**EDNAP Meeting**
**31 October 2012 – Brussels, Belgium**

# NIST Update

**John M. Butler**

NIST Human Identity Project Team
National Institute of Standards and Technology
Gaithersburg, Maryland USA

# NIST Human Identity Project Teams
## within the Applied Genetics Group

### *Forensic DNA Team*

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

John Butler

STRBase, Workshops & Textbooks

Mike Coble

Mixtures, mtDNA & Y

Becky Hill

Concordance & LT-DNA

Margaret Kline

SRM work, variant alleles & Cell Line ID

**Office Manager**
Patti Rohmiller

### Guest Researcher

Manuel **Fonde**vila Alvarez

### *Data Analysis Support*

Dave Duewer

### *DNA Biometrics Team*

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

Pete Vallone

Rapid PCR, Direct PCR & Biometrics

Erica Butts

ABI 3500 & DNA Extraction

Kevin Kiesler

PLEX-ID & NGS Exploration

http://www.cstl.nist.gov/strbase/NISTpub.htm

NIST
Applied Genetics

# APPLIED GENETICS Group

## *Major Programs Currently Underway*

- **Forensic DNA**
  - STRBase website
  - New loci and assays (26plex)
  - **STR kit concordance**
  - Ancestry SNP assays
  - Low-template DNA studies
  - **Mixture interpretation research and training**
  - STR nomenclature
  - Variant allele cataloging and sequencing
  - ABI 3500 validation
  - Training workshops to forensic DNA laboratories
  - Validation experiments, information and software tools
  - **Textbooks – 3rd ed.** (3 volumes)
- **Clinical Genetics**
  - Huntington's Disease SRM
  - CMV SRM
  - Exploring future needs
- **DNA Biometrics**
  - **Rapid PCR methods**
  - Testing of rapid DNA systems
  - Plex-ID mtDNA base composition
- **Cell Line Authentication**
  - **ATCC documentary standard** (Margaret Kline & John Butler served on this international committee)

# Aiding Cell Line Authentication

Katsnelson, A. (2010) *Nature News*, 465: 537 (3 June 2010)

## Biologists tackle cells' identity crisis

DNA fingerprinting scheme aims to make sure researchers are working on the right cells.

Ever since biologists learned how to grow human cells in culture half a century ago, the cells have been plagued by a problem of identity: many commonly used cell lines are not actually what researchers think they are.

Cell-line misidentification has led to mistakes in the literature, misguided research based on those results and millions wasted in grant money. Last year, *Nature* described the situation as a scandal[1].

But a universal system for determining the identity of cell lines may now be in view. Next month, a working group led by the American Type Culture Collection (ATCC), a nonprofit biological repository based in Manassas, Virginia, that stores 3,600 cell lines from more than 150 species, plans to unveil standard-

ATCC
LEADING BIOLOGICAL STANDARDS

ATCC® Standards Development Organization

**Designation: ASN-0002**

**Authentication of Human Cell Lines: Standardization of STR Profiling**

The working group, composed of representatives from academia, government and industry,

a universally accepted approach will allow different facilities to compare their cell lines with each other, he adds.

Fingerprinting has its limits, cautions Michael Johnson, a cancer researcher at Georgetown University in Washington DC. "Just because a cell fingerprints out as the same [as another cell] doesn't mean they will behave the same," he says, noting that a cell's properties can also be affected by the way it has been grown, the number of times it has been cultured anew and small genetic changes that wouldn't show up in a fingerprint test. One classic example, he notes, is an immortalized breast cell line called MCF10A, which can form organized hollow structures similar to those found in mammary tissue; MCF10A cells currently distributed by

http://www.nature.com/news/2010/100602/pdf/465537a.pdf

# Highlights Since Last EDNAP Meeting

- InDel work published
- PLEX-ID report available
- TrueAllele evaluation continues…
- New DNA mixture training materials
- New autosomal STR and Y-STR loci & kits
  - NIST U.S. population data set completed
- SRM 2372 recertification underway
- Rapid DNA efforts
- *Interpretation* book being written

# Insertion/Deletion (InDel) Markers

Manuel **Fonde**vila Alvarez

**Guest Researcher from Spain (Jan 2011 to July 2012)**

**Main Points**:

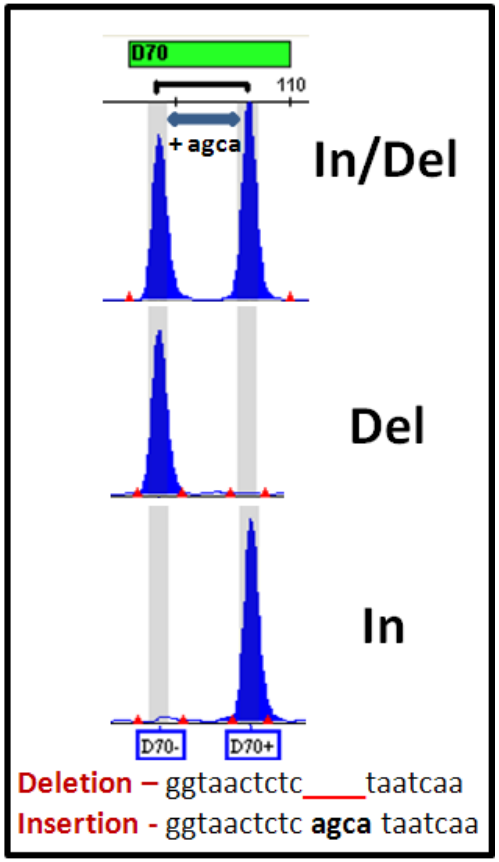
- InDels (insertion-deletion) or DIPs (deletion-insertion polymorphisms) are short length polymorphisms, consisting of the presence or absence of a short (typically 1-50 bp) sequence
- Like SNPs, InDels have low mutation rate (value to kinship analysis), small amplicon target sizes (value with degraded 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**

Int J Legal Med (2012) 126:725–737
DOI 10.1007/s00414-012-0721-7

ORIGINAL ARTICLE

**Forensic performance of two insertion–deletion marker assays**

M. Fondevila · C. Phillips · C. Santos · R. Pereira · L. Gusmão · A. Carracedo · J. M. Butler · M. V. Lareu · P. M. Vallone



# Performance Assessment of Plex-ID

Kevin Kiesler

Abbott Ibis Biosciences
**Plex-ID System**


NIST Report to the FBI:
Plex-ID Electrospray Time-of-Flight Mass Spectrometer for Mitochondrial DNA Base Composition Profiling

*Experiments performed and report written by: Kevin Kiesler, M.S. (NIST)*

*Under the direction of: Dr. Peter Vallone (NIST)*

- **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 Plex-ID platform mid-October 2011
- **136 page NIST report available on STRBase**

