

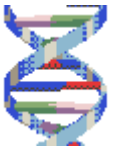


SWGDM (Fredericksburg, VA)
January 19, 2012

NIST Update

John M. Butler

NIST Applied Genetics Group
National Institute of Standards and Technology
Gaithersburg, Maryland



NIST Human Identity Project Teams

within the Applied Genetics Group

Forensic DNA Team

Guest Researcher

DNA Biometrics Team

Funding from the **National Institute of Justice (NIJ)** through NIST Office of Law Enforcement Standards

Funding from the **FBI S&T Branch** through NIST Information Access Division



John Butler



Mike Coble



Becky Hill



Margaret Kline



Manuel Fondevila Alvarez



Pete Vallone



Erica Butts



Kevin Kiesler

STRBase, Workshops & Textbooks

Concordance & LT-DNA Mixture, mtDNA & Y

SRM work, variant alleles & Cell Line ID

Data Analysis Support



Dave Duewer

Rapid PCR, Direct PCR & Biometrics

ABI 3500 & DNA Extraction

PLEX-ID & NGS Exploration



Office Manager Patti Rohmiller



NIST STRBase Website

<http://www.cstl.nist.gov/biotech/strbase/>



John Butler

Forensic STR Information

- [STRs101: Brief Introduction to STRs](#)
- [Core Loci: FBI CODIS Core STR Loci and European Core Loci](#)
- [STR Fact Sheets \(observed alleles and PCR product sizes\)](#)
- [Multiplex STR kits](#)
- [Sequence Information \(annotated\)](#)
- [Variant Allele Reports](#) ◆
- [Tri-Allelic Patterns](#) ◆
- [Mutation Rates for Common Loci](#)
- [Published PCR primers](#)
- [Y-chromosome STRs](#) ◆
- [Low-template DNA Information](#) *Updated*
- [Mixture Interpretation](#) *NEW*
- [Kinship Analysis](#) *NEW*
- [miniSTRs \(short amplicons\)](#) ◆
- [Null Alleles](#) - discordance observed between STR kits ◆
- [STR Reference List](#) - now 3400 references ◆

[Cataloged as of Dec 2011](#)

605 variant alleles

305 tri-allelic patterns

**We invite labs to supply
information on variant
and tri-alleles observed**

Forensic DNA Typing Textbook

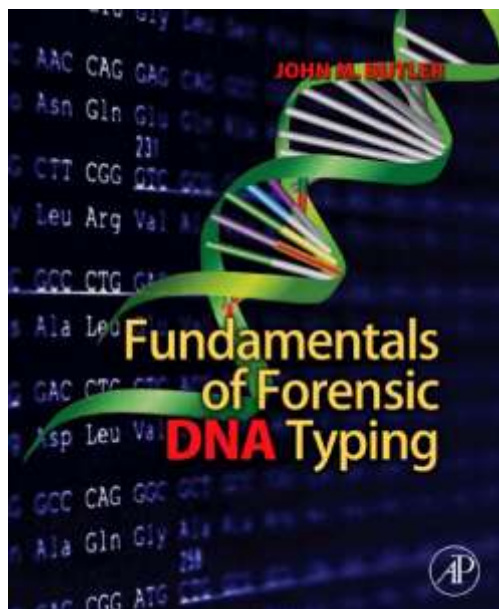
3rd Edition is Three Volumes

Now part of my job at NIST (no royalties are received)



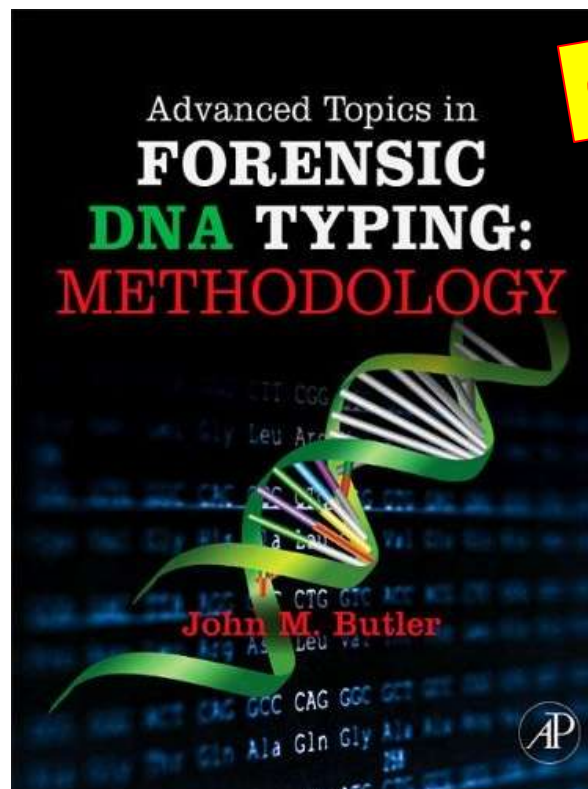
John Butler

*For beginning students,
general public, & lawyers*



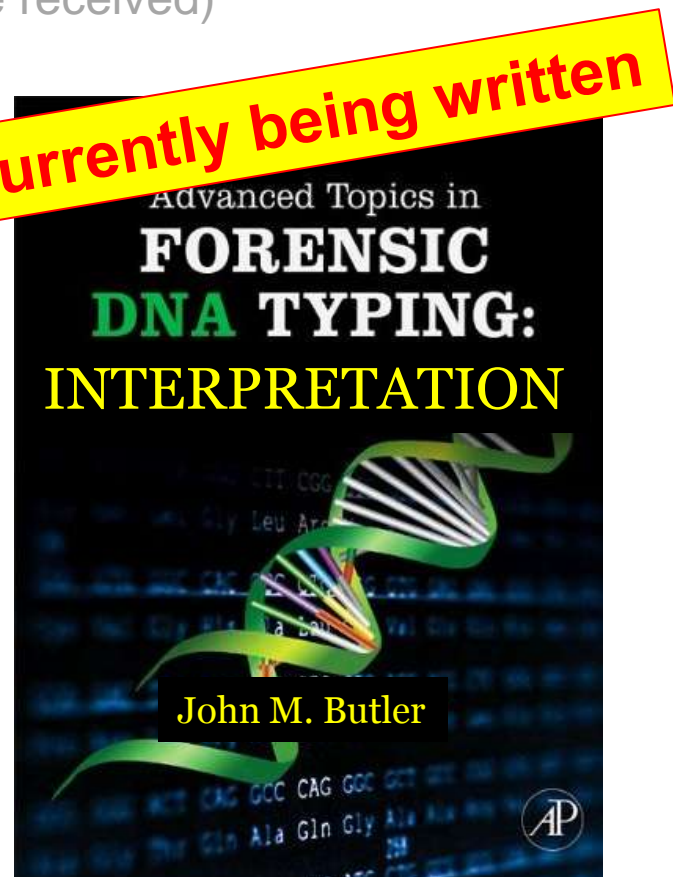
Sept 2009

~500 pages



August 2011

~700 pages



Fall 2012

~500 pages

Current NIST Projects

Short Overviews...

<http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm>

NIST SRM 2391c



Margaret Kline



Becky Hill

Main Points:

- Traceable physical reference materials to ensure accurate and comparable measurements between laboratories
- Helps meet ISO 17025 needs for traceability to a national metrology institute

- <http://www.nist.gov/srm>

- **SRM 2391c released Aug 2011**

Presentations/Publications:

- *Profiles in DNA* article (Sept 2011)
- ISFG 2011 and ISHI 2011 posters
- Forensic Sci. Int. Genet. Suppl. Ser. (2011)

<http://www.promega.com/resources/articles/profiles-in-dna>

The Latest and Greatest NIST PCR-Based DNA Profiling Standard: Updates and Status of...

