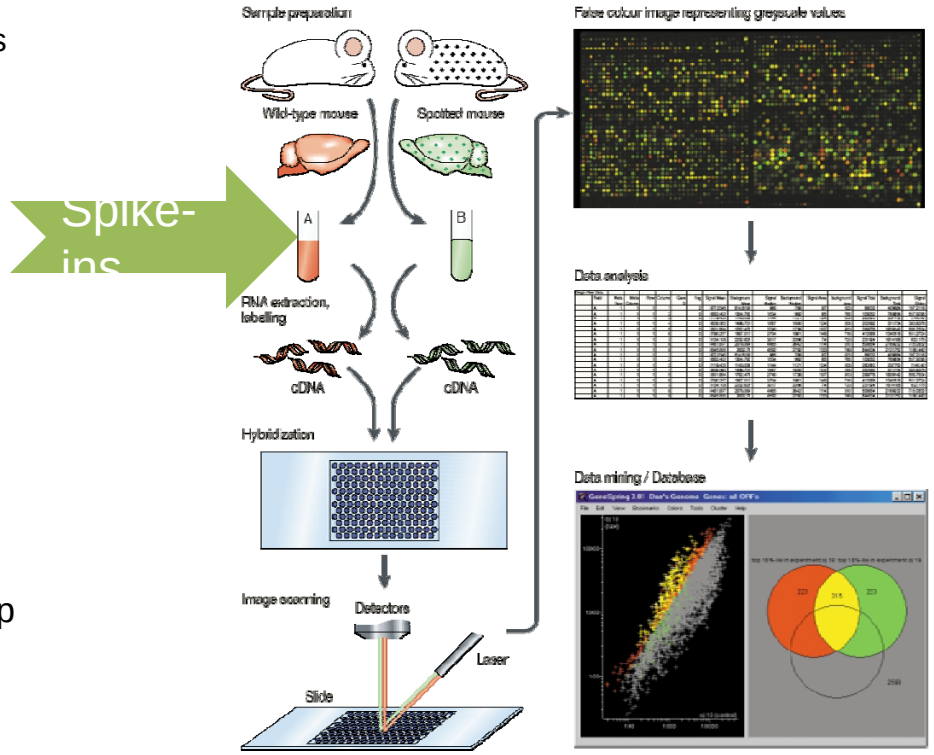


Slide from **Marc Salit**, Multiplexed Biomolecular Science Group

External RNA Control Consortium

- Industry-initiated, NIST-hosted, stakeholder coupled
  - Janet Warrington, VP Clinical Genomics at Affymetrix
  - all major microarray technology developers
  - other gene expression assay developers
    - collaborative study
    - probe content on commercial array platforms
- Use reference material approach to transfer accuracy of NIST measurements and ensure harmony amongst users
  - Long-term useful for gene expression, not tied to microarray measurement approach
- Novel aspects
  - Certification of sequence
    - developing new metrological framework for certifying sequence as property, consistent with ISO/REMCO definition of CRM
    - focus on confidence in sequence
  - SRM to be template; work with SDO to develop documentary standard for CRM production
    - CRMs to be commercially available

- Model for future work in this area
  - sets up our work to be well-coupled to stakeholder needs
  - keeps us relevant and tied in
  - lets us develop SRMs that are stable and generic

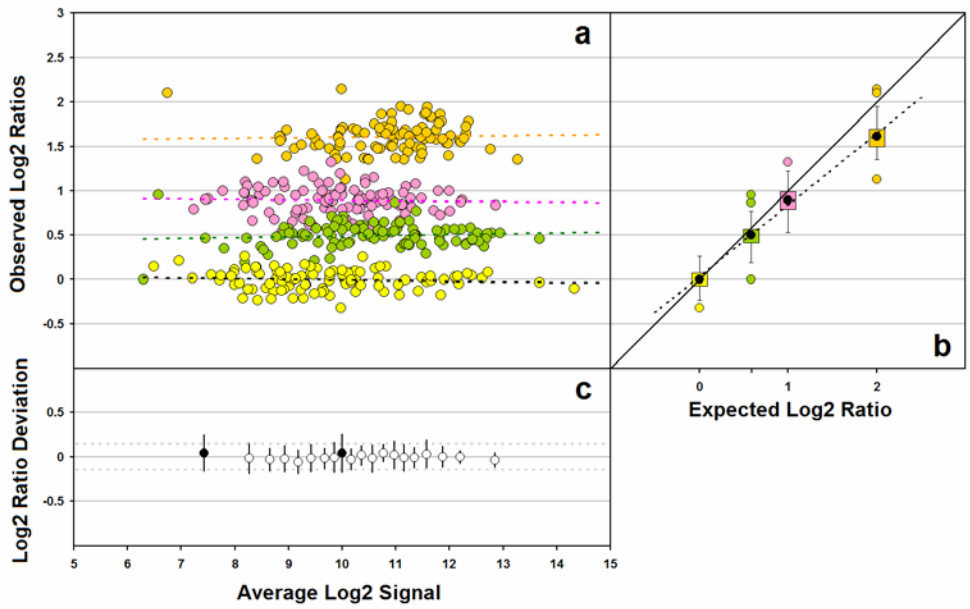


Slide from **Marc Salit**, Multiplexed Biomolecular Science Group

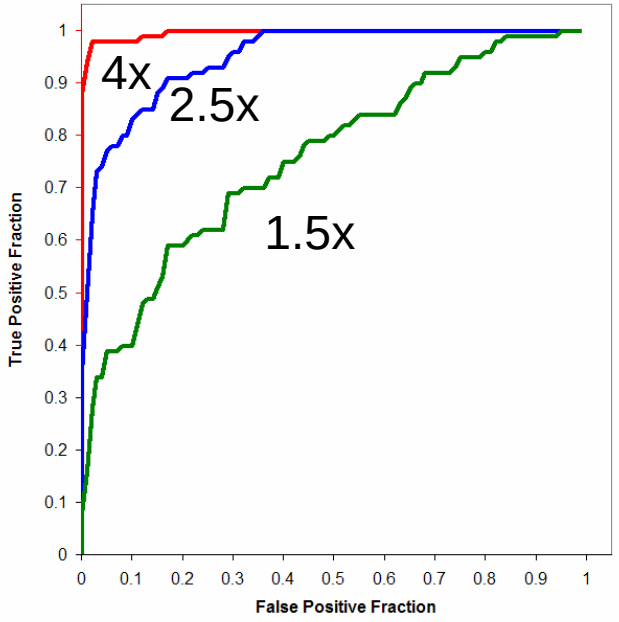
RNA Control Set

enable objective performance measures for microarray gene expression

Analytical performance
wrt signal level and ratio



Receiver-Operating Characteristic Curves



Approach developed in conjunction with Pine and Thompson, FDA
– results of simulation shown –

Slide from **Marc Salit**, Multiplexed Biomolecular Science Group

NIST SRM 2374 – DNA Sequence Library for External RNA Controls

- NIST developing reference material of 96 control sequences
  - SRM will be plasmid DNA with control sequences as inserts
  - sequence is certified property
  - sequencing at NIST and multiple partner labs
  - sequencing with Sanger and next-gen “UHTS” approach(es)
- Developed sequence library from submission by ERCC members & synthesis
  - evaluated performance of RNA controls on variety of platforms
  - selected 96 well performing sequences
- Preparing SRM
  - cloned sequence library into common vector
    - suitable for use in accurate preparation of RNA controls
  - Prepared 400 units
    - 96 tubes in each
- Certifying ~100,000 bases
  - Sanger sequencing complete at CBI, NIST
  - alternate sequencing approaches underway
  - quality measures developed to permit estimation of sequence reliability
    - based on *de novo* assembly at alternate sites
    - integration of data from multiple labs

Some Issues Faced When Developing Reference Materials

- Initial selection of material (SRM components) was for a specific purpose usually and may not address every need in the future (a new locus may not exhibit a diverse set of alleles)

- The forensic community uses commercial STR typing kits – and only wants a confirmation of the allele calls against an allelic ladder – should we fully sequence every sample?

- Some genetic loci will not be able to have every allele sequenced (e.g., due to locus duplication)

- There are lots of loci that could be “certified” – how do we decide which ones to include in future certificate updates?

Applied Genetics

Thank You for Your Attention…

Contact Information

john.butler@nist.gov

301-975-4049