**http://www.cstl.nist.gov/strbase/pub_pres/NIST-report-on-PlexID.pdf**

# TrueAllele Mixture Software Evaluation

Mike Coble

**Main Points**:

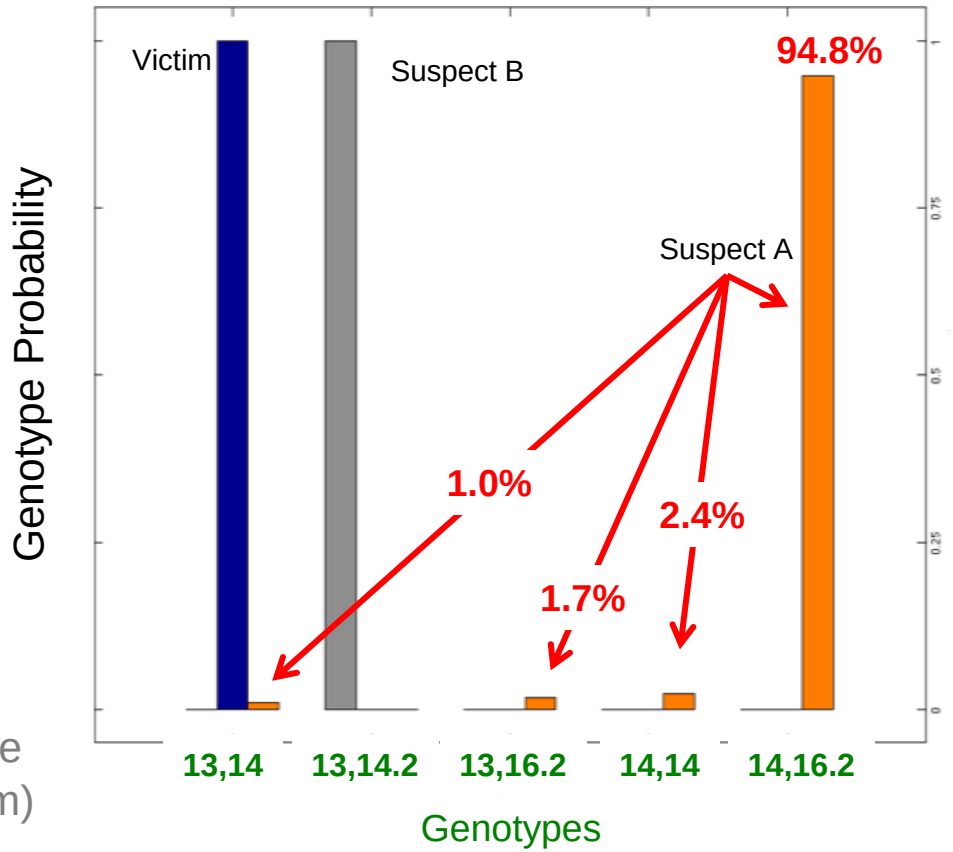
- Exploring the capabilities and limitations of a probabilistic genotyping approach
- Studying TrueAllele software with a number of different types of mixtures (including low-level and 3-4 person mixtures)
- Work being performed at NIST independently of Cybergenetics

**Presentations/Publications**:

- ISFG 2011 presentation
- Numerous mixture workshop talks (see http://www.cstl.nist.gov/strbase/mixture.htm)

D19S433 result from one replicate of 50,000 simulations

*3 person mixture conditioning on the victim*




See also Perlin et al. (2011) Validating TrueAllele DNA mixture interpretation. *J. Forensic Sci.* 56(6):1430-1447

# Mixture Training Workshops


John Butler


Mike Coble

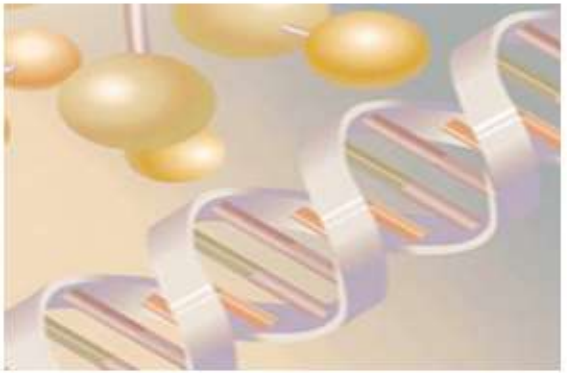

- Collaborators from Boston University (formerly Cellmark)
- ISHI 2012 workshop covered issues with thresholds, statistics, probabilistic genotyping, complex mixtures, court testimony, and assumptions made
  - Audience response systems (clickers) used to gather data from participants
- Slides are available on STRBase

**http://www.cstl.nist.gov/strbase/mixture.htm**

# SWGDAM Website and Resources Available

**http://www.swgdam.org/resources.html**

- Home
- ByLaws
- Members
- Committees
- Meetings
- Publications



## Additional Resources

Beginning with the development or/and revision of its next draft guidance document(s), SWGDAM will make a "Draft for Comment" or other work product available for the purpose of receiving comments from the general public. This "Draft for Comment" solicitation will be open for a minimum of 60 days, usually through SWGDAM.org. SWGDAM will make all reasonable efforts to advise the forensic DNA community of the open comment period for a proposed guidance document or standard, guideline, best practice, study, or other recommendation and/or finding via as many avenues as possible to include posting notices through discipline-specific and related professional organizations. SWGDAM strongly encourages all interested parties to regularly monitor SWGDAM.org for the posting of such draft documents as well. All public comments received by SWGDAM will forwarded to the appropriate SWGDAM Committee for review and consideration as a part of its formal business practice for the development of the guidance documents or other work product.

*The following information resources have been produced and reviewed by members of the* ***Mixture Committee*** *of SWGDAM and are available at*
www.cstl.nist.gov/biotech/strbase/mixture/SWGDAM-mixture-info.htm

**Link to** http://www.cstl.nist.gov/biotech/strbase/mixture/SWGDAM-mixture-info.htm

# Mixture Training Materials

## Reviewed by SWGDAM Mixture Committee

## SWGDAM Mixture Committee Resource Page

---

*The following information resources have been produced and reviewed by members of the **Mixture Committee** of the Scientific Working Group on DNA Analysis Methods (SWGDAM) -- see* http://www.swgdam.org/resources.html *for additional information.*

---

### Mixture Training Examples

- Download "Mixture 6" PowerPoint show (56 Mb)

  - with voice-over by Bruce Heidebrecht (Maryland State Police); may work best if file is first saved to your computer

- Download "Mixture IQAS2904" PowerPoint show (35 Mb)

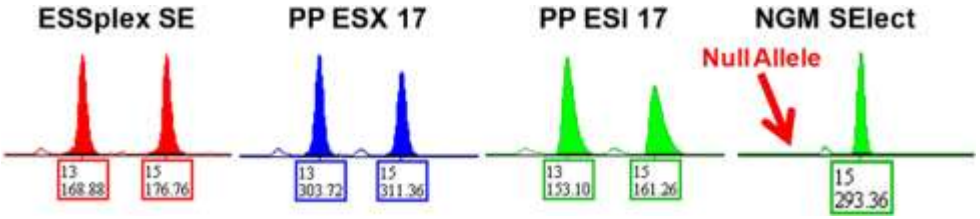
  - with voice-over by Bruce Heidebrecht (Maryland State Police); may work best if file is first saved to your computer