The Latest and Greatest NIST PCR-Based DNA Profiling Standard: Updates and Status of Standard Reference Material® (SRM) 2391c

Article

Figures & Tables

✉ 🖨️ ➕ Share

Margaret C. Kline, Carolyn R. (Becky) Hill, Jamie L. Almeida, Erica L.R. Butts, Michael D. Coble and John M. Butler

National Institute of Standards and Technology, Applied Genetics Group, Gaithersburg, Maryland, USA
2011

NIST SRM 2391c

Selling since
Aug 16, 2011
Current price: \$614



Produced with an entirely new set of genomic DNA samples.

9947A & 9948 are NOT included.

https://www-s.nist.gov/srmors/view_detail.cfm?srm=2391C

Description of Components in SRM 2391c

| Component | Description | Quantity ^a |
|-----------|--|---------------------------|
| A | 50 μ L of anonymous female genomic DNA | 1.4 – 1.9 ng DNA/ μ L |
| B | 50 μ L of anonymous male genomic DNA | 1.3 – 1.5 ng DNA/ μ L |
| C | 50 μ L of anonymous male genomic DNA | 1.3 – 2.0 ng DNA/ μ L |
| D | 50 μ L of mixed-source (Components A and C) | 1.4 – 2.0 ng DNA/ μ L |
| E | Two 6 mm punches of CRL-1486 cells spotted on 903 paper | ~75,000 cells per punch |
| F | Two 6 mm punches of HTB-157 cells spotted on FTA paper | ~75,000 cells per punch |

^a DNA concentrations and cell counts are nominal values and are **not** intended for use as quantitative standards.

STR Genotyping kits and primer mixes used at NIST to certify SRM 2391c

| Kit Provider | | | Primer Mixes |
|--------------------------|--------------------|---------------|--------------|
| <i>Life Technologies</i> | <i>Promega</i> | <i>Qiagen</i> | <i>NIST</i> |
| Identifiler | Powerplex 16 | ESSplex | 26plex |
| Identifiler Plus | Powerplex 16 HS | IDplex | miniSTRs |
| NGM | Powerplex ESX 17 | | |
| NGM SElect | Powerplex ESI 17 | | |
| COfiler | Powerplex ES | | |
| Profiler | Powerplex S5 | | |
| Profiler Plus | Powerplex Y | | |
| Profiler Plus ID | FFFL | | |
| SGM Plus | | | |
| SEfiler | | | |
| MiniFiler | | | |
| Yfiler | | | |

All results are concordant across all kits.

In total there is data for 51 autosomal STRs and 17 Y-STRs

Insertion/Deletion (InDel) Markers

NIST



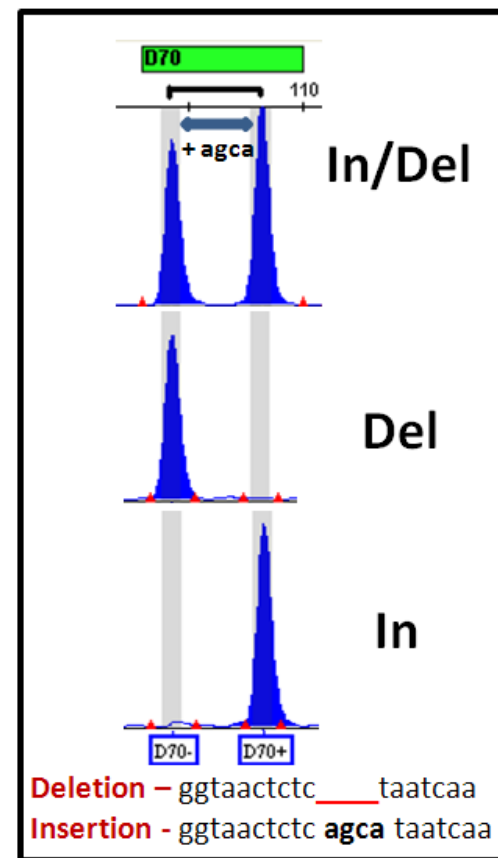
Manuel Fondevila
Alvarez
Guest Researcher
from Spain

Main Points:

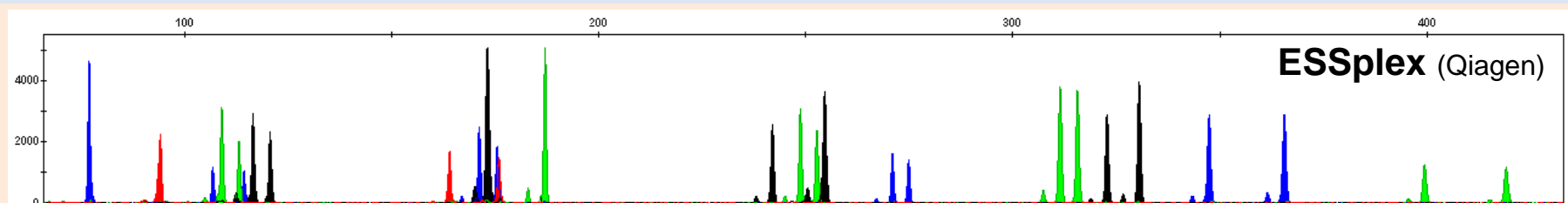
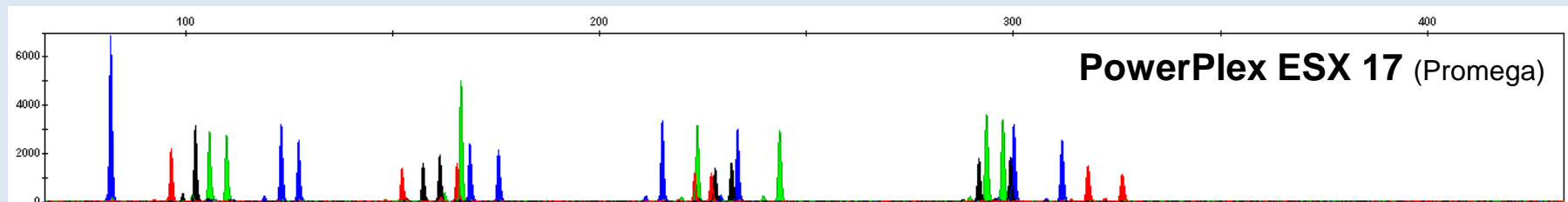
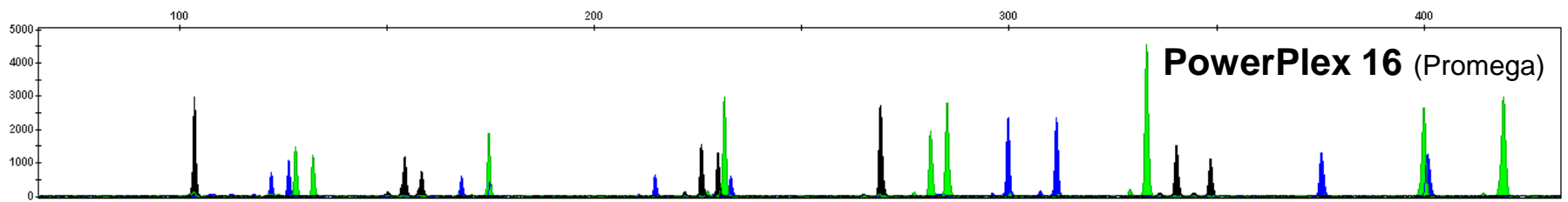
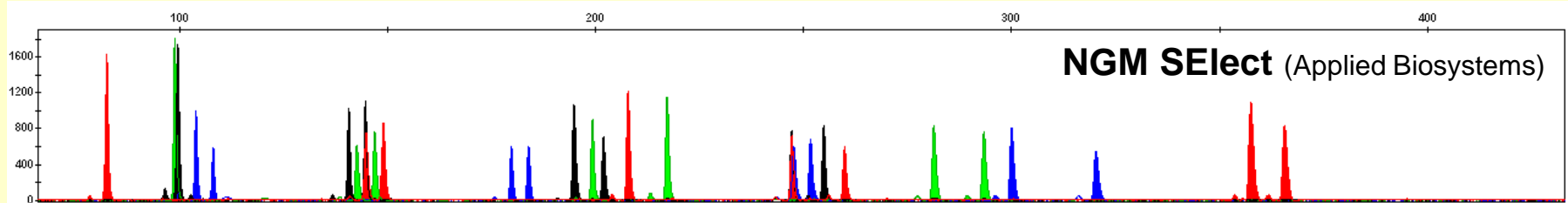
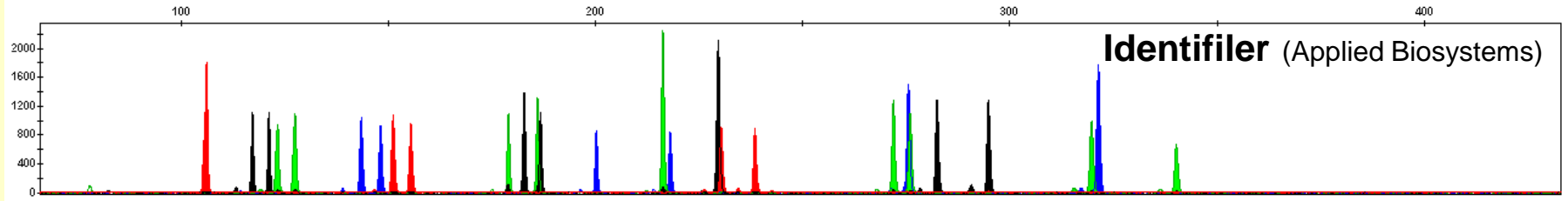
- InDels (insertion-deletion) or DIPs (deletion-insertion polymorphisms) are short length polymorphisms, consisting of the presence or absence of a short (typically 1-50 bp) sequence
- Like SNPs, InDels have low mutation rate (value to kinship analysis), small amplicon target sizes (value with degraded DNA), and can be highly multiplexed
- Can be analyzed on CE instruments like STRs
- Studied **commercial 30plex** (Qiagen DIplex) and a **home-brew 38plex** in U.S. population samples