**http://www.cstl.nist.gov/biotech/strbase/mixture/SWGDAM-mixture-info.htm**

# STR Kit Concordance Studies

Becky Hill

## D18S51 Comparisons




D18S51 null allele with the NGM SElect kit as compared to the ESSplex SE kit, PowerPlex ESX 17 and ESI 17 systems

*Kits are kindly provided by **Applied Biosystems, Promega, and Qiagen** for concordance testing performed at NIST*

Forensic Science International: Genetics Supplement Series 3 (2011) e188–e189

Contents lists available at ScienceDirect

Forensic Science International: Genetics Supplement Series

journal homepage: www.elsevier.com/locate/FSIGSS

Concordance testing comparing STR multiplex kits with a standard data set

Carolyn R. Hill *, Margaret C. Kline, David L. Duewer, John M. Butler

U.S. National Institute of Standards and Technology, NIST 100 Bureau Drive, Gaithersburg, MD 20899-8314, USA

- Examined NIST samples across >20 STR kits and in-house assays covering 29 autosomal STR loci
- **99.90% concordance observed to-date**
  - 1,225 total differences due to primer binding site mutations from 1,176,994 allele comparisons (as of Oct 2012)
- Information provided back to kit developers to redesign primers or add extra ones – often prior to kit release

# Aiding Improvements with SE33 Primers

Forensic Science International: Genetics Supplement Series 3 (2011) e502–e503

Contents lists available at ScienceDirect

Forensic Science International: Genetics Supplement Series

journal homepage: www.elsevier.com/locate/FSIGSS

## SE33 variant alleles: Sequences and implications

John M. Butler[a,*], Carolyn R. (Becky) Hill[a], Margaret C. Kline[a], Ingo Bastisch[b], Volker Weirich[c], Robert S. McLaren[d], Douglas R. Storts[d]

[a] U.S. National Institute of Standards and Technology, Gaithersburg, MD, USA
[b] Bundeskriminalamt (BKA), Wiesbaden, Germany
[c] LKA, Mecklenburg-Vorpommern, Germany
[d] Promega Corporation, Madison, WI, USA

Home » Resources » Articles & Publications » *Profiles in DNA* »

http://www.promega.com/resources/articles/profiles-in-dna/2012/improved-primer-pair-for-the-se33-locus-in-the-powerplex-esi-17-pro-system/

## Improved Primer Pair for the SE33 Locus in the PowerPlex® ESI 17 Pro System

Robert S. McLaren[1], Jaynish Patel[1], Douglas R. Storts[1], Carolyn R. Hill[2*], Margaret C. Kline[2] and John M. Butler[2]

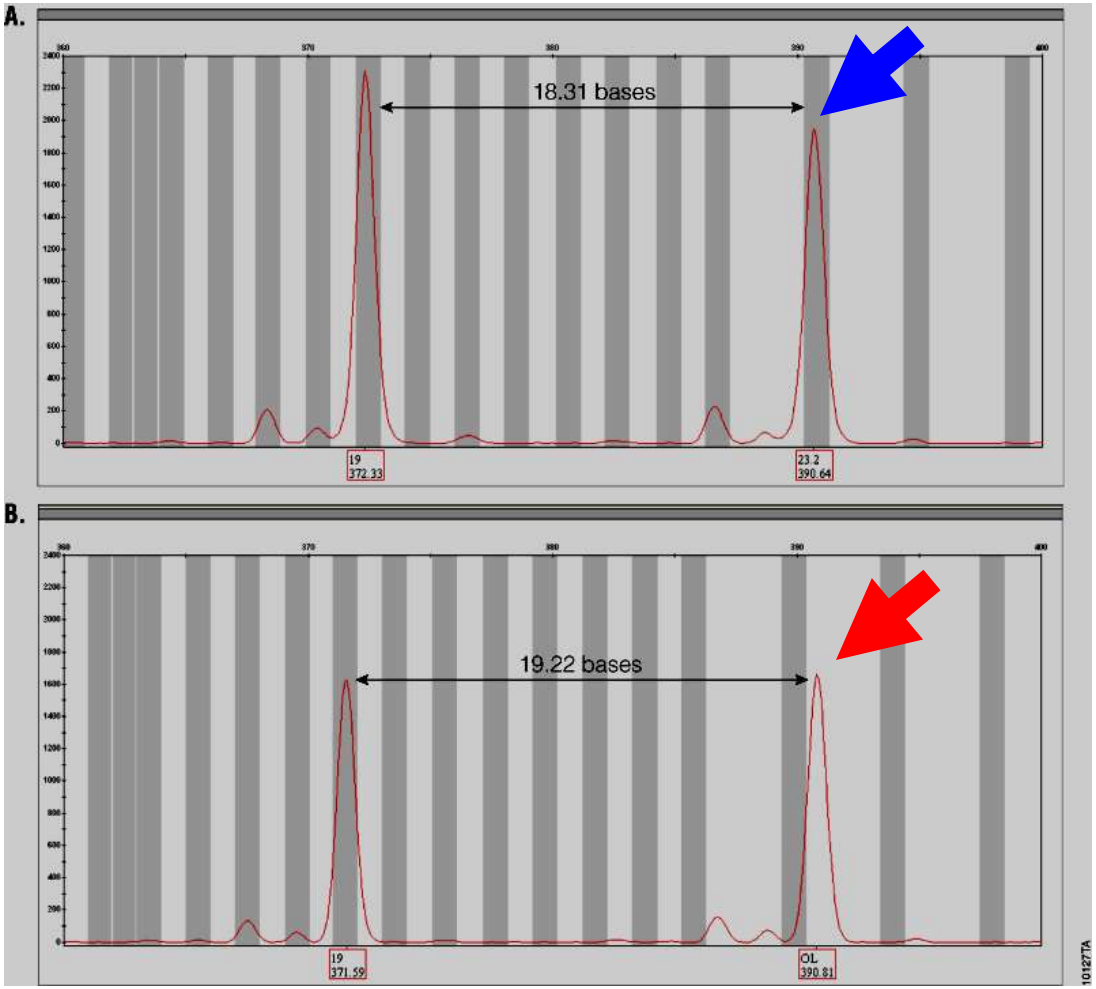
[1]Promega Corporation

[2]Human Identity Project Team, National Institutes of Standards and Technology

Publication Date: 2012

***A developmental validation article has also been prepared and submitted***

# PowerPlex ESI 17 Pro vs ESI 17 **SE33 Results**




PowerPlex **ESI 17 Pro**

**SE33 allele 23.2**

Reverse primer is **inside** of hairpin region

PowerPlex **ESI 17**

**SE33 allele "23.3"**

Reverse primer is **outside** of hairpin region

The SE33 locus range is shown for both PowerPlex® ESI 17 Pro (Panel A) and ESI 17 (Panel B) amplifications of DNA sample GT37190. Peak labels show allele calls (top) and sizes in bases (bottom). The off-ladder peak seen with PowerPlex® ESI 17 is correctly called as 23.2 with the PowerPlex® ESI 17 Pro System

http://www.promega.com/resources/articles/profiles-in-dna/2012/improved-primer-pair-for-the-se33-locus-in-the-powerplex-esi-17-pro-system/