Presentations/Publications:

- FSI Genetics Suppl. Series 2011 article
- ISFG 2011 poster and ISHI 2011 presentation



Same DNA Sample Tested with Five STR Kits



Kit Concordance Comparisons



Becky Hill

| <u>Kits compared</u> | <u>Samples</u> | <u>Loci compared</u> | <u>Comparisons</u> | <u># Differences</u> | <u>Concordance (%)</u> |
|----------------------|----------------|----------------------|--------------------|----------------------|------------------------|
| SGM-ID | 1436 | 11 | 15,796 | 1 | 99.994 |
| ID-ProPlus | 1427 | 10 | 14,270 | 1 | 99.993 |
| ID-IDplex | 669 | 16 | 10,704 | 19 | 99.822 |
| ID-PP16 | 662 | 14 | 9,268 | 4 | 99.957 |
| ID-MiniFiler | 1308 | 9 | 11,772 | 27 | 99.771 |
| SGM-NGM | 1436 | 11 | 15,796 | 4 | 99.975 |
| ID-NGM | 1449 | | | | |
| ProPlus-NGM | 1427 | | | | |
| SGM-ESI | 1436 | | | | |
| ProPlus-ESX | 1427 | | | | |
| ESI-ESX | 1455 | | | | |
| ESI-ESSplex | 1445 | | | | |
| ESX-ESSplex | 1445 | | | | |
| ESI-NGMSElect | 715 | | | | |

128 kit-to-kit comparisons
1,104,031 allele comparisons
1224 differences observed
~99.9% concordance
(many corrected now)

*Kits (except Identifiler) were kindly provided by **Applied Biosystems, Promega, and Qiagen** for concordance testing performed at NIST*

Recent Training Workshops



John Butler



Mike Coble



- AAFS (February 22, 2011)
 - **Mixture Interpretation (with 6 other speakers)**



- ISFG (August 30, 2011)
 - **CE Fundamentals and Troubleshooting**



- Int. Symp. Human Ident. (October 3, 2011)
 - **Mixture Interpretation (with Boston University)**



- Int. Symp. Human Ident. (October 6, 2011)
 - **Troubleshooting Laboratory Systems**

Slide handouts available at
<http://www.cstl.nist.gov/strbase/training.htm>

TrueAllele Mixture Software Evaluation



Mike Coble

Main Points:

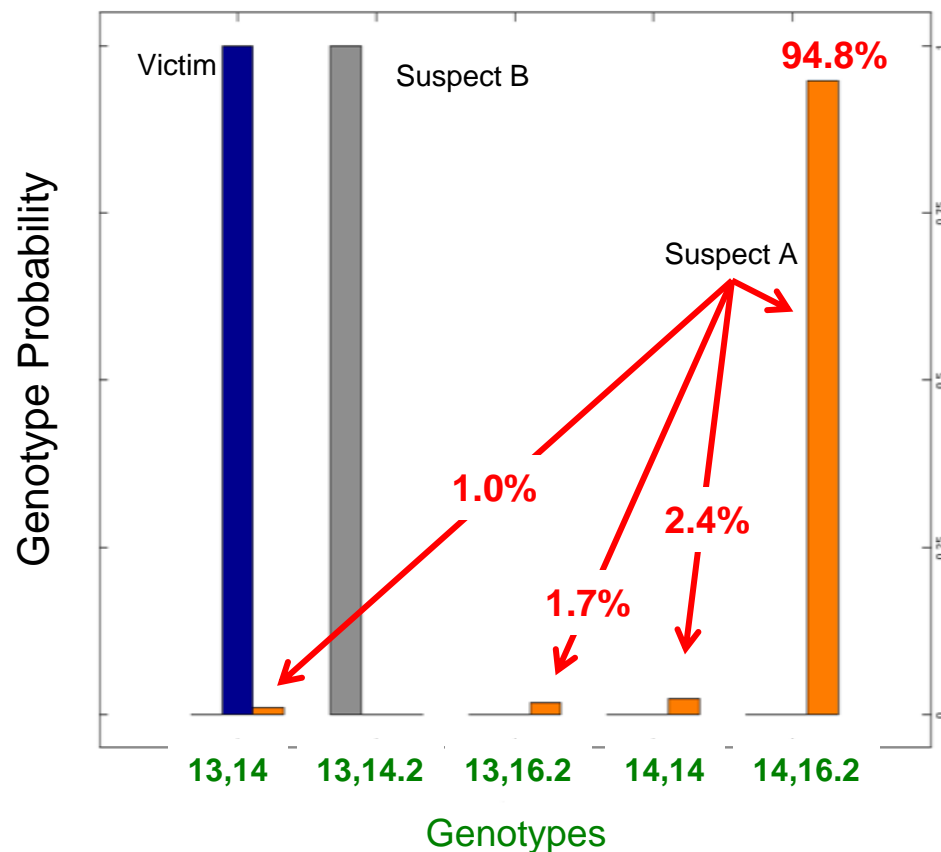
- Exploring the capabilities and limitations of a probabilistic genotyping approach
- Studying TrueAllele software with a number of different types of mixtures (including low-level and 3-4 person mixtures)
- Work being performed at NIST independently of Cybergenetics

Presentations/Publications:

- ISFG 2011 presentation
- ISHI 2011 mixture workshop

D19S433 result from one replicate of 50,000 simulations

3 person mixture conditioning on the victim



Rapid PCR and Rapid DNA Testing



Pete Vallone

Main Points:

- **Performing research on reducing the total time required for STR typing**
 - Focusing on the multiplex amplification of commercial STR kits with faster polymerases and thermal cyclers
 - Single-source reference samples (sensitivity > 200 pg)
- **Designing testing plans for rapid DNA typing devices**
 - NIST will be examining rapid DNA instruments with FBI collaboration
- **Exploring direct PCR protocols** with FTA and 903 papers

Presentations/Publications:

- Vallone et al. (2008) FSI Genetics - on rapid PCR
- ISFG 2011 and ISHI 2011 presentations by Tom Callaghan (FBI)
- ISFG 2011 presentation and poster on direct PCR

ABI 3500 Validation Studies



Erica Butts

Main Points:

- The 3500 has proven to be reliable, reproducible and robust in our hands – we have provided feedback to ABI to improve use
- Produces excellent DNA sequencing results
- Signal strength is different compared to ABI 3130xl and requires studies to set analytical and stochastic thresholds
- **Dye-specific analytical thresholds** resulted in less allelic and full locus dropout than applying one analytical threshold to all dyes
- RFID tracking decreases flexibility in our research experience

Presentations/Publications:

- MAAFS talk (May 2011)
- ABI road show talks (July & Aug 2011)
- ISFG presentation (Sept 2011)
- ISHI poster (Oct 2011)

ABI 3500 Open Letter Update



John Butler



Concerns Expressed in 3/31/11 Open Letter

1. RFID tags
2. New .hid file structure requires new software
3. Short shelf life of reagents – would like to see data for expiration times

At the Promega ISHI meeting (Oct 2011), ABI described data for studies around reagent expiration through a poster at their booth. Sailus, Wheaton, Fisher, Calandro. “Understanding the Consumables on the 3500 Genetic Analyzers in the context of a Human Identification (HID) Laboratory”

They have promised that **polymer and buffer expiration dates will no longer be a hard stop** but only a warning with the future Windows 7 software upgrade (3500 Data Collection v1.3).

Performance Assessment of PlexID



Kevin Kiesler Pete Vallone

Abbott Ibis Biosciences
PLEX-ID System

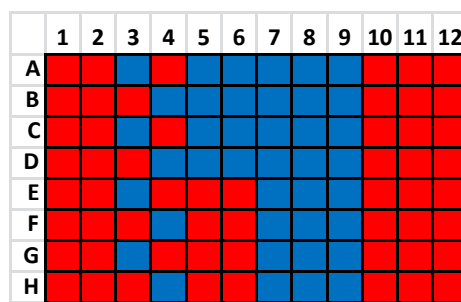


- **In collaboration with FBI**
- **Evaluating ESI-TOF mass spectrometer for mtDNA**
- Base composition of the control region determined from 8 triplex PCRs
- Started running the PlexID platform mid-October 2011
- Scheduled to complete experiments in February 2012

Contamination Check

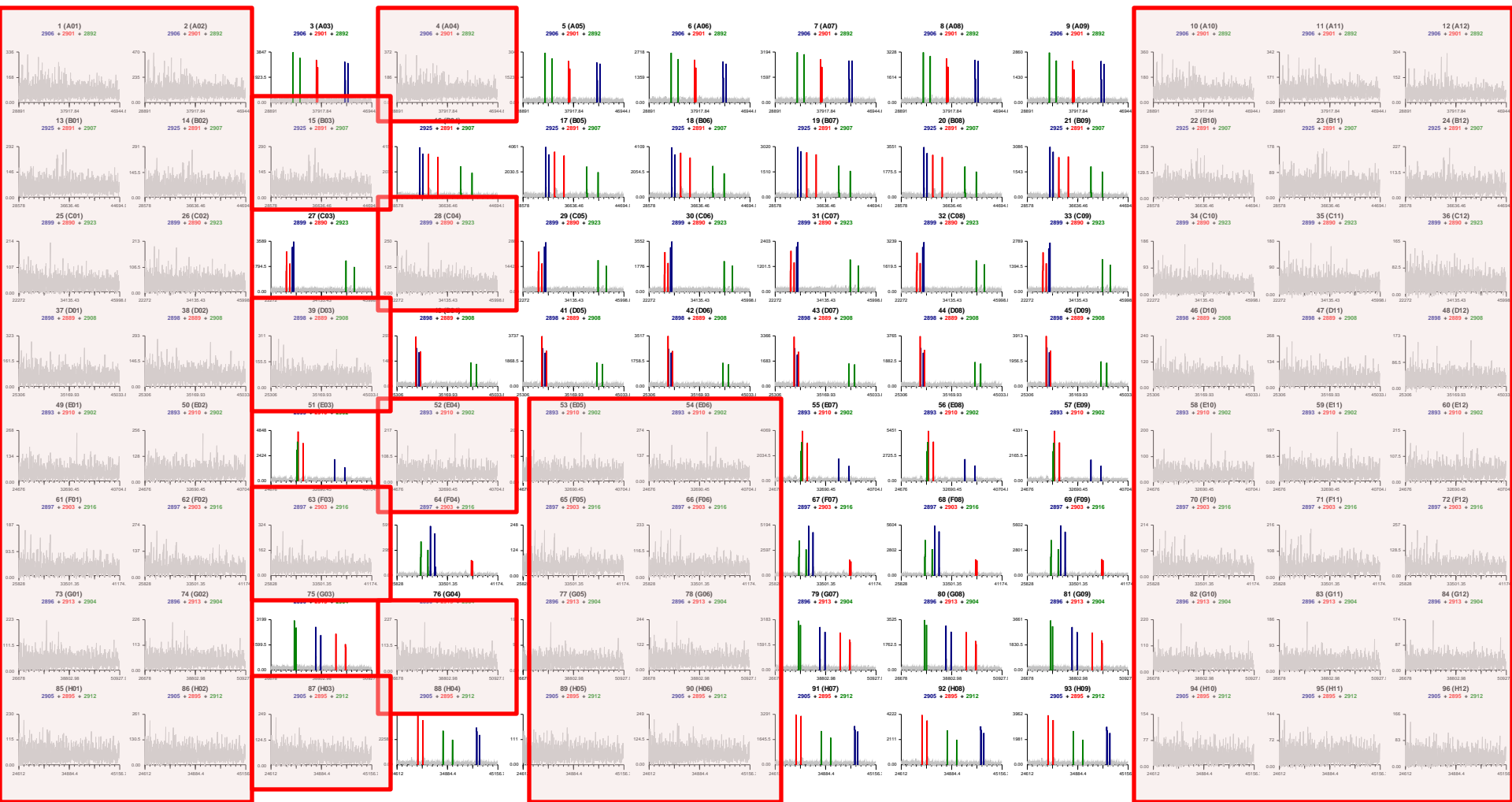
Checks run weekly on the PlexID to monitor baseline noise and potential contamination

No signal detected in 'red' wells



Red = empty well
Blue = positive control

Reagents Well-to-well Injector Desalter Carousel



PLEX-ID Evaluations Performed Thus Far...

| Experiment | Number of Plates | Number of Unique Samples |
|----------------------------------|------------------|----------------------------|
| Mixtures | 20 | 3 |
| Concordance | 33 | 247 |
| Sensitivity / Limit of Detection | 10 | 3 |
| Contamination | 8 | 1 |
| Total | 71 | 6816 wells examined |

- Mixtures can be detected with minor component present at 5-10%
- Concordance with Sanger sequencing (98.8%) (n=247)
- Limit of detection \approx 2.5 pg/well
- 1-2 plates run daily on the platform since mid-October

Future Projects Planned

- New book in progress on interpretation issues
- Additional mixture software evaluation
- Rapidly mutating Y-STR loci (European collaboration)
- More concordance testing with new STR kits
- Complete PLEX-ID mass spec validation with mtDNA base composition (FBI collaboration)
- Rapid DNA test device evaluation (FBI collaboration)
- Exploration of Next-Generation Sequencing
- Digital PCR for human DNA quantitation

Characterizing New STR Loci



John Butler



Becky Hill

Main Points:

- In April 2011, the FBI announced plans to expand the core loci for the U.S. beyond the current 13 CODIS STRs
- Our group is collecting U.S. population data on new loci and characterizing them to aid understanding of various marker combinations
- We are collecting all available information from the literature on the 24 commonly used autosomal STR loci

Presentations/Publications:

- AAFS 2011 presentation
- Hill et al (2011) *FSI Genetics* 5(4): 269-275
- Hares (2012) Expanding the U.S. core loci... *FSI Genetics* 6(1): e52-e54
- Butler & Hill (2012) *Forensic Sci Rev* 24(1): 15-26

Article in the January 2012 issue of *Forensic Science Review*

Available at <http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm>

Biology and Genetics of New Autosomal STR Loci Useful for Forensic DNA Analysis

REFERENCE: Butler JM, Hill CR: Biology and genetics of new autosomal STR loci useful for forensic DNA analysis; *Forensic Sci Rev* 24:15; 2012.

ABSTRACT: Short tandem repeats (STRs) are regions of tandemly repeated DNA segments found throughout the human genome that vary in length (through insertion, deletion, or mutation) with a core repeated DNA sequence. Forensic laboratories commonly use tetranucleotide repeats, containing a four base pair (4-bp) repeat structure such as GATA. In 1997, the Federal Bureau of Investigation (FBI) Laboratory selected 13 STR loci that form the backbone of the U.S. national DNA database. Building on the European expansion in 2009, the FBI announced plans in April 2011 to expand the U.S. core loci to as many as 20 STRs to enable more global DNA data sharing. Commercial STR kits enable consistency in marker use and allele nomenclature between laboratories and help improve quality control. The STRBase website, maintained by the U.S. National Institute of Standards and Technology (NIST), contains helpful information on STR markers used in human identity testing.

Key Words: Autosomal genetic markers, CODIS STRs, core loci, DNA typing, European Standard Set, expanded U.S. core loci, short tandem repeat (STR), STR kits.

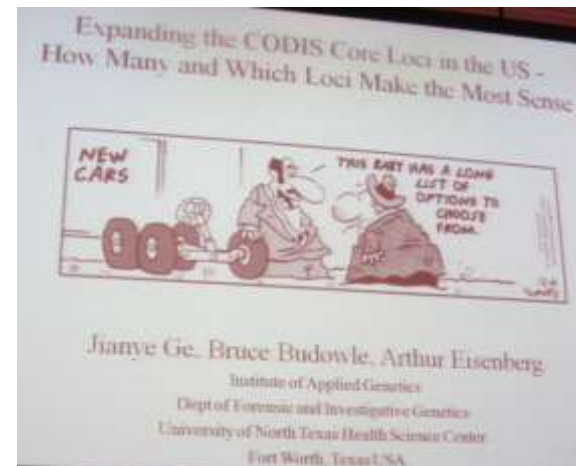
Discusses the 24 autosomal STR loci available in commercial kits

What concerns have been raised?

Recent Public Criticism of Efforts to Expand the CODIS Core Loci in the U.S.



October 4, 2011 Presentation at Promega's International Symposium on Human Identification



October 16, 2011 Follow-up article by BBC News on ISHI presentation



<http://ishinews.com/>

FBI's DNA Database Upgrade Plans Come Under Fire

© OCTOBER 16, 2011 9:00 AM FEATURED ARTICLE, GENERAL SESSION UPDATES, RECENT ARTICLE

By: Paul Rincon, Science editor Published October 16, 2011 by the BBC News Website A major upgrade of the Federal Bureau of Investigation's (FBI) DNA database system has come under fire from members of the forensic science community. The Codis system is used to generate the genetic profiles stored in the US national DNA [...]

<http://www.bbc.co.uk/news/science-environment-15311718>

Recent Publication by Budowle et al. Summarizing Criticisms Raised in ISHI Talk

Jianye Ge, Arthur Eisenberg, Bruce Budowle
Investigative Genetics 2012, 3:1 (6 January 2012)

**Developing criteria and data to determine best options for
expanding the core CODIS loci**

Research [Open Access](#)

[Developing criteria and data to determine best options for expanding the core CODIS loci](#)

Jianye Ge, Arthur Eisenberg, Bruce Budowle

Investigative Genetics 2012, 3:1 (6 January 2012)

[Abstract](#) | [Provisional PDF](#) | [PubMed](#) | [Editor's summary](#)

The rapid expansion of DNA databases over the past few years has led to numerous problems associated with their use. In order to increase the efficiency of the Combined DNA Index System, expansions to the core loci have been suggested. Budowle et al evaluate these proposed expansions and consider the consequences of the selection of core genetic markers for forensic DNA databases.

Available at <http://www.investigativegenetics.com>

Concerns Raised in Public Criticisms of Expanded CODIS Core Loci Selection

- **Not enough data behind decisions** – need more community involvement rather than a small committee making decisions
- **Casework needs** should drive decisions
- **Large loci fail in casework samples and should be avoided** – miniSTR capabilities are preferred
- **Large multiplexes** may adversely impact performance
- **DYS391 is a poor choice** and AMEL Y nulls are not a significant concern
- **Y-STRs should be included** as core loci to benefit familial searches of the future
- **No definition of performance goals** are provided

What data exist behind decisions made so far and what additional data are there for consideration to help address concerns raised?

The 11 STR Loci Beyond the CODIS 13

5 new European loci

| STR Locus | Location | Repeat Motif | Allele Range* | # Alleles* |
|----------------|----------|--------------|---------------|------------|
| D2S1338 | 2q35 | TGCC/TTCC | 10 to 31 | 40 |
| D19S433 | 19q12 | AAGG/TAGG | 5.2 to 20 | 36 |
| Penta D | 21q22.3 | AAAGA | 1.1 to 19 | 50 |
| Penta E | 15q26.2 | AAAGA | 5 to 32 | 53 |
| D1S1656 | 1q42 | TAGA | 8 to 20.3 | 25 |
| D12S391 | 12p13.2 | AGAT/AGAC | 13 to 27.2 | 52 |
| D2S441 | 2p14 | TCTA/TCAA | 8 to 17 | 22 |
| D10S1248 | 10q26.3 | GGAA | 7 to 19 | 13 |
| D22S1045 | 22q12.3 | ATT | 7 to 20 | 14 |
| SE33 | 6q14 | AAAG‡ | 3 to 49 | 178 |
| D6S1043 | 6q15 | AGAT/AGAC | 8 to 25 | 25 |

*Allele range and number of observed alleles from Appendix 1, J.M. Butler (2012) *Advanced Topics in Forensic DNA Typing: Methodology*; ‡SE33 alleles have complex repeat structure

Concern: Large loci fail in casework samples and should be avoided – miniSTR capabilities are preferred

- We agree that miniSTRs (smaller amplicons) work best with degraded DNA that is often present in casework samples
- **How often are high molecular weight loci failing?**
- What data exist on success rates of loci for profiles stored in Forensic Index of CODIS based on PCR product size?

Palm Beach Sheriff's Office Crime Lab

LDIS Forensic Unknowns – PowerPlex 16 data

Single-source

2,452 profiles total

- Loss of Penta D: 633
- Loss of Penta E: 323
- Loss of FGA: 202
- **Loss of all 3 loci: 130**

$$130/2452 = 5.3\%$$

FGA loss = 8.2%

Mixtures

841 profiles total

- Loss of Penta D: 297
- Loss of Penta E: 296
- Loss of FGA: 179
- **Loss of all 3 loci: 55**

$$55/841 = 6.5\%$$

FGA loss = 21.3%

Larger loci are lost in a fraction of casework samples...

Additional Data from VA and CA

| | Virginia | California |
|--|-----------------------|------------------------|
| # Forensic Unknowns (single-source profiles) | 13,488 | 37,024 |
| No FGA (largest of current CODIS 13 core loci) | 68 (0.5%) | 1,936 (4.5%) |
| Profiles missing at least one locus | 1,609 (12%) | 4,440 (12%) |

Data courtesy of George Li, Brad Jenkins, and Ken Konzak

Will Performance with Large Multiplexes Be Adversely Impacted with Additional Loci Added?