# **NIST 1036** U.S. Population Samples

- 1032 males + 4 females
  - 361 Caucasians (2 female)
  - 342 African Americans (1 female)
  - 236 Hispanics
  - 97 Asians (1 female)

***Unrelated* samples**
All known or potential related individuals (based on autosomal & lineage marker testing) have been removed from the 1036 data set (e.g., only sons were used from father-son samples)

- Anonymous donors with self-identified ancestry
  - Interstate Blood Bank (Memphis, TN) – obtained in 2002
  - Millennium Biotech, Inc. (Ft. Lauderdale, FL) – obtained in 2001
  - DNA Diagnostics Center (Fairfield, OH) – obtained in 2007

- **Complete profiles with 29 autosomal STRs + PowerPlex Y23**
  - **Examined with multiple kits** and in-house primer sets enabling concordance

- Additional DNA results available on subsets of these samples
  - mtDNA control region/whole genome (AFDIL)
  - >100 SNPs (AIMs), 68 InDel markers, X-STRs (AFDIL)
  - NIST assays: miniSTRs, 26plex, >100 Y-STRs, 50 Y-SNPs

**Data available on STRBase**: **http://www.cstl.nist.gov/biotech/strbase/NISTpop.htm**

# Benefits of NIST 1036 Data Set

- **Elimination of potential null alleles due to primer binding site mutations** through extensive concordance testing performed with different PCR primer sets from all available commercial STR kits

- **Ancestry testing performed** on DNA samples with autosomal SNPs, Y-SNPs, and mtDNA sequencing to verify self-declared ancestry categorization

- **Related individuals removed** based on Y-STR and mtDNA results

# 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 29 commonly used autosomal STR loci

**Presentations/Publications**:

- 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

# Determination of Additional CODIS Core Loci

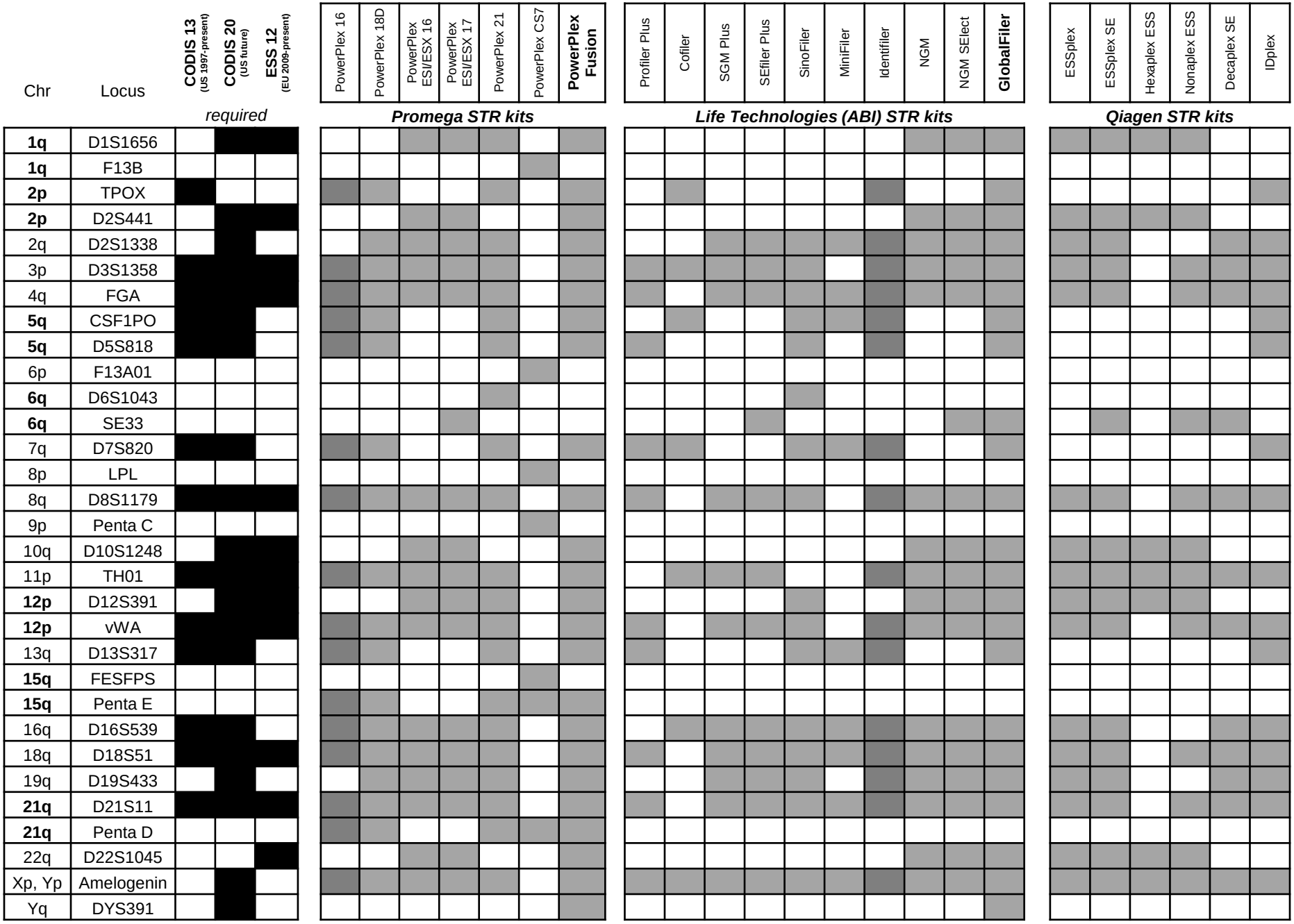
D.R. Hares (2012) Expanding the CODIS Core Loci in the United States. *Forensic Sci. Int. Genet.* 6: e52-e54
*Addendum to expanding the CODIS core loci in the United States*, Forensic Sci. Int. Genet. (2012) doi:10.1016/j.fsigen.2012.01.003

| What | Why | Who/How | When |
|---|---|---|---|
| Form a Working Group (WG) to discuss initial selection | Establishes target goals | CODIS Core Loci Working Group with FBI Chair and 5 members; Web meetings | May 2010 - present |
| Announce proposed additional CODIS core loci | Sets desired target goals and informs manufacturers | WG Chair; Publish proposed listing of CODIS core loci | April 2011 online (published Jan 2012) |
| Ongoing Progress Reports | Provides updates for DNA community | WG Chair; Present updates on status of CODIS Core Loci project at meetings | 2010-2012 |
| Implementation Considerations & Strategy | Identify issues for implementation and timeline | WG | June 2011 - present |
| Manufacturers develop prototype kits | Creates tools to meet target goals | Manufacturers; Provide status reports to WG for timeline | 2011-2012 |
| Test and validate prototype kits | Examines if target goals can be met | Validation Laboratories; Follow QAS compliant validation plan | Beginning in 2012 |
| Review and evaluate data from validation | Evaluates if desired performance is obtained | NIST, SWGDAM and FBI; Provide feedback, if any, to Manufacturers | In conjunction with and at the conclusion of validation |
| Selection of new CODIS core loci | Allows protocols to be established | FBI; seek input from DNA community and stakeholders; Notify Congress | After evaluation of validation data and kit production factors |
| Implementation of new CODIS core loci at the National DNA Index System | Enables target goals to be met | All NDIS-participating labs | ~ 24 months after selection of new CODIS core loci |