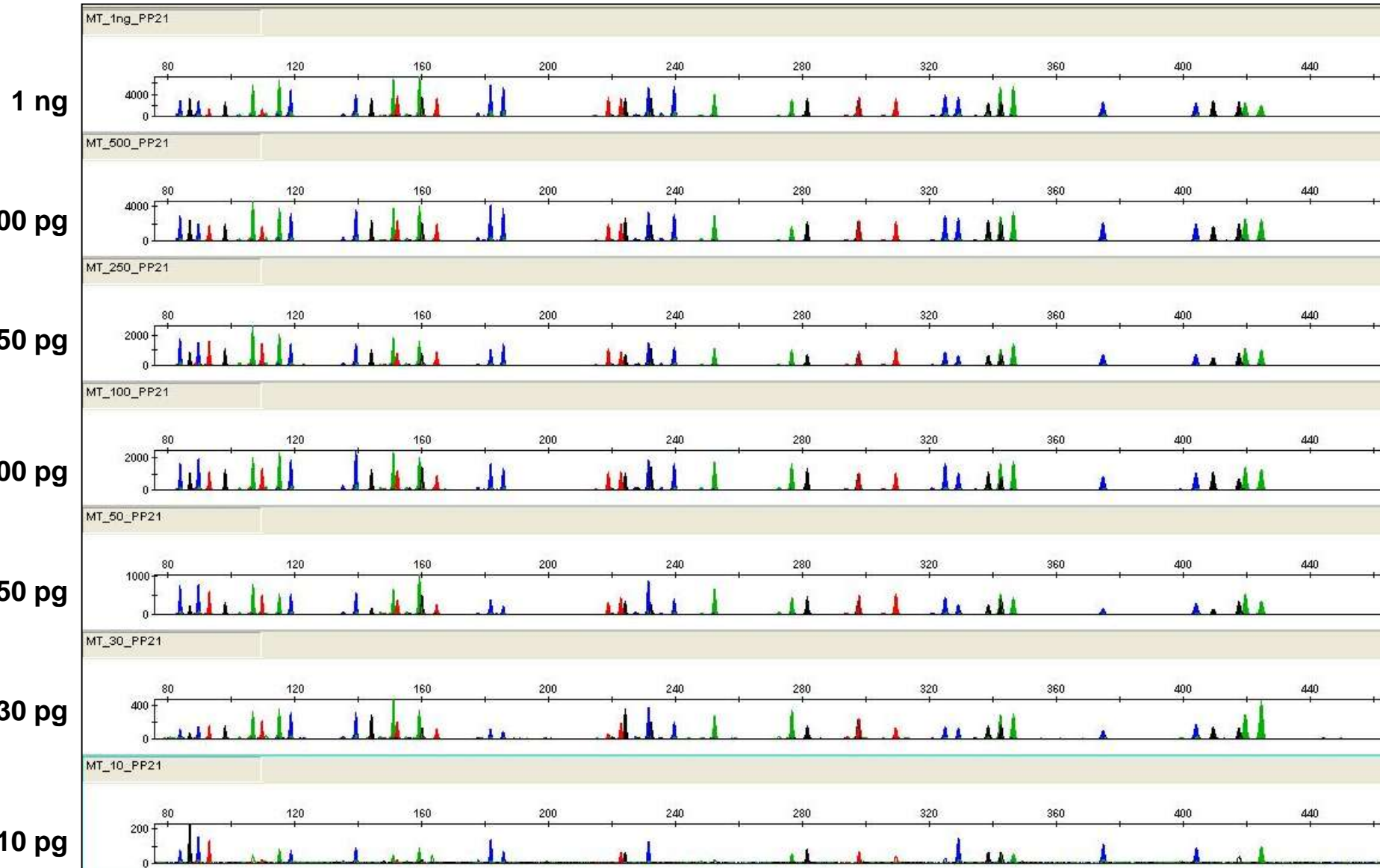
- **There has been significant improvement in kit development in recent years**
 - In addition, 6-dye capability of ABI 3500 instruments may play a role in future kits...
- What assay or kit data exist with 20plex (or greater) STR multiplexes?
 - NIST 26plex
 - PowerPlex 21 data collected at NIST

PowerPlex 21

- Promega STR kit to be released in early 2012
 - NIST has been working with this kit since spring 2011 primarily for concordance testing and has permission from Promega to discuss results
- **Contains 20 autosomal STRs + amelogenin**
- **Enables examination of performance characteristics** similar to a future U.S. megaplex containing at least 20 loci

DNA Dilution Series with PowerPlex 21

As expected with any STR kit/assay, allele dropout occurs below 100 pg...



Measurement of Allele Dropout and Extreme Peak Height Imbalance for 2 STR Kits

Three fully heterozygous (except PT83 at Penta D) pristine DNA samples were examined in a dilution series with PowerPlex 21 and Identifiler Plus. Results are ordered by amplicon size and dye color.

PowerPlex 21 - 30 cycles (5s@2kV)

| Sample | D | D | D | D | P | D | D | P | D | D | D | D | D | D | DNA amount | | | | | | | |
|--------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|------------|---|---|-------|---|---|---|------|
| MT | 3 | 1 | 6 | 1 | E | 1 | D | 2 | C | E | | D | D | D | 8 | 1 | 1 | 10 pg | | | | |
| PT83 | S | S | S | 3 | N | 6 | 1 | S | S | N | | 2 | 7 | 5 | S | 2 | 9 | | | | | |
| PT84 | A | 1 | 1 | 1 | S | T | S | 8 | 1 | F | T | T | 1 | S | S | T | 1 | | S | S | | |
| MT | 3 | 6 | 0 | 3 | A | 5 | S | 3 | 1 | A | H | v | S | 8 | 8 | P | 1 | | 3 | 4 | F | |
| PT83 | E | 5 | 5 | 4 | 1 | 3 | 5 | 3 | P | | 0 | W | 1 | 2 | 1 | O | 7 | | 9 | 3 | G | |
| PT84 | L | 8 | 6 | 3 | 7 | E | 9 | 1 | 8 | O | D | 1 | A | 1 | 0 | 8 | X | 9 | 1 | 3 | A | 1 ng |
| MT | | | | | | | | | | | | | | | | | | | | | | |
| PT83 | | | | | | | | | | | | | | | | | | | | | | |
| PT84 | | | | | | | | | | | | | | | | | | | | | | |
| MT | | | | | | | | | | | | | | | | | | | | | | |
| PT83 | | | | | | | | | | | | | | | | | | | | | | |
| PT84 | | | | | | | | | | | | | | | | | | | | | | |
| MT | | | | | | | | | | | | | | | | | | | | | | |
| PT83 | | | | | | | | | | | | | | | | | | | | | | |
| PT84 | | | | | | | | | | | | | | | | | | | | | | |
| MT | | | | | | | | | | | | | | | | | | | | | | |
| PT83 | | | | | | | | | | | | | | | | | | | | | | |
| PT84 | | | | | | | | | | | | | | | | | | | | | | |

Identifiler Plus - 28 cycles (10s@3kV)

| Sample | D | D | D | D | D | D | D | D | D | D | D | D | D | D | DNA amount | | |
|--------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|------------|---|------|
| MT | 8 | D | D | C | 3 | 1 | 1 | 2 | 1 | | | | D | D | 10 pg | | |
| PT83 | S | 2 | 7 | S | S | 3 | 6 | S | 9 | | | | 1 | 5 | | | |
| PT84 | 1 | 1 | S | F | 1 | T | S | S | 1 | S | | | T | 8 | | A | S |
| MT | 1 | S | 8 | 1 | 3 | H | 3 | 5 | 3 | 4 | v | P | S | M | | 8 | F |
| PT83 | 7 | 1 | 2 | P | 5 | 0 | 1 | 3 | 3 | 3 | W | O | 5 | E | | 1 | G |
| PT84 | 9 | 1 | 0 | O | 8 | 1 | 7 | 9 | 8 | 3 | A | X | 1 | L | 8 | A | 1 ng |
| MT | | | | | | | | | | | | | | | | | |
| PT83 | | | | | | | | | | | | | | | | | |
| PT84 | | | | | | | | | | | | | | | | | |
| MT | | | | | | | | | | | | | | | | | |
| PT83 | | | | | | | | | | | | | | | | | |
| PT84 | | | | | | | | | | | | | | | | | |
| MT | | | | | | | | | | | | | | | | | |
| PT83 | | | | | | | | | | | | | | | | | |
| PT84 | | | | | | | | | | | | | | | | | |
| MT | | | | | | | | | | | | | | | | | |
| PT83 | | | | | | | | | | | | | | | | | |
| PT84 | | | | | | | | | | | | | | | | | |