**http://www.fbi.gov/about-us/lab/codis/planned-process-and-timeline-for-implementation-of-additional-codis-core-loci**

# 29 autosomal STRs

# STR Loci Present in Current Commercial Kits




Butler, J.M. & Hill, C.R. (2013) *Topics on Forensic DNA Analysis: Current Practices & Emerging Technologies* (CRC Press), Figure 9.1

# Probability of Identity Values
## for Various STR Kits or Locus Combinations
based on NIST 1036 U.S. Population Samples

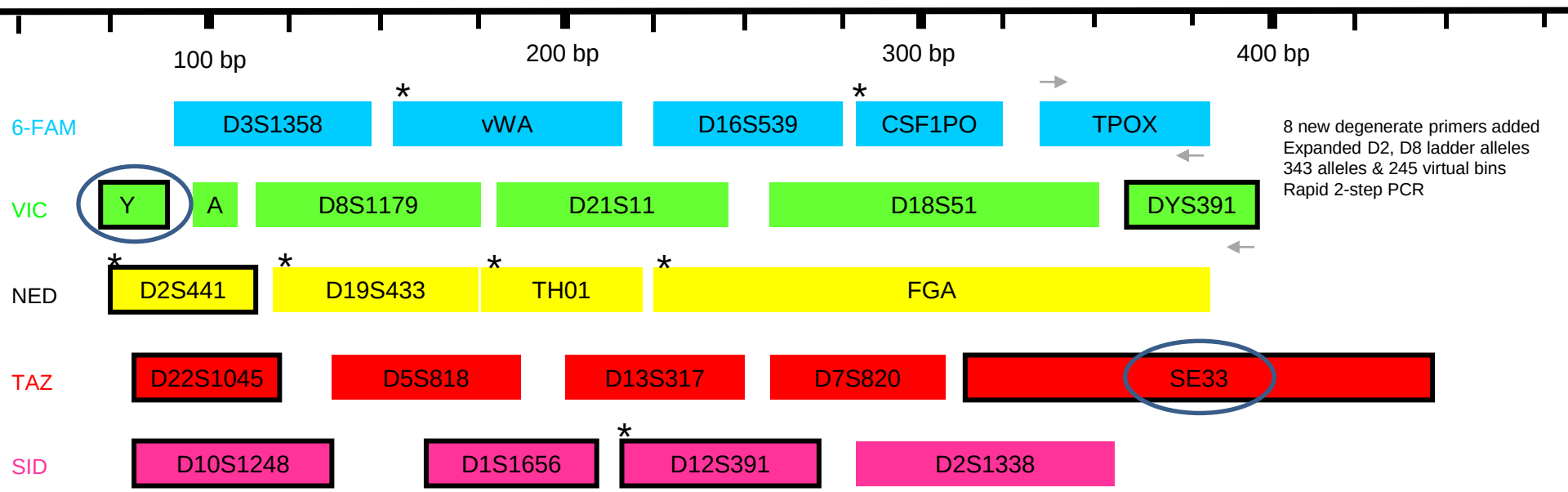
| STR Kit or Core Set of Loci | Total N=1036 | Caucasians (n=361) | African Am. (n=342) | Hispanics (n=236) | Asians (n=97) |
|---|---|---|---|---|---|
| **CODIS 13** | 5.02E-16 | 2.97E-15 | 1.14E-15 | 1.36E-15 | 1.71E-14 |
| Identifiler | 6.18E-19 | 6.87E-18 | 1.04E-18 | 2.73E-18 | 5.31E-17 |
| PowerPlex 16 | 2.82E-19 | 4.24E-18 | 6.09E-19 | 1.26E-18 | 2.55E-17 |
| PowerPlex 18D | 3.47E-22 | 9.82E-21 | 5.60E-22 | 2.54E-21 | 7.92E-20 |
| | | | | | |
| **ESS 12** | 3.04E-16 | 9.66E-16 | 9.25E-16 | 2.60E-15 | 3.42E-14 |
| ESI 16 / ESX 16 / NGM | 2.80E-20 | 2.20E-19 | 6.23E-20 | 4.03E-19 | 9.83E-18 |
| ESI 17 / ESX 17 / NGM SElect | 1.85E-22 | 1.74E-21 | 6.71E-22 | 3.97E-21 | 1.87E-19 |
| | | | | | |
| **CODIS 20** | 9.35E-24 | 7.32E-23 | 6.12E-23 | 8.43E-23 | 4.22E-21 |
| GlobalFiler | 7.73E-28 | 1.30E-26 | 3.20E-27 | 2.27E-26 | 1.81E-24 |
| PowerPlex Fusion | 6.58E-29 | 2.35E-27 | 1.59E-28 | 2.12E-27 | 1.42E-25 |
| | | | | | |
| **All 29 autosomal STRs** | 2.24E-37 | 7.36E-35 | 3.16E-37 | 2.93E-35 | 4.02E-32 |
| **29 autoSTRs + DYS391** | 1.07E-37 | 3.26E-35 | 1.77E-37 | 1.29E-35 | 2.81E-32 |

Hill et al. ISHI 2012 poster #84 (see STRBase); Butler et al. (2012) *Profiles in DNA (in press)*

# **STR Kit Layouts** by Dye Label and PCR Product Size

## Life Technologies/Applied Biosystems **GlobalFiler** (6-dye – LIZ600 size standard)

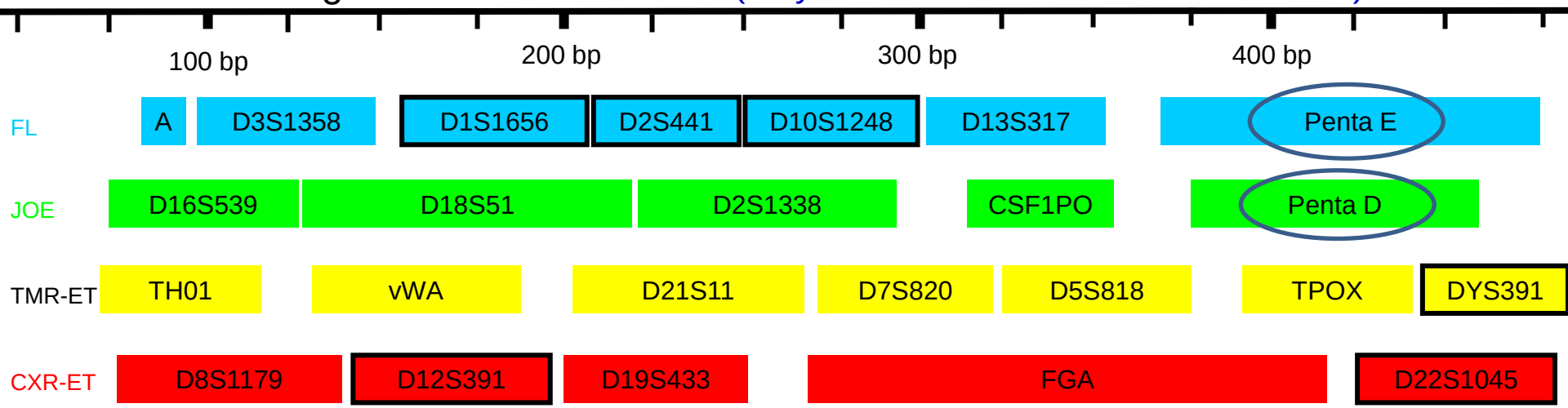
24plex




8 new degenerate primers added
Expanded D2, D8 ladder alleles
343 alleles & 245 virtual bins
Rapid 2-step PCR

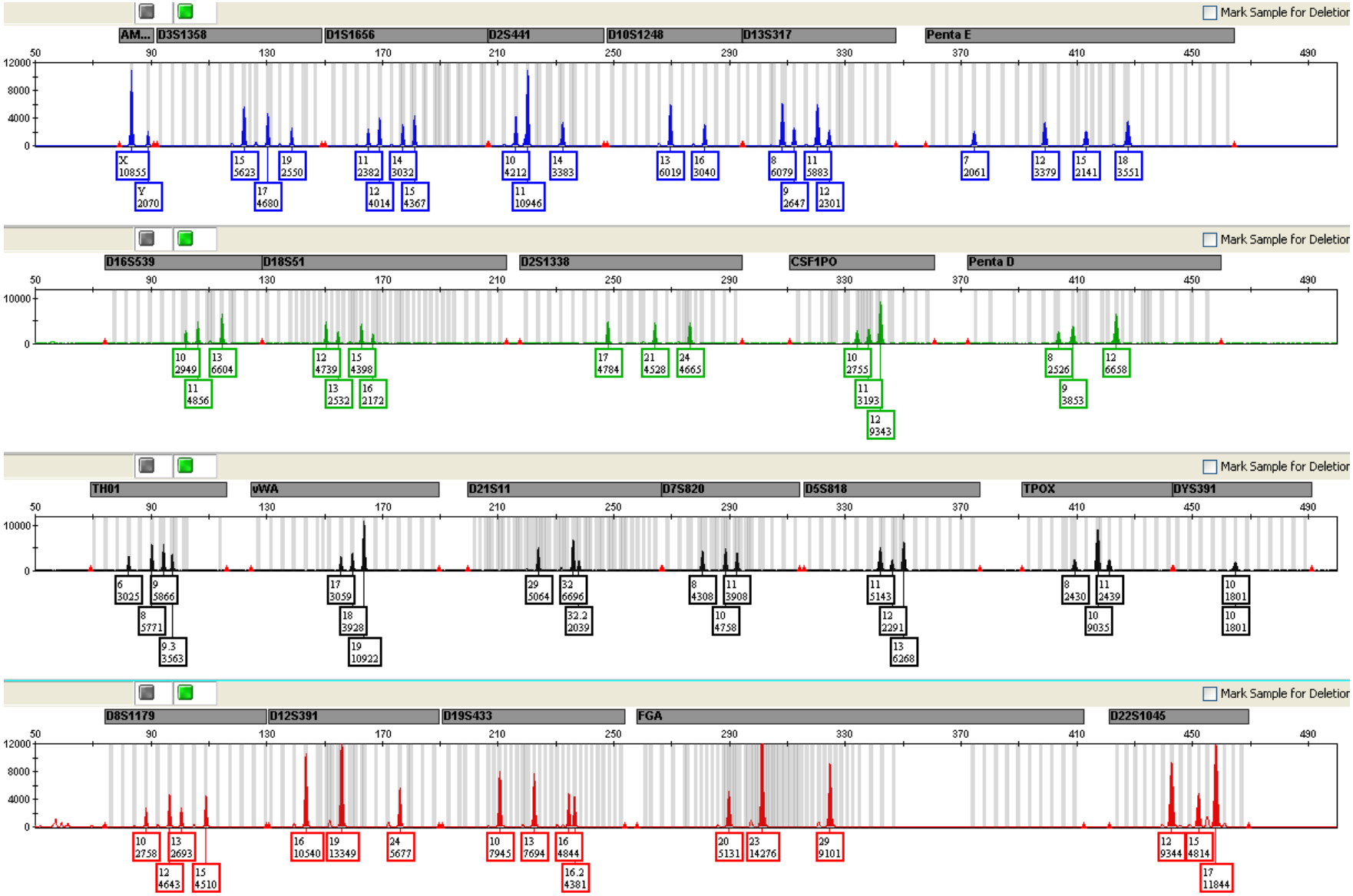
## Promega PowerPlex **FUSION** (5-dye – CC5 internal lane standard 500)

24plex



# DNA Mixture with **PowerPlex Fusion** (Promega)

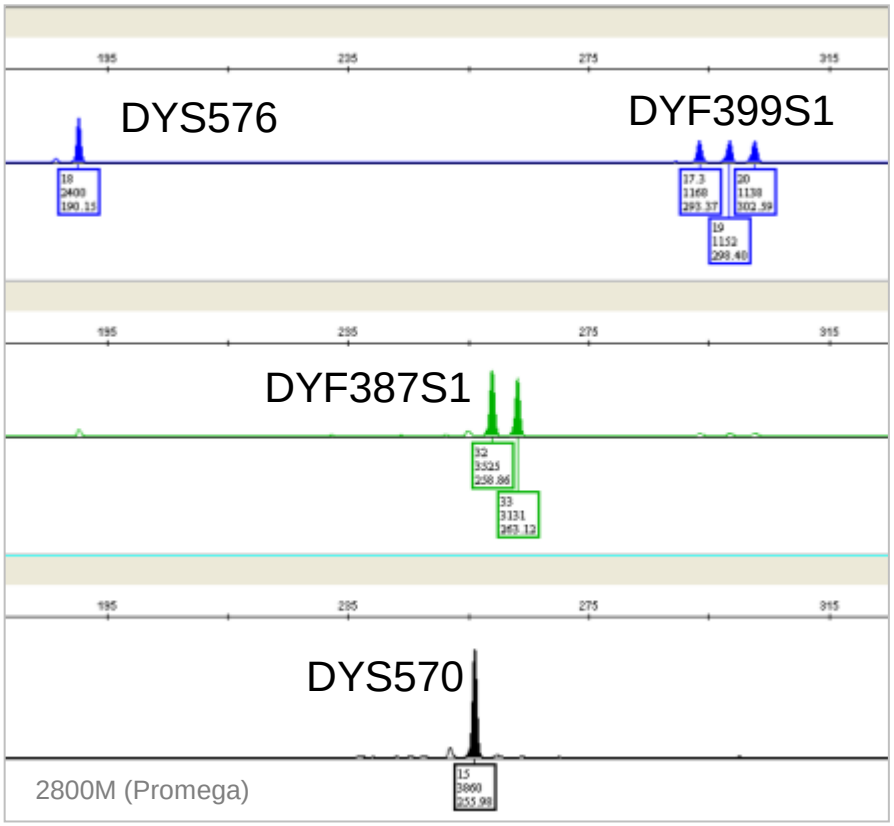
**24plex assay**



# Rapidly Mutating Y-STR Loci

Mike Coble

Becky Hill



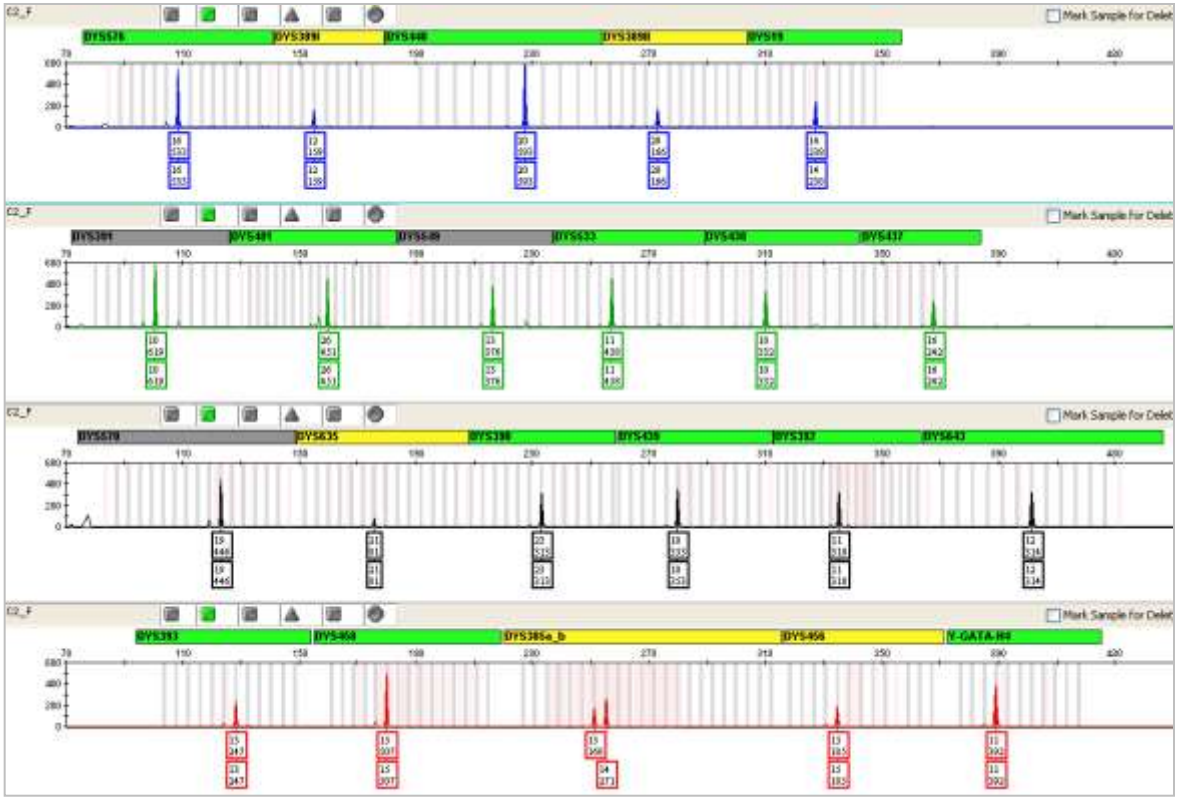

- Part of RM Y-STR Study Group organized by Manfred Kayser (Erasmus University, The Netherlands)
- Supplied data from 1,296 U.S. samples (634 population + 331 father/son pairs)
- Publication with RM Y-STR Study Group is forthcoming

# PowerPlex Y23 Kit


Mike Coble


Becky Hill

125pg male + 400ng female **(3200x female)**



**Kit found to be *sensitive* and *specific* to male DNA**

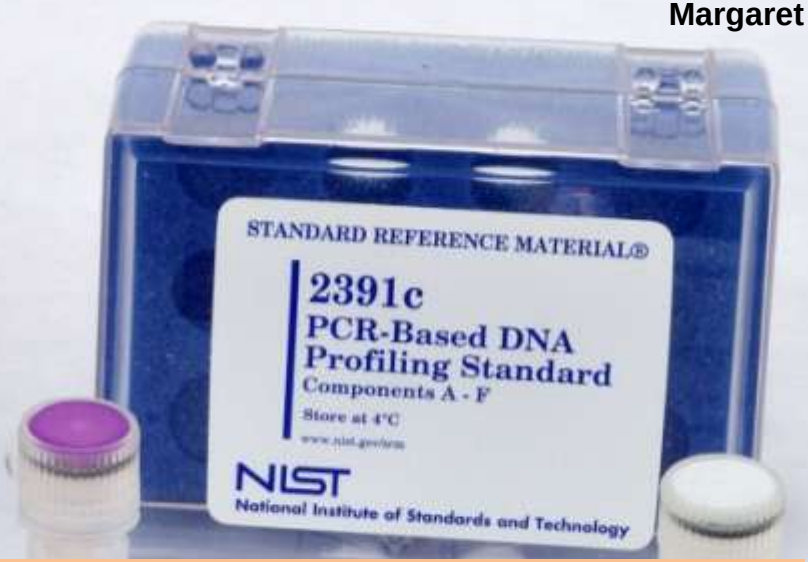
- Typed 1032 males from 4 U.S. population groups
- Data supplied to YHRD and USYSTR databases
- Publications are forthcoming
- Full dataset to be released on STRBase

# NIST Reference Materials for Forensic DNA Measurement Assurance

Margaret Kline


**SRM 2372 is currently not available** because the dsDNA has unraveled, which impacts absorbance certification values. We are re-certifying the samples with aid of digital PCR measurements. **We hope to have it available again by early 2013.**

**DNA quantity measurement calibration**




**SRM 2391c currently does not cover the six additional Y-STR markers in PowerPlex Y23.** We plan to certify values for these markers by mid-2013.

**Autosomal and Y-chromosome short tandem repeat (STR) measurement calibration**

# Rapid DNA Efforts


**Pete Vallone** **Erica Butts**

Accelerated Nuclear DNA Equipment (ANDE) developed by **NetBIo**

http://ishinews.com/wp-content/uploads/2012/10/Rapid-DNA-Miles-1.58MB.pdf

RapidHIT 200 developed by **IntegenX**


http://integenx.com/wp-content/uploads/2010/06/RapidHIT-200.png

- Evaluating ANDE (NetBio) and IntegenX rapid DNA instruments
  - both instruments are capable of swab in → STR profile out in less than 90 minutes without user intervention
- Exploring rapid DNA techniques including direct PCR and rapid PCR
  - STR profiles generated in <2 hours with standard lab equipment and rapid protocols
  - See ISHI 2012 poster available on STRBase “Rapid DNA Testing Approaches for Reference Samples”

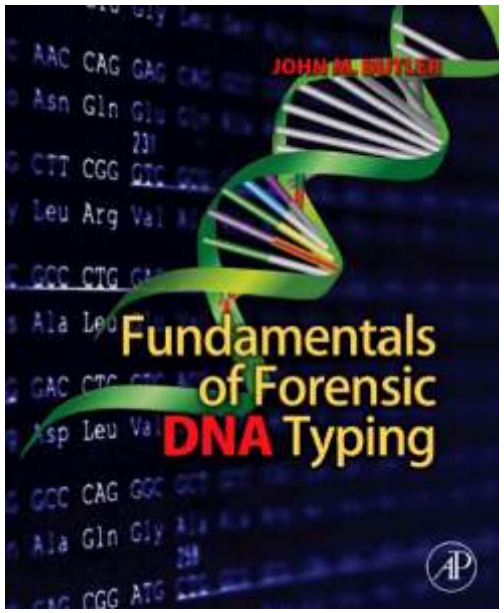
# *Forensic DNA Typing* Textbook

## 3rd Edition is Three Volumes

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

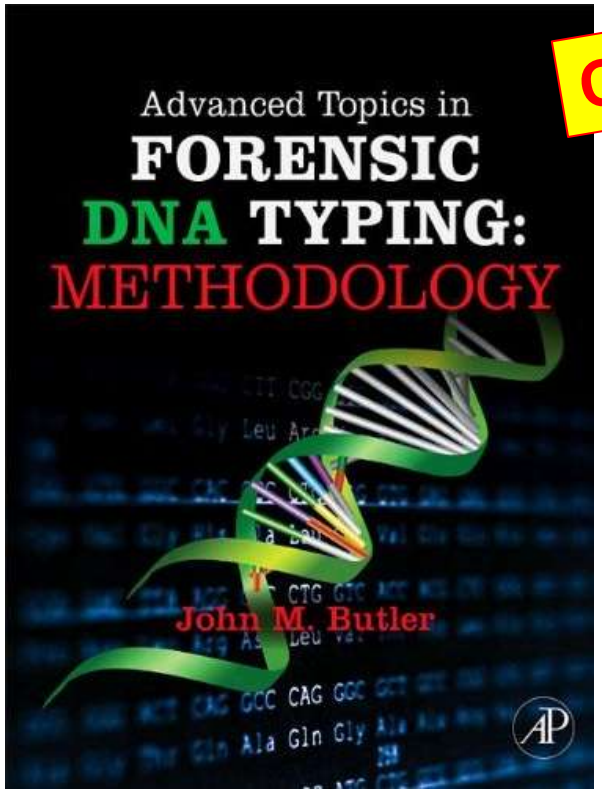
John Butler

*For beginning students, general public, & lawyers*




**Fall 2009**

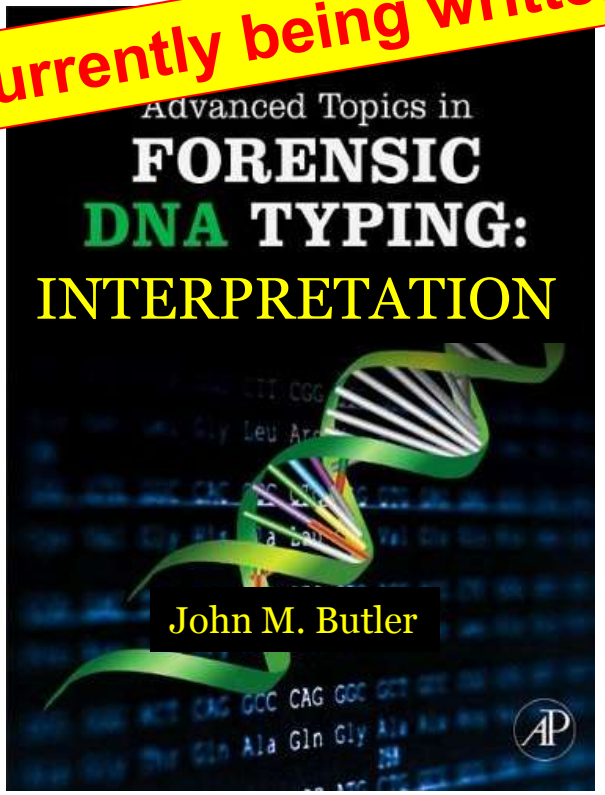
**~500 pages**




**Fall 2011**

**~700 pages**

**Currently being written**

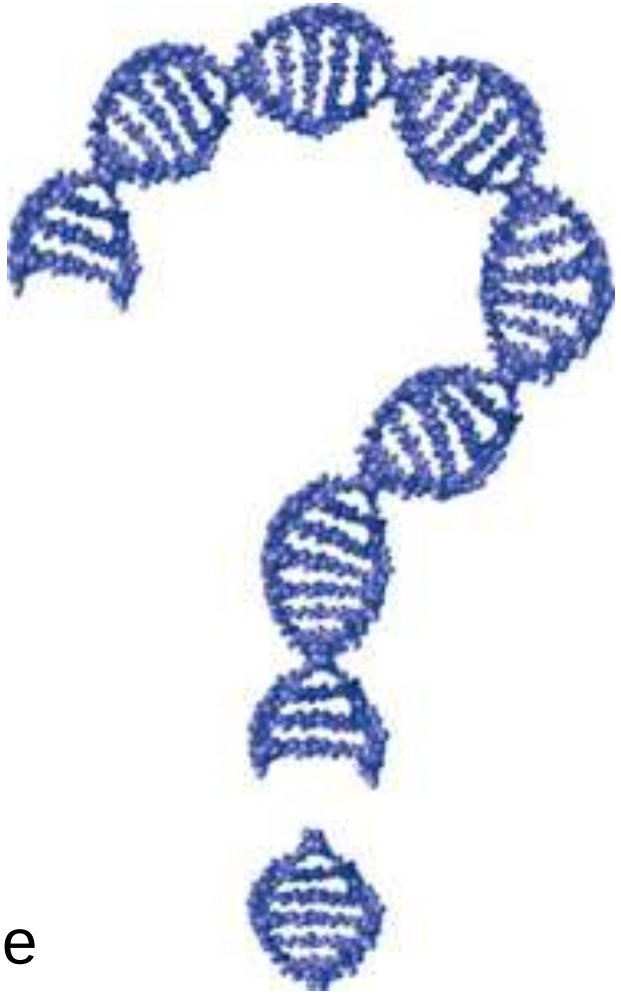



**Fall 2013**

**~500 pages**

# 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**