Legend
Green = both alleles detected
Light Green = PHR <60%
Yellow = allele dropout
Red = locus dropout
Black = allele drop-in

Having 5 additional loci did not adversely impact success rates

Total alleles possible = 875
 Total alleles present = 805

92% detected

Total alleles possible = 672
 Total alleles present = 619

92% detected

Concern: DYS391 is a poor choice and AMEL Y nulls are not a significant concern

- AMEL Y nulls happen...
 - Common practice in some labs is a follow-up test with Y-STRs to confirm that a sample is male
 - **Some labs have implemented an additional ChrY test (SRY) to confirm AMEL Y nulls**
- A further purpose of having a single Y-STR is to aid QC checks if further Y-STR testing is performed for familial searching or casework purposes
 - DYS391 result will enable a QC check to Yfiler or PowerPlex Y results like D3 and D7 did for Profiler Plus/COfiler (albeit a rather weak one because it is not very polymorphic)
 - By itself, DYS391 is not polymorphic enough to be helpful with any potential familial search filter

SRY Male-Specific Amplicon Used to Aid with Amelogenin Null Detection

Journal of Forensic Sciences, May 2009, 54(3): 551-555

TECHNICAL NOTE

J Forensic Sci, May 2009, Vol. 54, No. 3
doi: 10.1111/j.1556-4029.2009.01007.x
Available online at: www.blackwell-synergy.com

Vanja Kastelic,¹ B.S.; Bruce Budowle,² Ph.D.; and Katja Drobnič,¹ Ph.D.

Validation of *SRY* Marker for Forensic Casework Analysis

¹Forensic Science Centre, Ministry of the Interior, Vodovodna 95, Ljubljana, Slovenia.
²FBI Laboratory, Quantico, VA 22135.
Received 13 Mar. 2008; and in revised form 15 May 2008; accepted 14 July 2008.

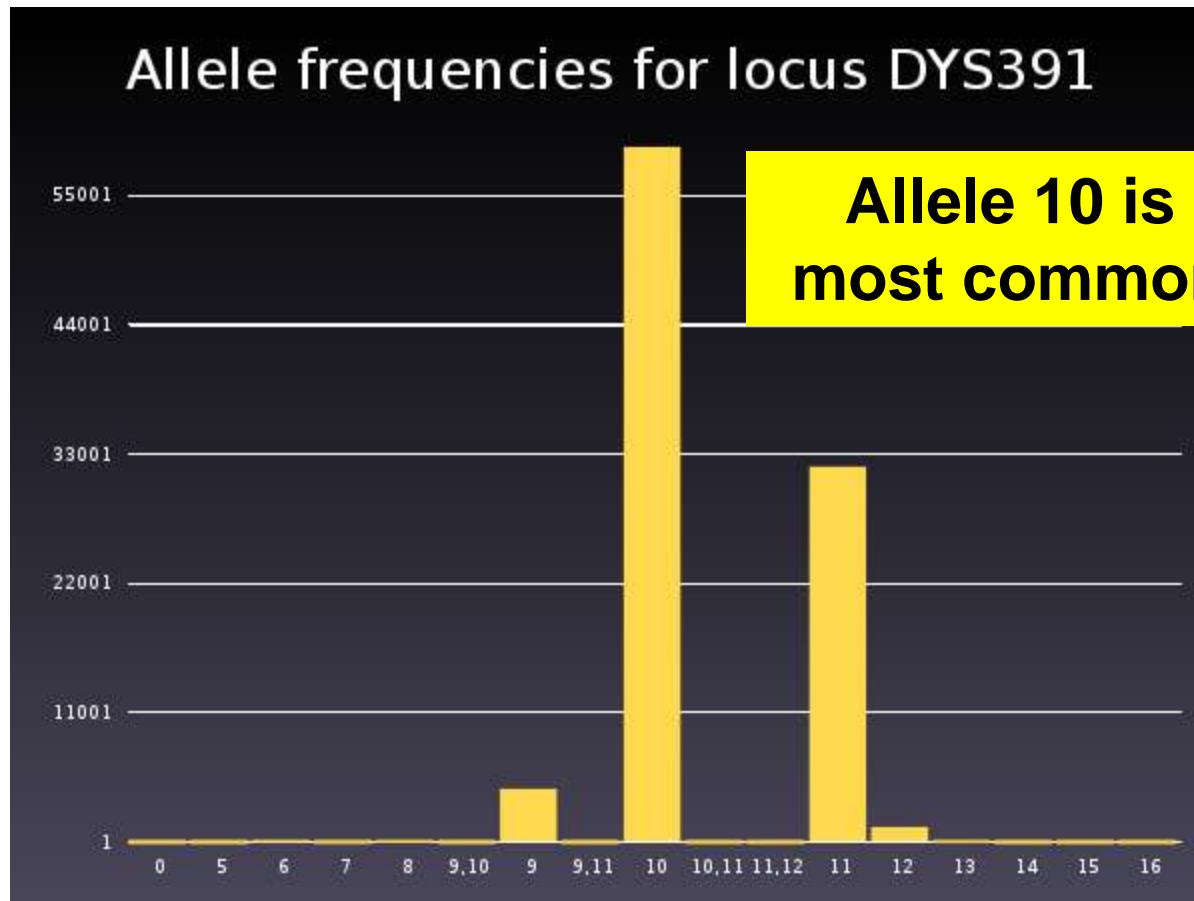
“Determining the gender of the source of forensic DNA evidence is based on the amelogenin test. However, at times the assay may not be indicative of gender assignment, because of deletions at the amelogenin site. ...The study herein addresses the validation of primers for the target SRY gene regarding specificity, sensitivity, and robustness.”

Why Consider DYS391?

- **DYS391 is located on the long arm of the Y-chromosome over 7 Mb away from amelogenin.** Thus, it is likely to be detected in the event of an amelogenin Y deletion that could make a male sample falsely appear as a female (X,-).
- DYS391 is not very polymorphic. From a data set of 97,575 haplotypes available on the Y-Chromosome Haplotype Reference Database, over half of them possess allele 10. However, only two null alleles have been reported and 0.01% duplication events (11 total) have been seen in over 700 different population groups from around the world. Thus, **it is a stable locus with a relatively narrow allele range.**
- DYS391 has a mutation rate of 0.26%, which is comparable to most autosomal STRs commonly in use. There have been 38 mutations observed so far in the 14,621 meioses reported in the literature and compiled on YHRD.

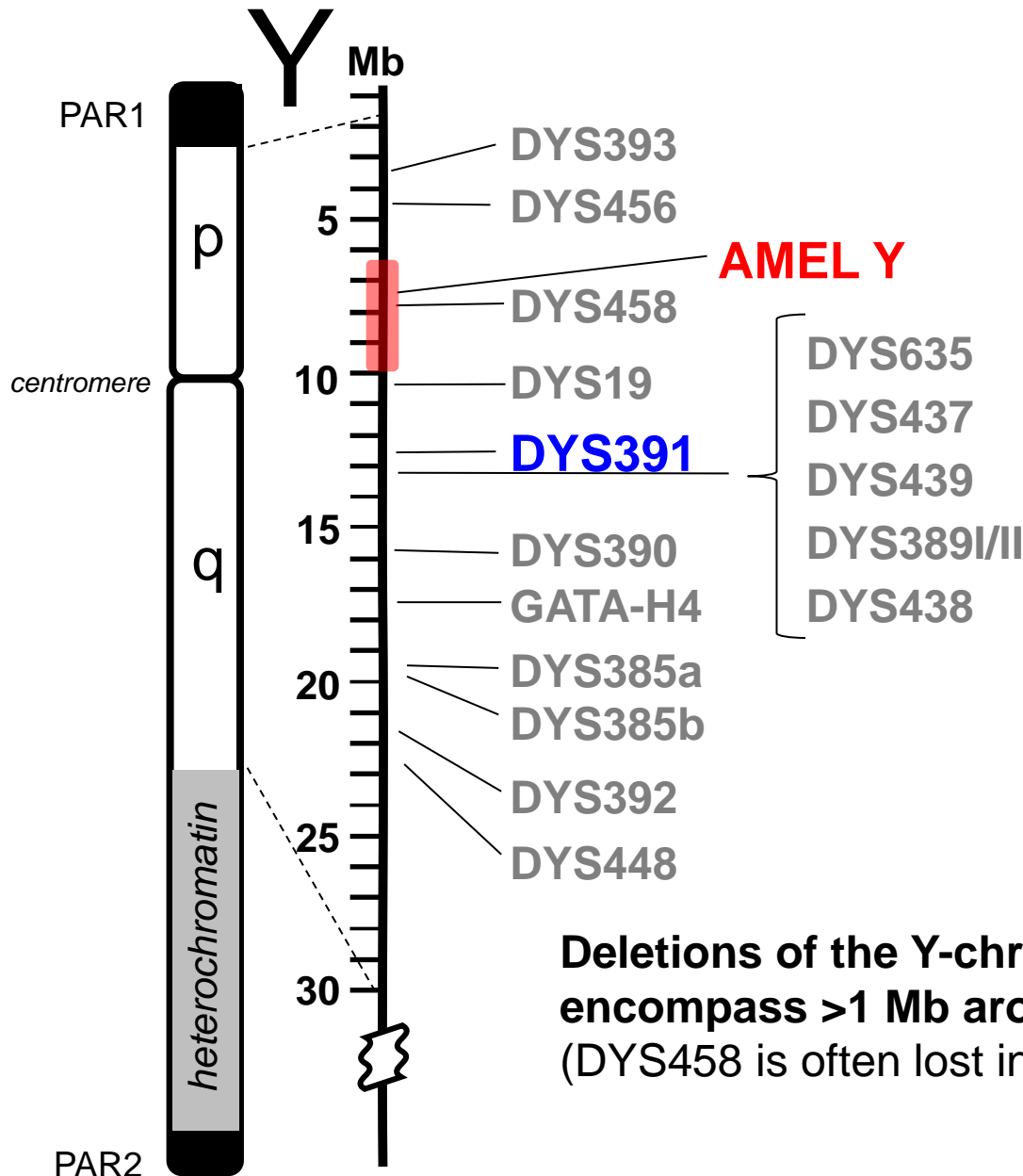
DYS391 Variability

YHRD (Y-chromosome Haplotype Reference Database) **data from 97,575 samples**



<http://www.yhrd.org/Research/Loci/DYS391>

Relative positions of 17 Y-STR loci commonly used in ChrY testing



Most commonly seen in males of Indian subcontinent origin

Chang *et al.* (2007) *Forensic Sci. Int.* 166: 115-120
12/649 Malaysian males showed no AMEL Y

Cadenas *et al.* (2007) *Forensic Sci. Int.* 166: 155-163
5/77 Nepal males showed no AMEL Y

Deletions of the Y-chromosome can encompass >1 Mb around the AMEL Y region (DYS458 is often lost in these situations)

Why mixing Y-STRs and autosomal STRs in a single DNA test is a bad idea...

Offender/arrestee reference samples

- Male samples: will work fine
- **Female samples: only autosomal STRs will amplify resulting in a waste of reagents compared to match probability produced**

Casework samples

- Mixtures: excess of either male or female DNA will result in poor STR typing results

Missing person samples

- Y-STRs will fail to work on female DNA samples

Do females represent a significant portion of the samples being examined in these specimen categories?

Not all DNA samples tested are male...

And if not male, then Y-STRs fail to amplify!

SDIS Offender/Arrestee Data:

- Virginia (371,000): ~**22% female***
 - California (1.9 million): ~**17% female***
 - Illinois (463,000): ~**16% female***
- *Determined to be female based on amelogenin results or meta data

Data kindly provided by George Li & Brad Jenkins, Ken Konzak, and Taylor Scott

Missing Persons:

- **NamUs** (<http://www.namus.gov/>; searched 11/4/11):
 - Unidentified persons: **20% female** (1699/8438 cases)
 - Missing persons: **36% female** (3278/9012 cases)
- **NDIS Statistics** (Aug 2011):
 - Unidentified human remains: 5,324
 - Missing person cases: 1,039

Per NDIS Custodian (11/4/11):
~45% females in MP cases
(by amelogenin results)

Summary

- It is vital that an expanded set of core loci be carefully considered and implemented to avoid adventitious hits on large and growing DNA databases.
- There is limited “electrophoretic real-estate” in constructing STR multiplex assays that will work in 5-dye instruments and contain PCR products <500 bp – 6-dye kits and instruments will help.
- The number of females in DNA database and missing persons cases make required use of multiple Y-STRs of questionable utility.
- **Data driven decisions are being made by the CODIS Core Loci Working Group.**
- The CODIS community will be involved in the implementation phase of adding new kits.

Thank you for your attention

Acknowledgments: Applied Biosystems, Promega, and Qiagen for STR kits used in concordance studies

Contact Information

John Butler

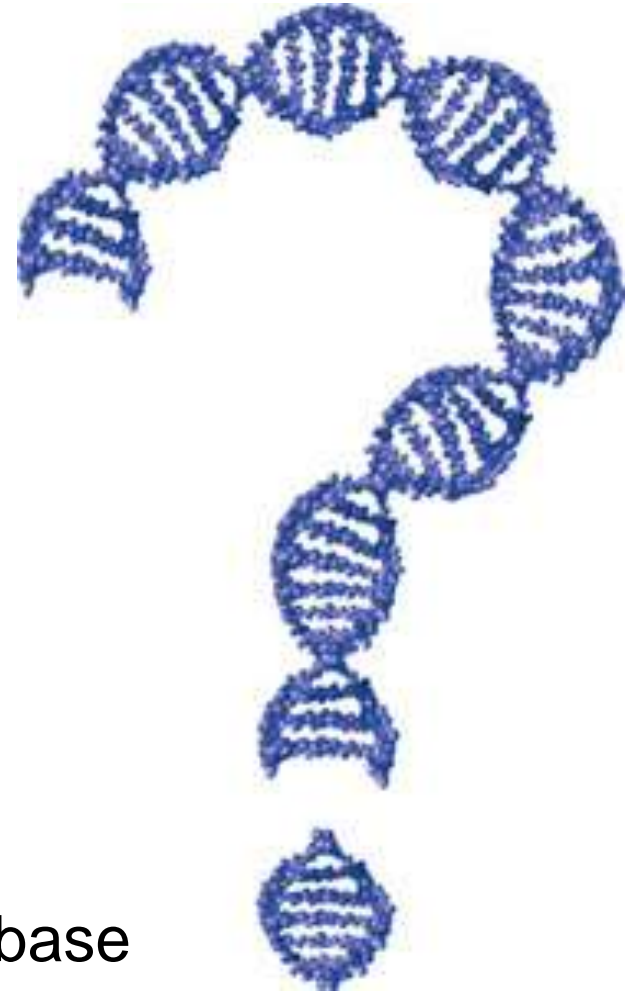
NIST Fellow

Group Leader of Applied Genetics

john.butler@nist.gov

301-975-4049

<http://www.cstl.nist.gov/biotech/strbase>



Our team publications and presentations are available at:
<http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm>