CAC Fall Meeting (Sacramento, CA) – October 25, 2011

CALIFORNIA ASSOCIATION OF CRIMINALISTS

NIST Update:

On-going research projects from a highly productive group

John M. Butler

NIST Applied Genetics Group

National Institute of Standards and Technology

Gaithersburg, Maryland

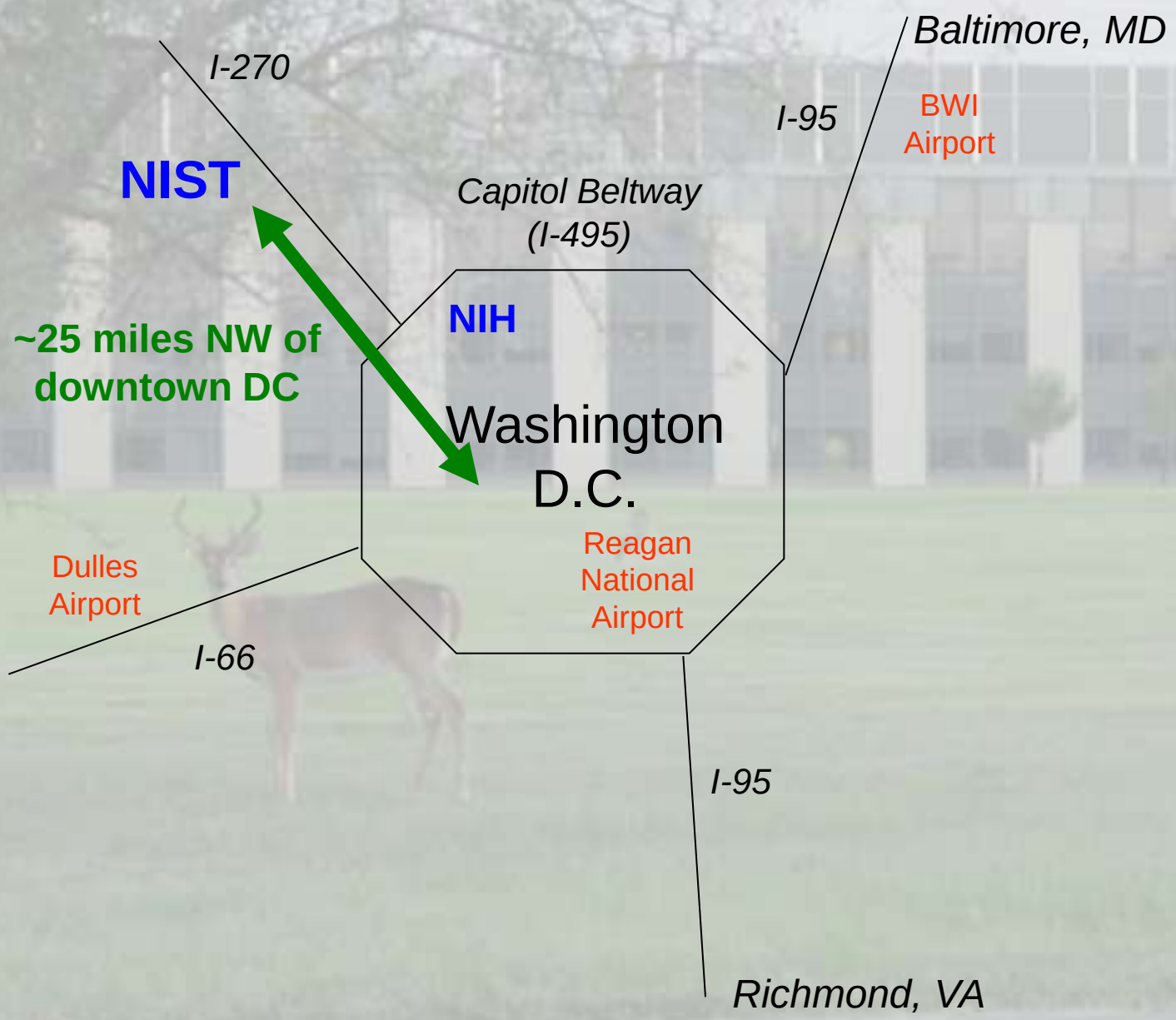
Presentation Topics

- Introduction to NIST and to Our Group
- STRBase website
- Textbooks
  - *Advanced Topics in Forensic DNA Typing: Methodology*
- **Group Research Overview**
  - Standard Reference Materials (SRMs)
- ABI 3500 open letter status update

National Institute of Standards & Technology (NIST)

- **Non-regulatory agency** established in 1901 in the US Department of Commerce.
- Mission to promote US innovation and industrial competitiveness by advancing measurement science, standards & technology.
- NIST develops a wide variety of physical standards, test methods, and standard reference data.

Location of NIST



Early Driver for U.S. Standards

1904

- Out-of-town fire companies arriving at a Baltimore fire cannot couple their hoses to the hydrants. 1526 buildings razed.

1905

- National Fire Protection Association adopted NBS-developed national hose coupling standard.

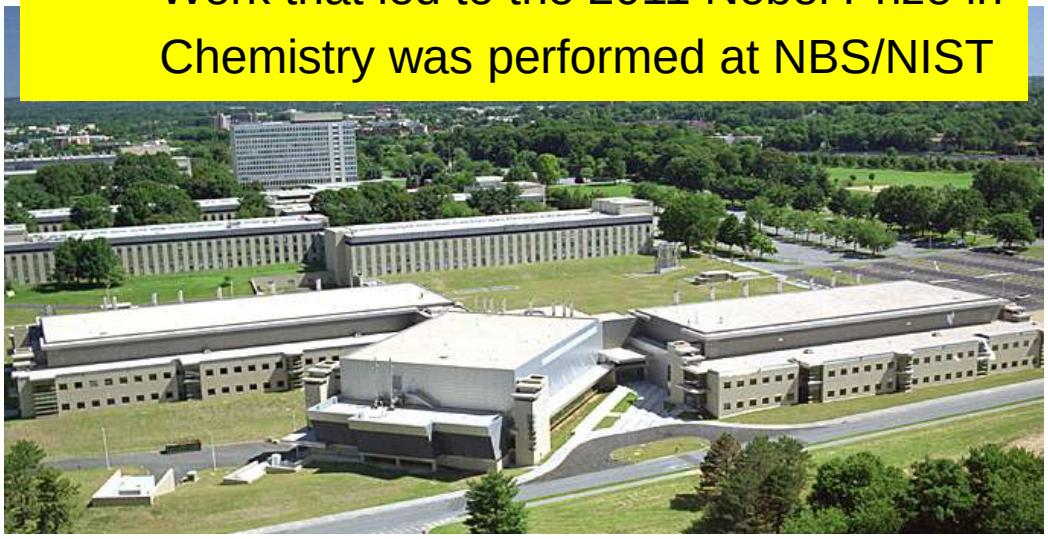
NBS = National Bureau of Standards (name changed to NIST in 1988)

NIST Today

Major Assets

- ~ 2,900 employees
- ~ 2600 associates and facilities users
- ~ 400 NIST staff on about 1,000 national and international standards committees
- **3 Nobel Prizes in Physics in past 15 years**

Work that led to the 2011 Nobel Prize in Chemistry was performed at NBS/NIST



Major Programs

- **NIST Laboratories**
- Baldridge National Quality Program
- Hollings Manufacturing Extension Partnership
- Technology Innovation Program

Joint NIST/University Institutes:

- JILA
- Joint Quantum Institute
- Institute for Bioscience & Biotechnology Research
- Hollings Marine Laboratory

NIST Organizational Structure and Where Our Group Fits…

Laboratories

- Engineering
- Physical Measurement
- Information Technology
- Material Measurement
- Center for Nanoscale Scien
- NIST Center for Neutron Re

The Laboratory programs at NIST were reorganized in October 2010 into four labs and two centers

Material Measurement Laboratory (MML)

Divisions

- Analytical Chemistry
- Biochemical Science
- Ceramics
- Chemical and Biochemical Reference Data
- Materials Reliability
- Measurement Services
- Metallurgy
- Polymers
- Surface and Microanalysis Science
- Thermophysical Properties

Biochemical Science Division (BSD)

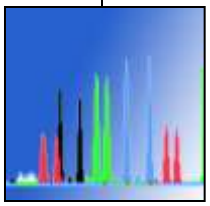
Groups

- Applied Genetics Group
- Bioassay Methods Group
- Cell Systems Science Group
- DNA Science Group
- Macromolecular Structure and Function Group
- Multiplexed Biomolecular Science Group
- Process Sensing Group

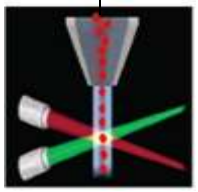
NIST Biochemical Science Division

Laurie Locascio
Division Chief

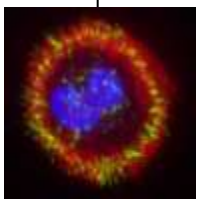
Process Sensing
Dean Ripple
Group Leader



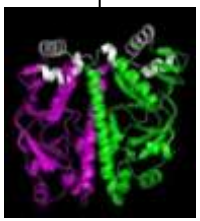
Applied Genetics
John Butler
Group Leader



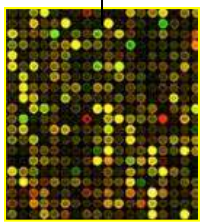
Bioassay Methods
Steve Choquette
Group Leader



Cell Systems Science
Anne Plant
Group Leader



Macromolecular Structure & Function
John Marino
Group Leader



Multiplexed Biomolecular Science
Marc Salit
Group Leader

Doing some Next Generation Sequencing using ABI SOLID

NIST
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)
- **FBI Science & Technology Branch** (DNA Biometrics)
- **NIST SRM Program** (SRM development and production)
- Base funding from Congress (clinical DNA)

NIST
Applied Genetics

NIST Applied Genetics Group

Group Leader

John Butler

Marcia Holden

Margaret Kline

Pete Vallone

Mike Coble

Ross Haynes

Becky Hill

Erica Butts

Kristen O'Connor

Kevin Kiesler

NIST Applied Genetics

Our FY2011 Group Productivity

(Oct 2010 to Sept 2011)

- **21 publications**
  - 20 articles + 1 book
- **77 presentations**
  - 65 talks (58 invited) + 12 posters (all available on STRBase)
- **10 training workshops**
  - Mixture interpretation (ISHI, AAFS, NFSTC, IN, HI, AZ, MI, Palm Beach, Houston)
  - Capillary electrophoresis (ISFG)
- **3 Standard Reference Materials (SRMs) completed**
  - 2391c (forensic STRs), 2393 (HD), 2366 (CMV)
- **10 committee assignments**
  - VA SAC, DOD DNA oversight, FBI new CODIS core loci, SWGDAM (mixture interpretation, rapid DNA, enhanced detection methods), NIST/NIJ evidence preservation TWG, JCTLM, NIJ DNA TWG, ATCC cell line authentication

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
- **Ag Biotech**
  - "universal" GMO detection/ quantitation (35S promoter)
- **DNA Biometrics**
  - **Rapid PCR methods**
  - Efforts to standardize testing of future portable DNA systems
  - Kinship analysis
- **Cell Line Authentication**
  - ATCC documentary standard

Aiding Cell Line Authentication

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

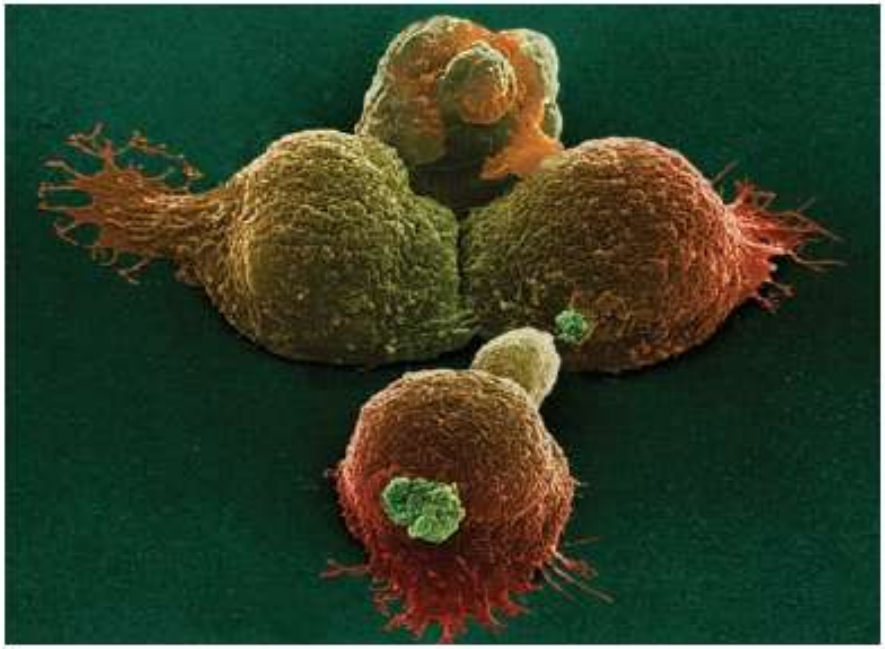
Biologists tackle cells' identity crisis

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-



Breast cancer cells: not always what they're supposed to be.

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

Support to Cell Line Authentication Efforts

Margaret Kline

John Butler

Timeline | **Key milestones in the effort to address cell line misidentification**

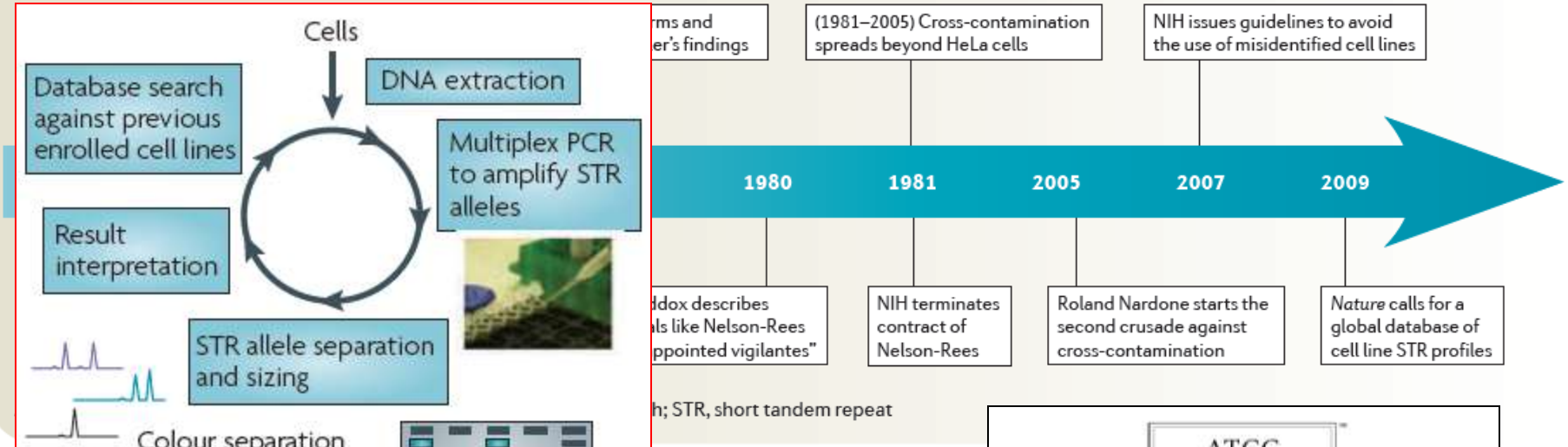
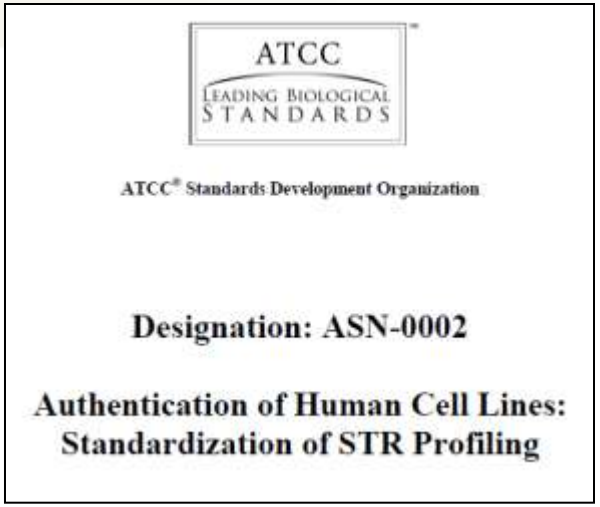


Figure 3 | **Short tandem repeat profiling methodology.** Short tandem repeat (STR) loci consist of repetitive DNA sequences with varying numbers of repeats. Each STR locus can be polymerase chain reaction (PCR) amplified and the amplified products labelled with fluorophores of different colours, making the products easy to distinguish by size and colour. Images courtesy of J. Butler, National Institute of Standards and Technology.



Masters. J.R.W., et al. (2010) Cell line misidentification: the beginning of the end. *Nature Rev. Cancer* 10: 441-448.

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

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

Office Manager
Patti Rohmiller

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

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* ◆

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

NIST Human Identity Team Projects

Funded by the National Institute of Justice

http://www.cstl.nist.gov/biotech/strbase/NIJprojects.htm

Projects

33 different projects are described

[Human DNA Quantitation] [Mitochondrial DNA] [Y Chromosome] [Compromised DNA Evidence] [Miniaturization and Automation] [General Tools and Information] [Non-Human DNA] [Alternative Forensic DNA Markers]

Alphabetical Listing of Projects

ABI 3100 performance with various STR typing systems (April 2001-June 2003)

ABI 3130xl upgrade evaluation (Sept 2005-May 2006)

AutoDimer: software to enable rapid multiplex PCR design (2000-2005) [see also software.htm]

Autosomal SNP loci (July 2002-present)

Autosomal STR loci: beyond the CODIS markers (Jan 2004-present) [see also newSTRs.htm]

Biomatrica dry storage device DNA stability studies (June 2007-present)

ABI 3100 Performance with Various STR Typing Systems

Participants: John M. Butler, Margaret C. Kline, Richard Schoske, and Peter M. Vallone

Purpose: T

Progress: W
typing kits
June 2002.
combination
combination
performance
were also ex
produced a
into routine

Publication

Jo
th

ABI 3130*xl* Upgrade Evaluation

Participants: Carolyn R. "Becky" Hill, Amy E. Decker, Peter M. Vallone, Margaret C. Kline, and John M. Butler

Collection

Progress:
ABI 3100
determine
pump (pic
concordan
We experi
injections

Publicati

B
(W

[Return

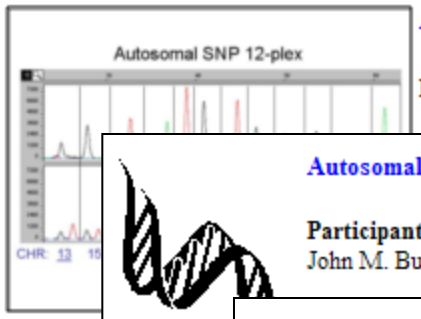
AutoDimer: Software Developed to Enable Rapid Multiplex PCR Design

Participants: Peter M. Vallone and John M. Butler

community.

Progress: Th
based diagnos
extension, and
primer selecti
sequences are
short DNA o
PCR primers
sets of pre-se
primer-primer
visually inspe
of interaction
primer-primer

The AutoDim
Approximatel
developed int



Autosomal SNP Assays

Participants: Peter M. Vallone, Amy E. Decker, and John M. Butler

Progress:
were type
sample, th
of SNP m
Hardy-We
distinguish
on testing
genotypes

The 12 se
exhibit mo
likely that
comparis
(STR) loc
unrelated

Autosomal STR Loci: Beyond the CODIS Markers

Participants: Carolyn R. "Becky" Hill, Michael D. Coble (now at AFDIL), Peter M. Vallone, Margaret C. Kline, and John M. Butler

Purpose: A
evidence imp
of time, degr
repeat (STR
System (CO
in a set of m
repeat flanki
allele ranges
amplicons le

Progress: N
selected base
comparing th
observed wi
information
characterized
their core ge
developed a

Biomatrica Dry Storage Device DNA Stability Studies

Participants: Margaret C. Kline

Project Timeframe: June 2007 to present

Purpose: The ability to ship and store DNA samples at room temperature could benefit laboratories This particular study has been designed to examine the effect of "shipping" well characterized genomic DNA extracts on Biomatrica SampleGard™ Dry Storage Devices.

Progress: The devices (three replicate plates) have been prepared at NIST with 20 µL of genomic DNA at concentrations of 1 ng/µL, 0.25 ng/µL and 0.05 ng/µL. NIST plans to analyze the selected genomic DNA extract before and after application on the storage device using appropriate DNA quantitation assay(s) such as Quantifiler and short tandem repeat (STR) genotyping methods such as Identifiler. Two plates are being shipped at ambient temperature back and forth multiple times between Maryland (NIST) and California (Biomatrica) in the middle of the summer via U.S. Postal Service in Barrier pouches supplied by Biomatrica. Two portable temperature/humidity recorders are being shipped along side the plates to enable monitoring environmental conditions. Sampling is being conducted at NIST with each arrival of the shipped plates and compared to a control plate stored at NIST for the duration of the study. The range of temperature and humidity changes experienced by the shipped samples will be tracked. The shipping and analysis process will be repeated until degradation of the samples is detected or the samples have been exhausted. Starting this study during the summer months is desirable to stress the system at extreme heat and humidity conditions commonly occurring during the shipping process.

Publications or Presentations Resulting From This Project:

[Return to NIJ Projects page] [Return to STRBase]

STRBase
…/NIJprojects.htm

Benefits of Website like STRBase

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

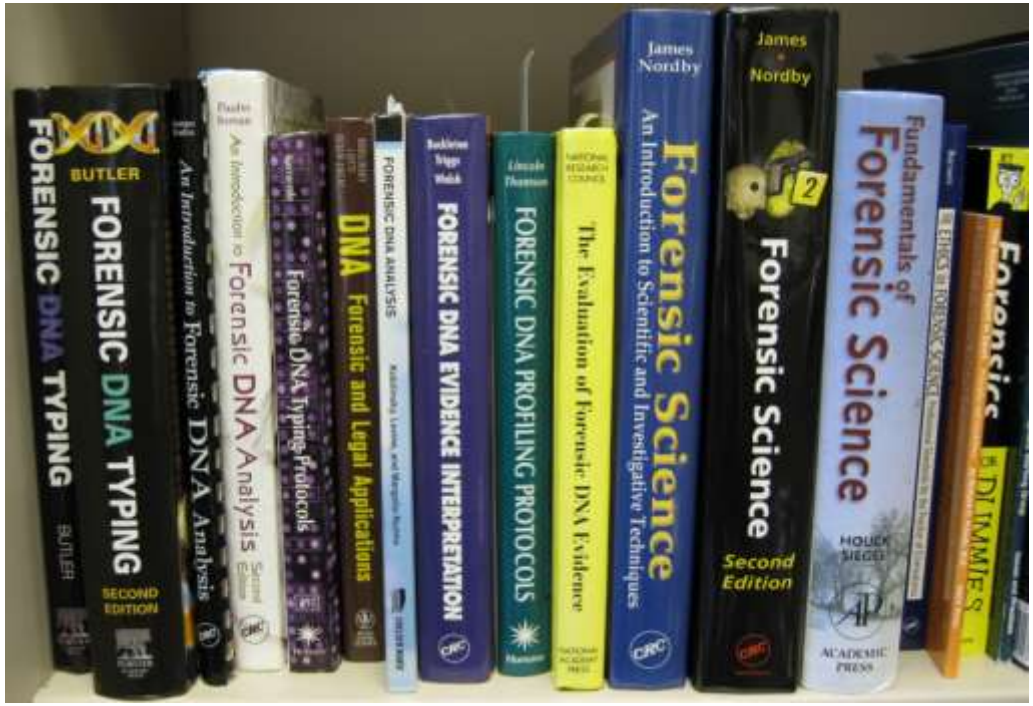
- Develops expertise when collecting information
- Requires NIST to stay up-to-date with field
- Provides transparency to our team's work
- Training tool and resource for the world
- Respected resource for >14 years
- ~10,000 pages of information available now
- >400,000 hits cumulative
- **Method for sharing information (PowerPoint files, population data, etc.)**

Forensic Science Publications

Forensic DNA Library

in my office and our group library

- We have purchased >300 books on topics related to forensic DNA analysis as of Oct 2011



Initially funded from 2002-2007 by PECASE award money

Fruits of a Good Literature Collection

Review Articles

J Forensic Sci, March 2006, Vol. 51, No. 2
doi:10.1111/j.1556-4029.2006.00046.x
Available online at: www.blackwell-synergy.com

John M. Butler,[1] *Ph.D.*

Genetics and Genomics of Core Short Tandem Repeat Loci Used in Human Identity Testing

analytical chemistry

Anal. Chem. 2011, 83, 4539–4556 REVIEW

pubs.acs.org/ac

575 references reviewed (121 on DNA)

Forensic Science

T. A. Brettell

Department of Chemical and Physical Sciences, Cedar Crest College, 100 College Drive, Allentown, Pennsylvania 18104-6196, United States

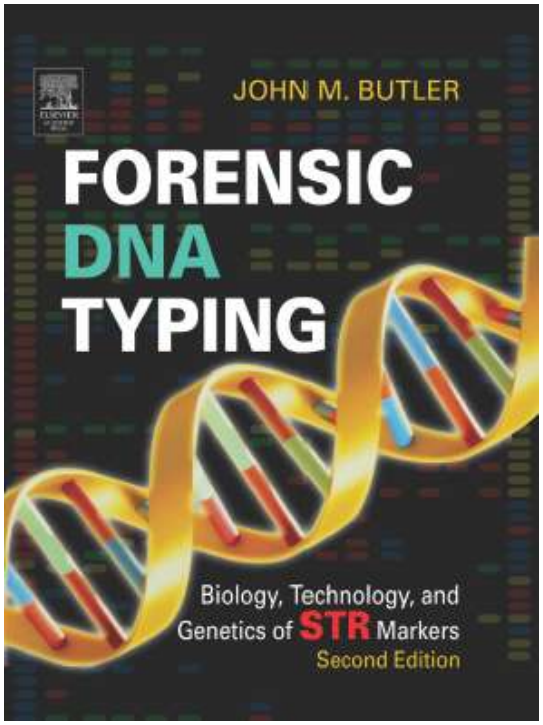
J. M. Butler

Biochemical Science Division, National Institute of Standards and Technology, Gaithersburg, Maryland 20899-8312, United States

J. R. Almirall

Department of Chemistry and Biochemistry and International Forensic Research Institute, Florida International University, University Park, Miami, Florida 33199, United States

Textbooks



2nd Edition 688 pp.
Feb 2005

Language Editions of Forensic DNA Typing

Chinese (2007)

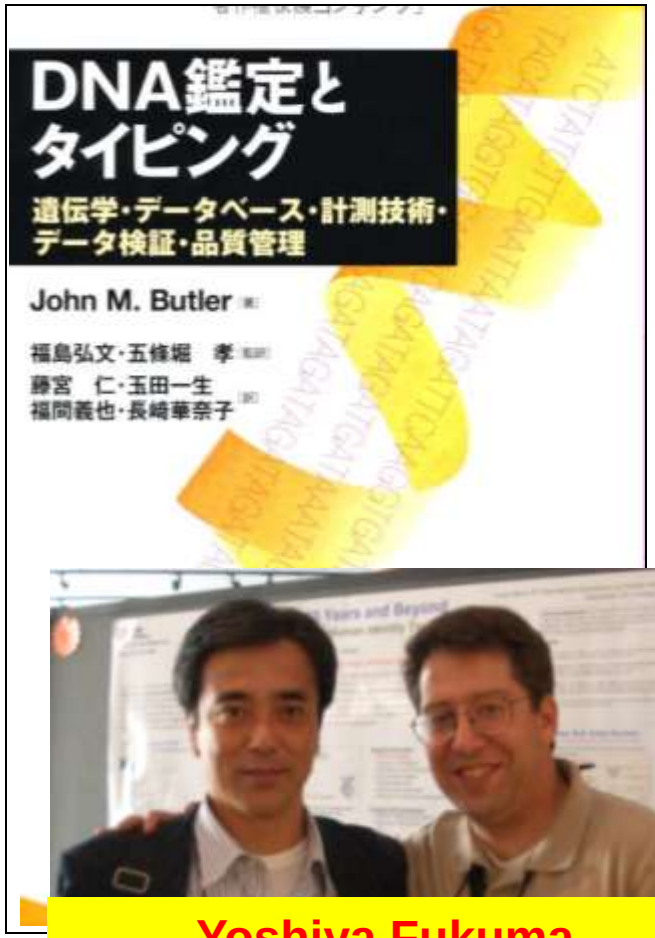
Translated by Y. Hou

http://www.sciencep.com/s_single.php?id=12683

Yiping Hou (Chinese translator)

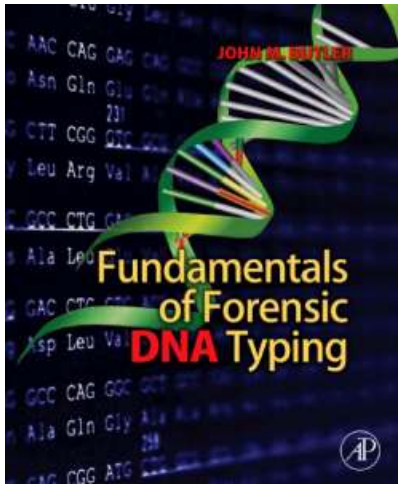
Japanese (2009)

Translated by Y. Fukuma



http://www.amazon.co.jp/gp/reader/4320056825

Yoshiya Fukuma
(Japanese translator)



Written as Part of My Job at NIST
(no royalties to be received)

Fundamentals of Forensic DNA Typing

Contribution of the National Institute of Standards and Technology, 2010.

Academic Press is an imprint of Elsevier
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA
525 B Street, Suite 1900, San Diego, California 92101-4495, USA
84 Theobald's Road, London WC1X 8RR, UK

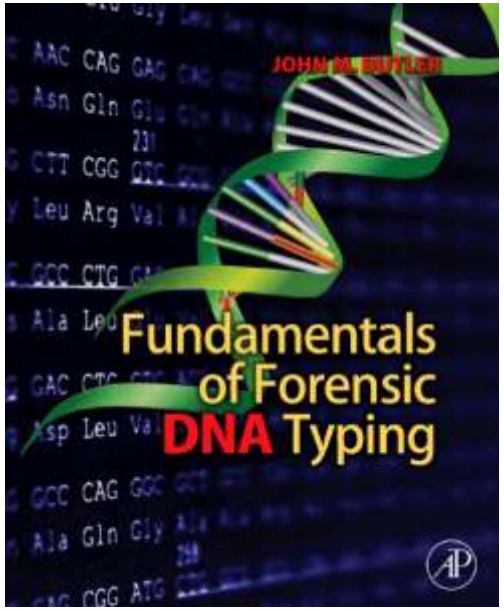
This work was funded in part by the National Institute of Justice (NIJ) through interagency agreement 2008-DN-R-121 with the NIST Office of Law Enforcement Standards. Points of view in this document are those of the author and do not necessarily represent the official position or policies of the U.S. Department of Justice. Certain commercial equipment, instruments, and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that any of the materials, instruments, or equipment identified are necessarily the best available for the purpose.

Forensic DNA Typing Textbook

3rd Edition is Three Volumes

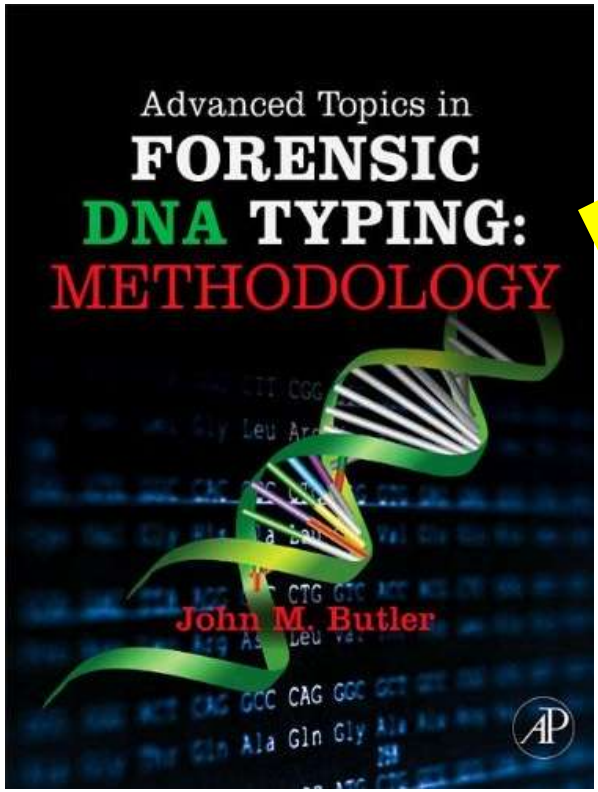
John Butler

For beginning students, general public, & lawyers



Sept 2009

~500 pages



August 2011

~700 pages

Currently being written

Advanced Topics in Forensic DNA Typing: **INTERPRETATION**

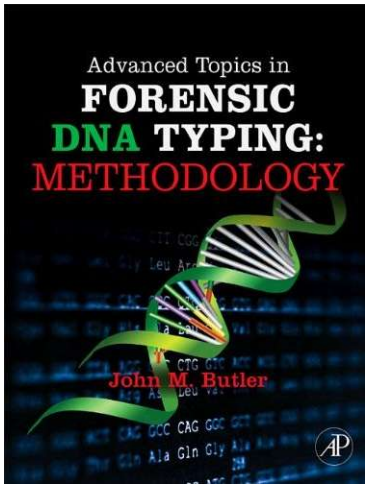
Fall 2012

~500 pages

New Material in *Advanced Topics: Methodology*

Released August 2011

>50% new material from previous editions



- Cites >1500 new references (>2800 ref. total)
- **New chapter** on legal aspects (Ch. 18)
  - expert witness prep, perspectives from lawyers
  - App. 4 (interviews): experts, prosecutors, & defense
- **New chapter** on X-chromosome markers (Ch. 15)
- **Extensive updates** on CE (Ch. 6), validation (Ch. 7), database issues (Ch. 8), disaster victim identification (Ch. 9), miniSTRs (Ch. 10), LTDNA (Ch. 11), SNPs (Ch. 12), Y-STRs (Ch. 13), mtDNA (Ch. 14), non-human DNA (Ch. 16), and new technology (Ch. 17)
- Coverage of all the new STR kits (Ch. 5)
- Listing of all known STR alleles for all 23 kit loci (App. 1)
- Most detail to-date on the Grim Sleeper case (D.N.A. Box 8.5)

Current NIST Projects

Short Overviews…

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

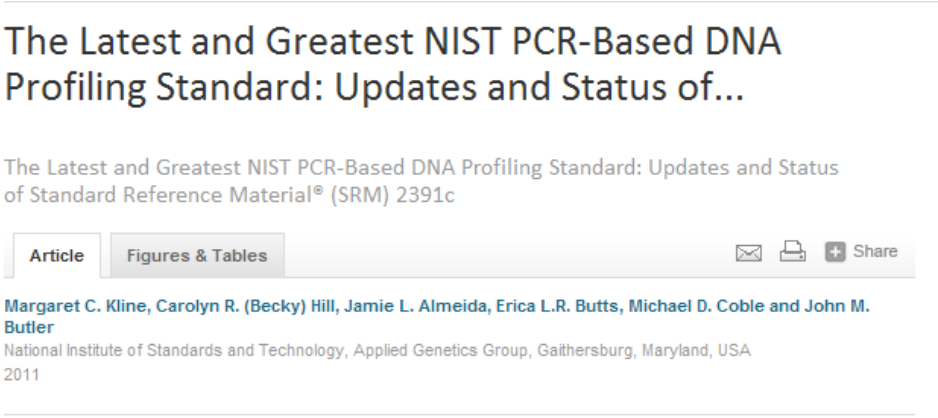
NIST SRM 2391c


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**


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

Presentations/Publications:

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

NIST Standard Reference Material (SRM) for Forensic DNA Testing

SRM 2391b (2003-2011)

- **48 autosomal STR loci** with certified values
- **10 liquid genomic DNA** components + **2 punches** (cells on 903 paper)
- All single source samples
- 4 males + 6 females
- 9947A & 9948 included

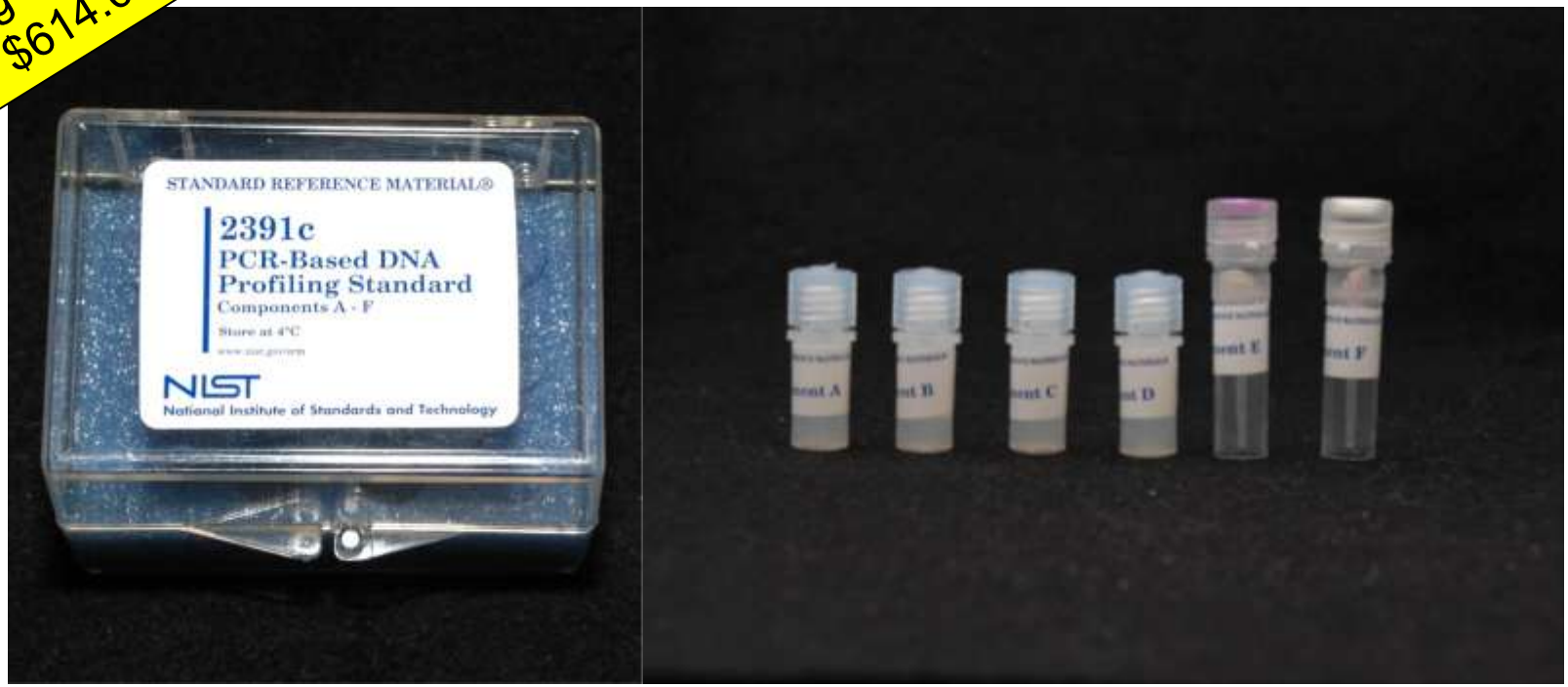
SRM 2391c (2011-future)

- **23 autosomal STR loci and 17 Y-STRs** certified
- **4 liquid genomic DNA** components + **2 punches** (cells on **FTA** & 903 paper)
- 5 single source + 1 mixture
- 3 males + 2 females (unique)
- All new samples
  - no 9947A or 9948

SRM 2391c to replace SRM 2391b and SRM 2395 (for Y-STRs)

Selling since
Aug 16, 2011
$614.00

NIST SRM 2391c



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 µL of anonymous **female** genomic DNA	1.4 – 1.9 ng DNA/µL
B	50 µL of anonymous **male** genomic DNA	1.3 – 1.5 ng DNA/µL
C	50 µL of anonymous **male** genomic DNA	1.3 – 2.0 ng DNA/µL
D	50 µL of **mixed-source** (Components A and C)	1.4 – 2.0 ng DNA/µL
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

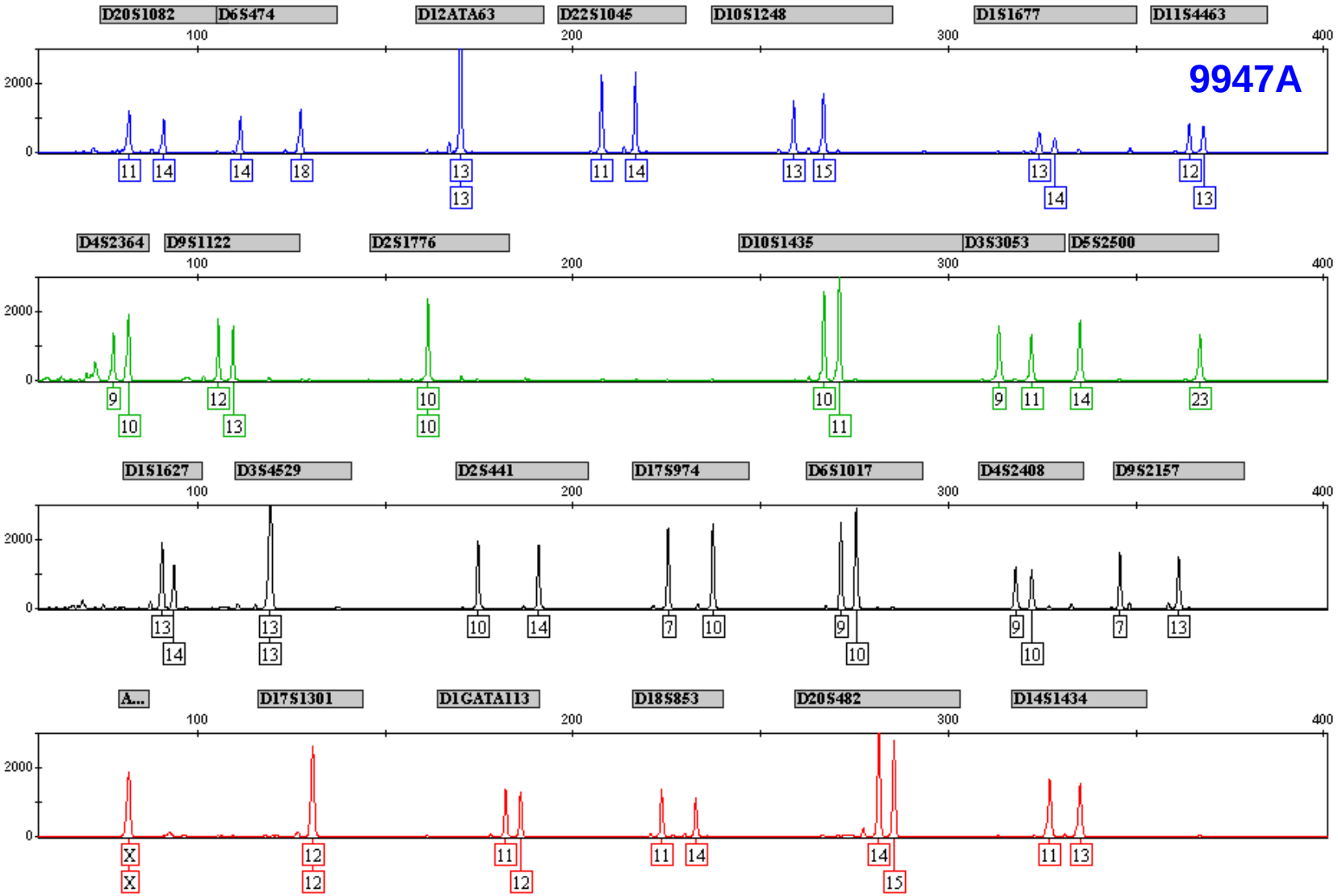
Kit Provider			Primer Mixes
Life Technologies	*Promega*	*Qiagen*	*NIST*
Identifiler	Powerplex 16	ESSplex	26plex
Identifiler Plus	Powerplex 16 HS	IDplex	miniSTRs
NGM	Powerplex ESX 17		
NGM SElect	Powerplex ESI 17		
COfiler	Powerplex ES		
Profiler	Powerplex S5		
Profiler Plus	**Powerplex Y**		
Profiler 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

NIST STR 26plex

Hill et al. (2009) *Journal of Forensic Sciences*, 54(5):1008-1015

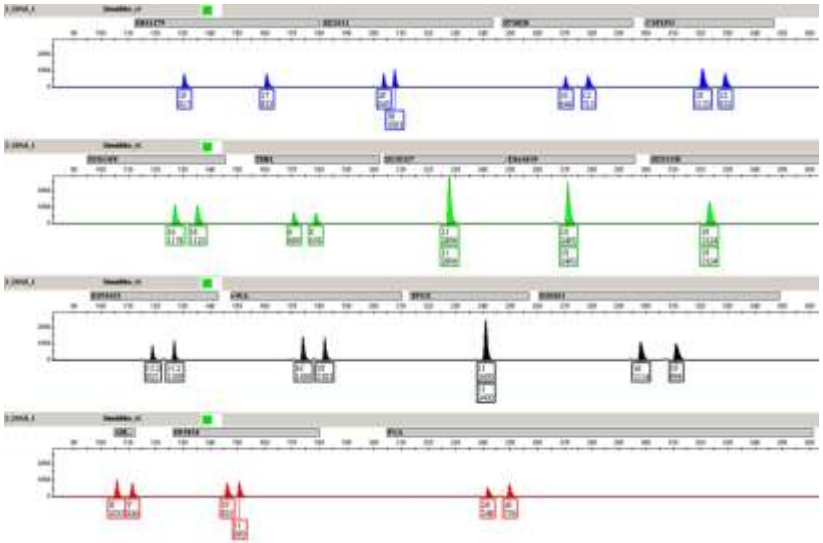


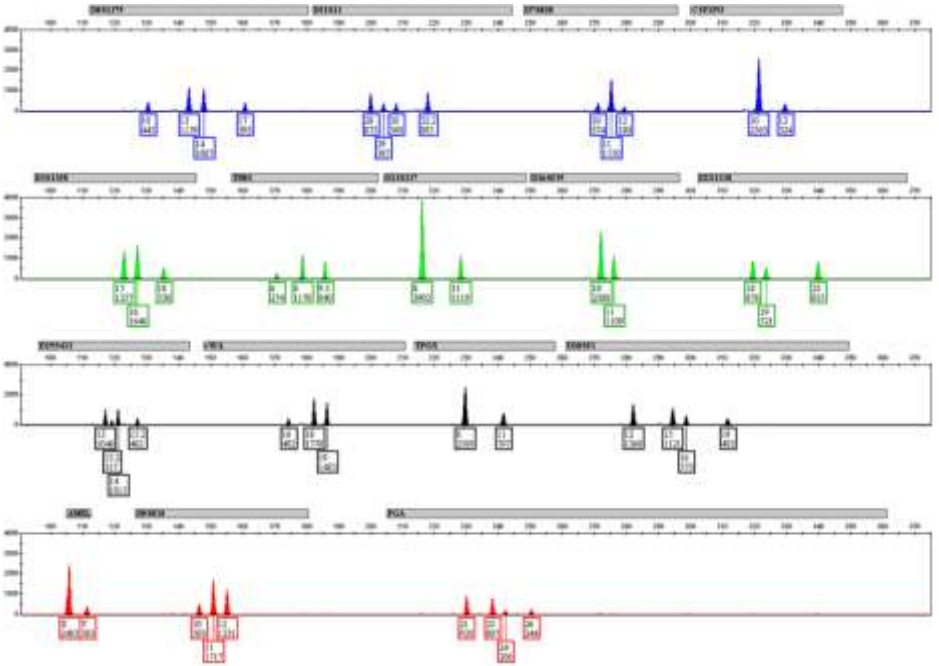
Gender identification + 25 autosomal STR loci in a single amplification

Component D

3 part A

1 part C





The certified ratio for Component D, the mass of Component A relative to that of Component C, is **3.1 ± 0.1 Component A / Component C.**

STR Kit Concordance Testing

Becky Hill

Main Points:

- When different primer sets are utilized, there is a concern that allele dropout may occur due to primer binding site mutations that impact one set of primers but not another
- To test SRM 2391b/2391c (PCR-based DNA Profiling Standard) components with all new STR multiplex kits and verify results against certified reference values
- To gain a better understanding of primer binding site mutations that cause null alleles

If no primer binding site mutations

Set 1 Amplicons Set 2 Amplicons

If a primer binding site mutation exists

Set 1 Amplicons Set 2 Amplicons

Presentations/Publications:

- *Profiles in DNA* article (Hill et al. 2010)
- ISFG 2011 and ISHI 2011 posters (Hill et al.)

Commercially Available STR Kits

Applied Biosystems (17)

- ~~AmpFlSTR Blue (1996)~~
- ~~AmpFlSTR Green I (1997)~~
- Profiler (1997)
- Profiler Plus (1997)
- COfiler (1998)
- SGM Plus (1999)
- **Identifiler** (2001)
- Profiler Plus ID (2001)
- ~~SEfiler (2002)~~
- Yfiler (2004)
- MiniFiler (2007)
- SEfiler Plus (2007)
- Sinofiler (2008) – China only
- **Identifiler Direct** (2009)
- NGM (2009)
- **Identifiler Plus** (2010)
- NGM SElect (2010)

Promega Corporation (13)

- PowerPlex 1.1 (1997)
- PowerPlex 1.2 (1998)
- PowerPlex 2.1 (1999)
- **PowerPlex 16** (2000)
- PowerPlex ES (2002)
- PowerPlex Y (2003)
- PowerPlex S5 (2007)
- **PowerPlex 16 HS** (2009)
- PowerPlex ESX 16 (2009)
- PowerPlex ESX 17 (2009)
- PowerPlex ESI 16 (2009)
- PowerPlex ESI 17 (2009)
- PowerPlex 18D (2011)
- PowerPlex 21 (2012)
- PowerPlex ESI 17 Pro (2012)

Qiagen (2010)

Primarily selling kits in Europe

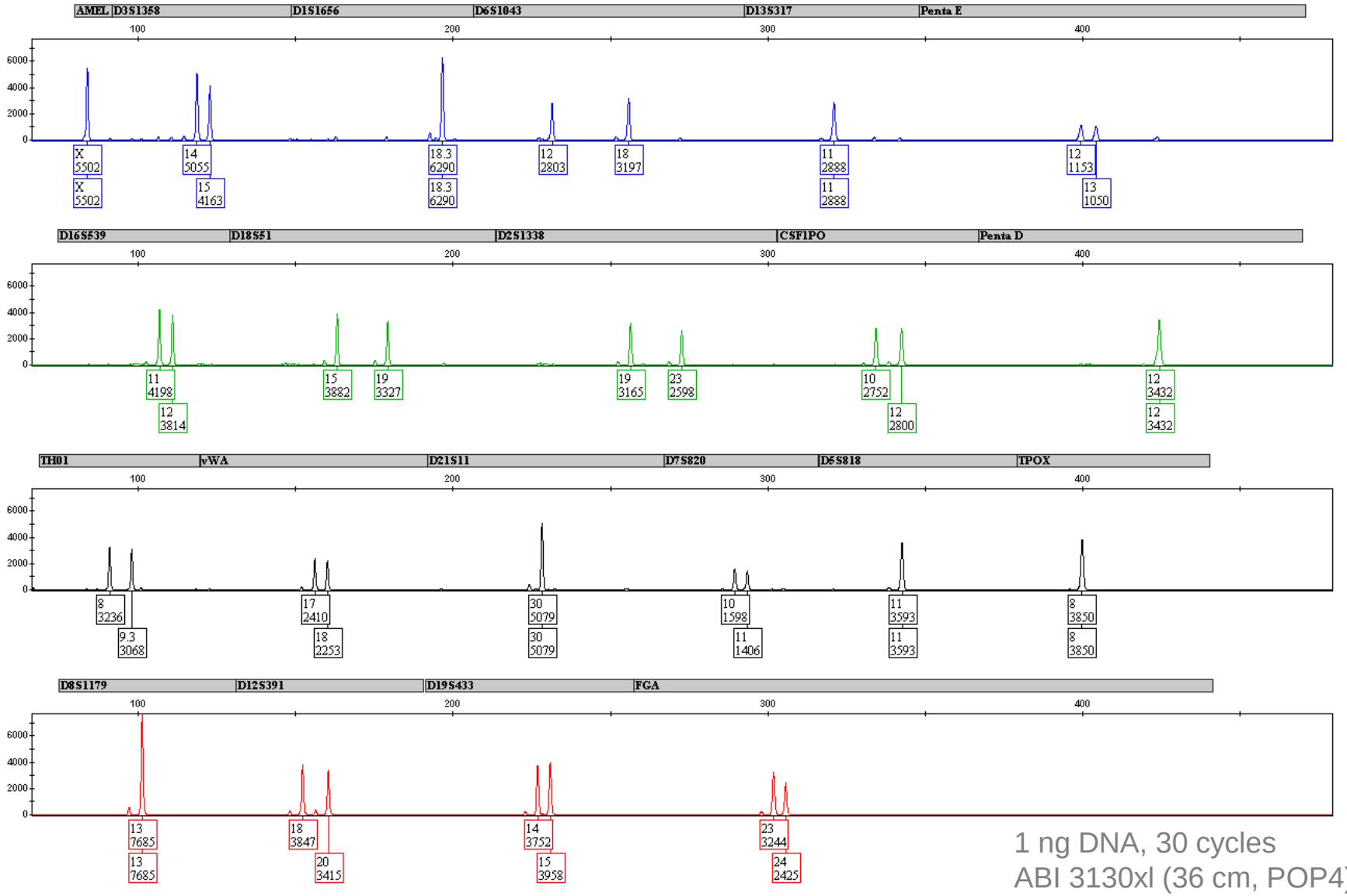
Due to patent restrictions cannot sell in U.S.

- ESSplex
- ESSplex SE
- Decaplex SE
- IDplex
- Nonaplex ESS
- Hexaplex ESS
- HD (Chimera)
- Argus X-12
- Argus Y-12
- DIPlex (30 InDels)

~1/3 of all STR kits were released in the last two years

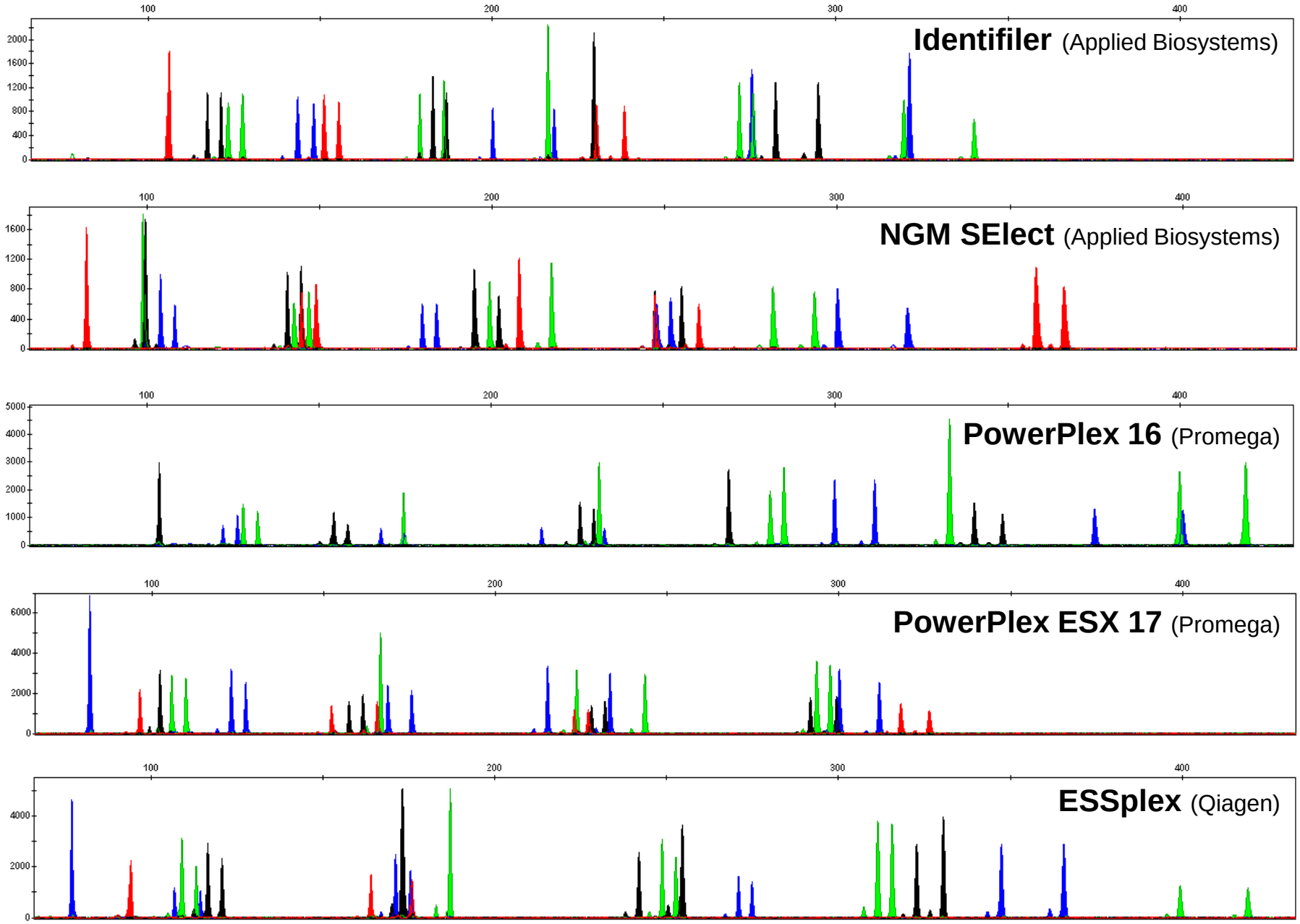
PowerPlex 21 NIST Result with 9947A

20 autosomal STR loci + amelogenin



1 ng DNA, 30 cycles
ABI 3130xl (36 cm, POP4)

Same DNA Sample Tested with Five STR Kits



STR Kit Concordance Testing

Profiles in DNA Article Published April 2010

Article Type: Feature

Volume 13 No. 1, April 2010

Strategies for Concordance Testing

Carolyn R. Hill, Margaret C. Kline, David L. Duewer and John M. Butler
National Institute of Standards and Technology, Biochemical Science Division, Gaithersburg, Maryland, USA

Concordance evaluatio
or "null alleles" present
commercial short tand
markers available to th
kits because the primer
(PCR) product sizes. W
may occur due to primer-binding-site mutations that affect one set of primers but not another.

4 S's of Concordance Testing

Standard samples (data on same samples)
Software (to check data concordance)
Sequencing (to understand null alleles)
STRBase (sharing with the community)

http://www.promega.com/profiles/1301/1301_08.html

NIST Standard Sample Sets

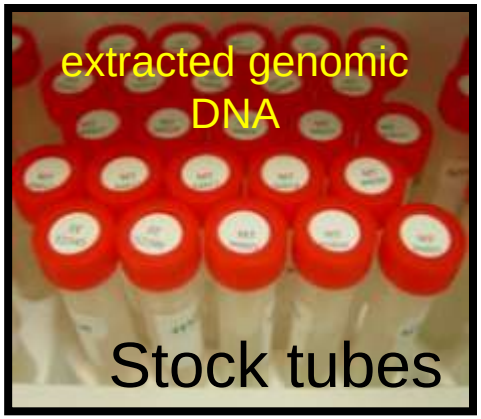
- **U.S. Population Samples** (**663 samples**)
  - Previously studied with Identifiler, MiniFiler, Yfiler, PP16, PP ESX/ESI 17, NGM, miniSTRs, and 23plex *(>200,000 allele calls)*
  - 260 African Americans, 260 Caucasians, 140 Hispanics, and 3 Asians

- **U.S. Father/Son pairs** (**800 samples**)
  - Previously studied with Identifiler, MiniFiler, Yfiler, PP ESX/ESI 17, NGM, 23plex
  - ~**100 fathers/100 sons for each group**: African Americans, Caucasians, Hispanics, and Asians

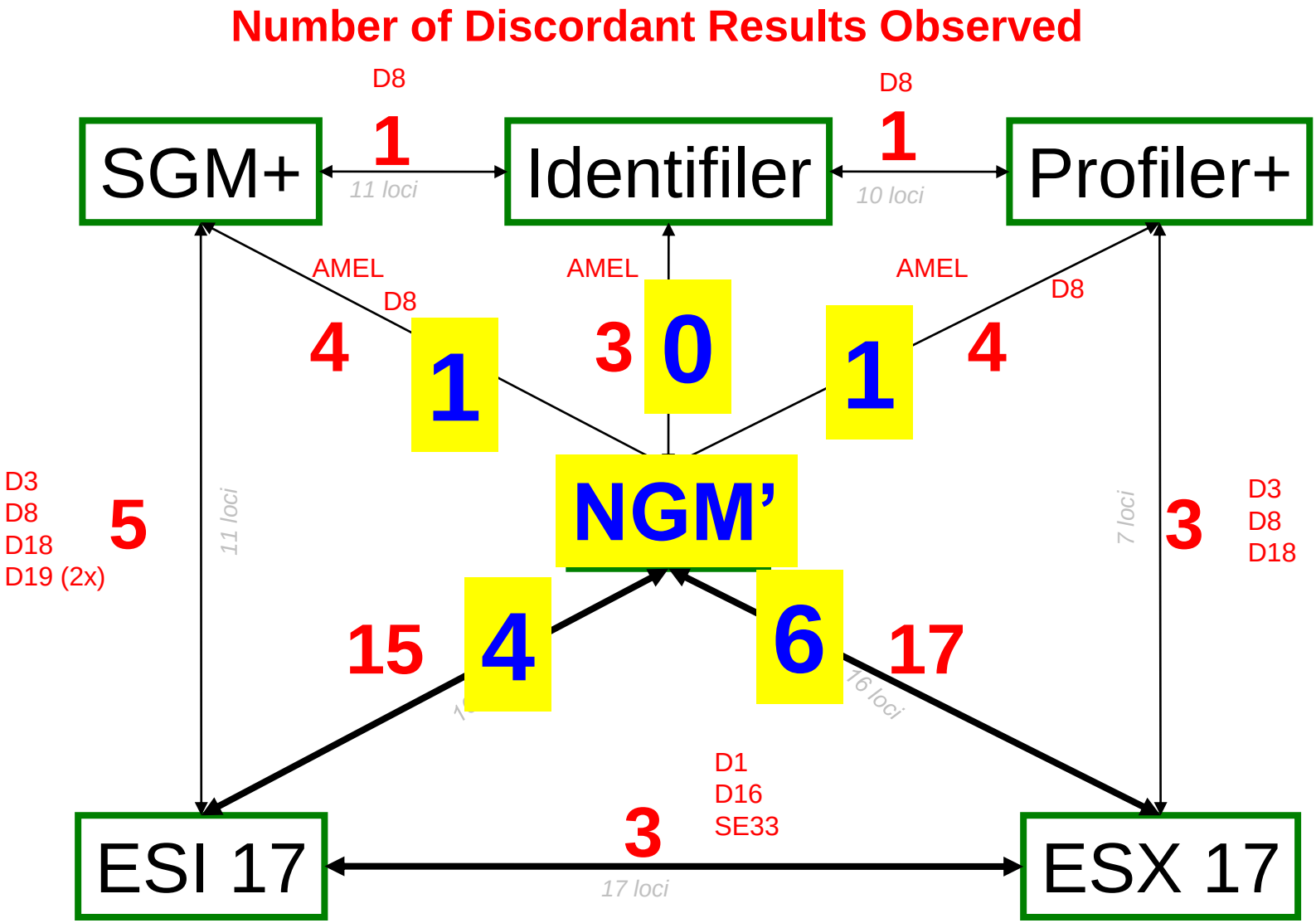
- **NIST SRM 2391b** PCR DNA Profiling Standard (**12 samples**)
  - Components 1-10 (includes 9947A and 9948): *well characterized*
  - ABI 007 and K562

>1450 total samples

http://www.cstl.nist.gov/biotech/strbase/NISTpop.htm



Initial Concordance Testing Summary



See also Hill, C.R., et al. (2010) Strategies for concordance testing. *Profiles in DNA (Promega)*, 13(1). Available at http://www.promega.com/profiles/1301/1301_08.html

Kit Concordance Comparisons

Kits compared	Samples	Loci compared	Comparisons	# Differences	Concordance (%)
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	11	15,939	3	99.981
ProPlus-NGM	1427				
SGM-ESI	1436				
ProPlus-ESX	1427				
ESI-ESX	1455				
ESI-ESSplex	1445				
ESX-ESSplex	1445				
ESI-NGMSElect	715				
ESX-NGMSElect	715				
ESS-NGMSElect	663	17	11,271	17	99.849
		TOTAL	240,156	186	99.923

> 1 million allele comparisons

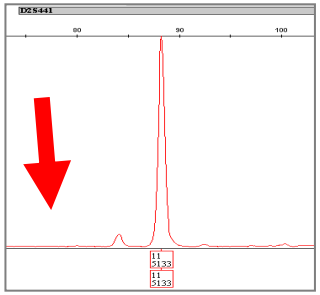
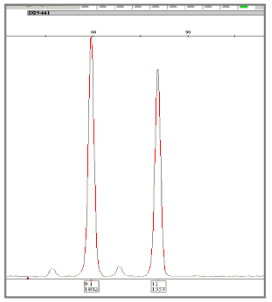
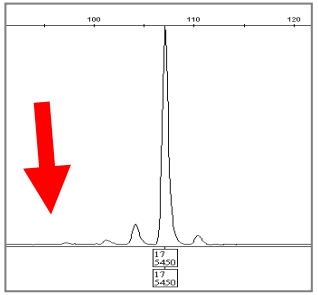
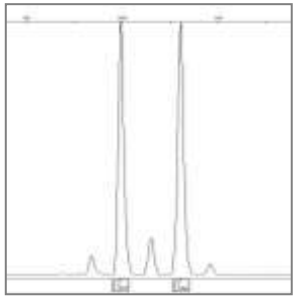
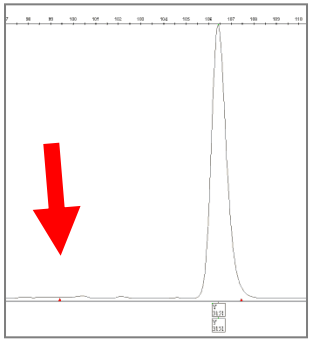
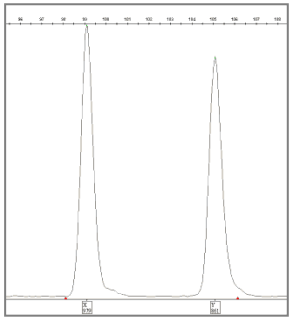
>1100 differences observed

~99.9% concordance

(many corrected now)

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

Extra (Degenerate) Primers Added with NGM SElect

	NGM (original)		NGM SElect and NGM’	
D2S441 9.1 allele missing in 7 Asians		11,11		9.1,11
D22S1045 15 allele missing in 4 samples		17,17		15,17
Amelogenin X allele missing in 3 samples		Y,Y		X,Y

Variant STR Allele Sequencing

Margaret Kline

Main Points:

- **STR allele sequencing has been provided free to the community** for the past ten years thanks to NIJ-funding
- Article provides primer sequences (outside of all known kit primers) for 23 autosomal STRs & 17 Y-STRs and full protocol for gel separations and sequencing reactions
  - 111 normal and variant alleles sequenced (at 19 STR & 4 Y-STRs)
  - 17 null alleles sequenced (with impact on various STR kit primers)

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Short communication

STR sequence analysis for characterizing normal, variant, and null alleles

Margaret C. Kline *, Carolyn R. Hill, Amy E. Decker [1], John M. Butler

National Institute of Standards and Technology, 100 Bureau Drive, M/S 8312, Gaithersburg, MD 20899, USA

Presentations/Publications:

- FSI Genetics article (Aug 2011) and numerous talks

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* (Aug 2011 issue)
- Butler & Hill (2011) *Forensic Sci Rev* (submitted)
- Hares (2011) Expanding the U.S. core loci… *FSI Genetics* (in press)

CODIS Core Loci Working Group

(formed in May 2010)

Douglas Hares – Chair
John Butler – NIST
Taylor Scott – ISP
Cecelia Crouse – PBSO
Brad Jenkins – VDFS
Ken Konzak – Cal DOJ

Announcing Plans to Expand the U.S. CODIS STR Core Loci

TICLE IN PRESS

Science International: Genetics xxx (2011) xxx–xxx

tents lists available at ScienceDirect

Forensic Science International: Genetics

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

Letter to the Editor

Expanding the CODIS core loci in the United States

Dear Editor:

After over a decade of operation, the National DNA Index System (NDIS) continues to grow in importance and size [1]. While the STR DNA technology has remained relatively consistent, other key aspects of the NDIS program have been reevaluated and revisions implemented. For example, based upon recommendations of the Scientific Working Group on DNA Analysis Methods, the Director of the Federal Bureau of Investigation (FBI) issued revised Quality Assurance Standards (QAS) for Forensic DNA

major reasons for expanding the CODIS core loci in the United States:

(1) To reduce the likelihood of adventitious matches [7] as the number of profiles stored at NDIS continues to increase each year (expected to total over 10 million profiles by the time of this publication). There are no signs that this trend will slow down as States expand the coverage of their DNA database programs and increase laboratory efficiency and capacity.

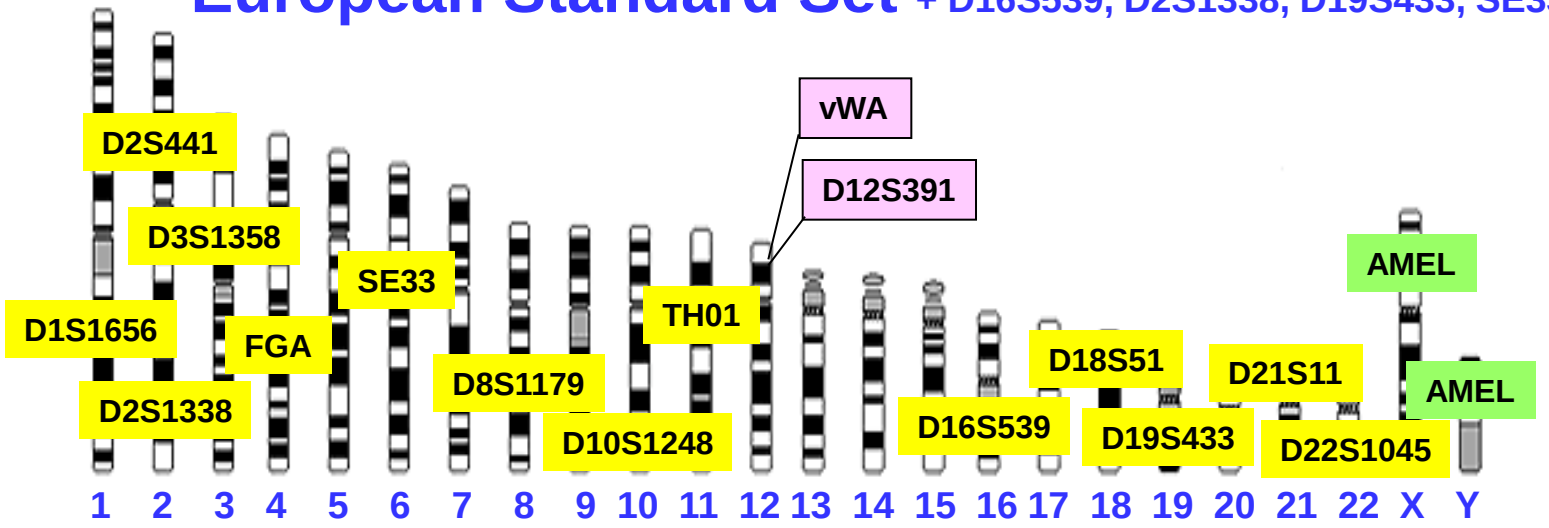
(2) To increase international compatibility to assist law enforcement data sharing efforts.

(3) To increase discrimination power to aid missing persons cases.

Common Forensic STR Loci

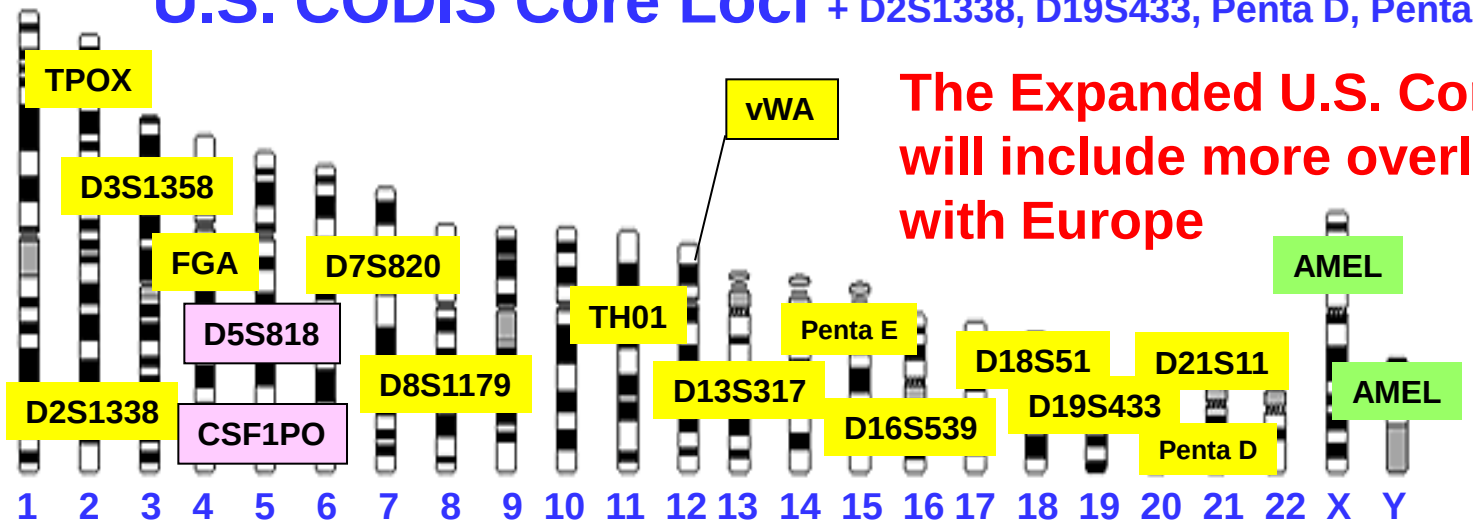
Europe

European Standard Set + D16S539, D2S1338, D19S433, SE33



United States

U.S. CODIS Core Loci + D2S1338, D19S433, Penta D, Penta E



The Expanded U.S. Core Set will include more overlap with Europe

The 11 STR Loci Beyond the CODIS 13

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

5 new European loci

**Allele range and number of observed alleles from Appendix 1, J.M. Butler (2011) Advanced Topics in Forensic DNA Typing: Methodology; ‡SE33 alleles have complex repeat structure*

SE33 (58 alleles observed)

343 genotypes observed
Heterozygosity = 0.9377

Allele	Total #	Total %	Populations, % Af Am	Asian	Cauc	Hisp
6.3						
7						
8						
10.2						
11	1	0.0			0.1	
11.2	2	0.1	0.2			
12	11	0.4	0.3		0.5	0.4
12.2	4	0.1	0.2			0.3
13	31	1.1	1.1		1.5	1.0
13.2	9	0.3	1.0			
14	85	2.9	5.1	0.2	2.5	2.4
14.2	10	0.3	0.4		0.4	0.3
15	102	3.5	3.9	1.2	3.9	3.9
15.2	8	0.3	0.3			0.7
16	144	5.0	4.8	4.7	4.0	6.7
16.2	5	0.2	0.3		0.1	0.1
16.3	2	0.1				0.3
17	205	7.1	9.3	4.0	6.2	7.3
17.2	1	0.0	0.1			
17.3	5	0.2	0.1		0.2	0.3
18	268	9.3	12.1	5.0	7.2	**11.0**
18.3	1	0.0			0.1	
19	250	8.7	**12.4**	6.2	6.6	8.0
19.2	8	0.3		0.2	0.4	0.4
20	216	7.5	10.9	9.2	5.4	4.8
20.2	20	0.7	0.3	1.2	1.1	0.3
21	108	3.7	4.6	6.7	2.4	2.7
21.2	48	1.7	1.1	1.7	2.4	1.3
22	42	1.5	1.3	1.7	1.5	1.3
22.2	65	2.3	0.4	3.2	3.8	1.9
23	12	0.4	0.6	1.0	0.2	0.1
23.2	91	3.2	2.2	4.2	4.3	2.1
24	1	0.0			0.1	
24.2	74	2.6	1.3	6.2	2.2	2.5
25.2	109	3.8	2.6	6.9	4.0	3.1
26	1	0.0	0.1			
26.2	163	5.6	6.1	5.2	4.3	7.1
27	1	0.0				0.1
27.2	225	7.8	4.3	**10.4**	**9.5**	8.6
27.3	2	0.1				0.3
28	2	0.1	0.1	0.2		
28.2	180	6.2	4.4	7.9	7.4	6.1
28.3	2	0.1	0.1		0.1	
29	1	0.0		0.2		
29.2	147	5.1	2.7	5.7	6.3	6.3
29.3	1	0.0		0.2		
30	1	0.0				0.1
30.2	111	3.8	1.6	3.2	5.8	4.6
31	3	0.1	0.1		0.2	
31.2	52	1.8	1.5	2.5	2.2	1.3
32	1	0.0			0.1	
32.2	25	0.9	0.4	0.7	1.3	0.9
33	2	0.1			0.1	0.1
33.2	11	0.4	0.3		0.5	0.4
34	9	0.3	0.3		0.7	
34.2	1	0.0			0.1	
35	1	0.0	0.1			
36	2	0.1	0.2			

SE33 Internal Sequence Variation

Same Length,
Different Internal Sequence

Allele (Repeat #)	ABI SEfiler	Promega ESX 17	Promega ESI 17	Repeat Motif Patterns															Reference
				AAAG	AG	AAAG	AG	**AAAG**	AAAAAG	AG	AGAAAG	AAAG	AAAAAG	AAAG	G	AAGG	AAAG/ANAG	AG	
				5' flanking				**central repeat**							3' flanking				
28.2 (a)	299 bp	360 bp	402 bp	2	1	3	1	8	1	0	0	19	0	0	1	1	2	1	Rolf *et al.* (1997)
28.2 (b)	299 bp	360 bp	402 bp	2	1	3	1	9	0	0	0	18	0	0	1	1	2	1	Rolf *et al.* (1997)
28.2 (c)	299 bp	360 bp	402 bp	2	1	3	1	9	0	0	0	15	0	0	1	1	2	1	Rolf *et al.* (1997)
28.2 (d)	299 bp	360 bp	402 bp	2	1	3	1	9	1	0	0	18	0	0	1	1	2	1	Rolf *et al.* (1997)
28.2 (e)	299 bp	360 bp	402 bp	2	1	3	1	10	1	0	0	17	0	0	1	1	2	1	Rolf *et al.* (1997)
28.2 (f)	299 bp	360 bp	402 bp	2	1	3	1	11	1	0	0	16	0	0	1	1	2	1	Rolf *et al.* (1997)
28.2 (g)	299 bp	360 bp	402 bp	2	1	3	1	12	1	0	0	15	0	0	1	1	2	1	Rolf *et al.* (1997)
28.2 (h)	299 bp	360 bp	402 bp	2	1	3	1	13	1	0	0	14	0	0	1	1	2	1	Rolf *et al.* (1997)
28.2 (i)	299 bp	360 bp	402 bp	2	1	3	1	14	1	0	0	13	0	0	1	1	2	1	Rolf *et al.* (1997)
28.2 (j)	299 bp	360 bp	402 bp	2	1	3	1	14	1	0	0	13	0	0	1	3	0	1	Rolf *et al.* (1997)
28.2 (k)	299 bp	360 bp	402 bp	2	1	3	1	16	1	0	0	11	0	0	1	1	2	1	Rolf *et al.* (1997)
28.3	300 bp	361 bp	403 bp	2	1	3	1	10	1	0	0	12	+A	4	1	1	2	1	Dauber *et al.* (2009)
29	301 bp	362 bp	404 bp	2	1	0	0	15	1	0	0	16	0	0	1	1	2	1	Dauber *et al.* (2009)
29.2 (a)	303 bp	364 bp	406 bp	2	1	3	1	8	1	0	0	20	0	0	1	1	2	1	Rolf *et al.* (1997)
29.2 (b)	303 bp	364 bp	406 bp	2	1	3	1	9	0	0	1	19	0	0	1	1	2	1	Rolf *et al.* (1997)
29.2 (c)	303 bp	364 bp	406 bp	2	1	3	1	9	1	0	0	19	0	0	1	1	2	1	Rolf *et al.* (1997)
29.2 (d)	303 bp	364 bp	406 bp	1	1	3	1	10	1	0	0	19	0	0	1	1	2	1	Rolf *et al.* (1997)
29.2 (e)	303 bp	364 bp	406 bp	2	1	3	1	11	0	5	0	16	0	0	1	1	2	1	Rolf *et al.* (1997)
29.2 (f)	303 bp	364 bp	406 bp	1	1	3	1	11	1	0	0	18	0	0	1	1	2	1	Rolf *et al.* (1997)
29.2 (g)	303 bp	364 bp	406 bp	2	1	3	1	11	1	0	0	17	0	0	1	1	2	1	Rolf *et al.* (1997)
29.2 (h)	303 bp	364 bp	406 bp	2	1	3	1	12	1	0	0	16	0	0	1	1	2	1	Rolf *et al.* (1997)
29.2 (i)	303 bp	364 bp	406 bp	2	1	3	1	13	0	0	1	15	0	0	1	3	0	1	Rolf *et al.* (1997)
29.2 (j)	303 bp	364 bp	406 bp	2	1	3	1	13	1	0	0	15	0	0	1	1	2	1	Rolf *et al.* (1997)
29.2 (k)	303 bp	364 bp	406 bp	2	1	3	1	14	1	0	0	14	0	0	1	1	2	1	Rolf *et al.* (1997)
29.2 (l)	303 bp	364 bp	406 bp	2	1	3	1	16	1	0	0	12	0	0	1	1	2	1	Rolf *et al.* (1997)
29.2 (m)	303 bp	364 bp	406 bp	2	1	3	1	11	1	0	0	17	0	0	1	1	2	1	D41-TTG-deletion -- Kline *et al.* (2010)

Allele 28.2
(11 sequences)

Allele 29.2
(13 sequences)

25 Alleles Reported in the Literature for **D1S1656**

15 NIST observed alleles circled in red

Allele (Repeat #)	Promega ESX 17	Promega ESI 17	ABI NGM	Repeat Structure $[TAGA]_4[TGA]_{0-1}[TAGA]_nTAGG[TG]_5$	Reference
8	133 bp	222 bp	171 bp	$[TAGA]_8[TG]_5$	Phillips *et al.* (2010)
9	137 bp	226 bp	175 bp	$[TAGA]_9[TG]_5$	Phillips *et al.* (2010)
10 (a)	141 bp	230 bp	179 bp	$[TAGA]_{10}[TG]_5$	Lareu *et al.* (1998)
10 (b)	141 bp	230 bp	179 bp	$[TAGA]_{10}TAGG[TG]_5$	Phillips *et al.* (2010)
11	145 bp	234 bp	183 bp	$[TAGA]_{11}[TG]_5$	Lareu *et al.* (1998)
12 (a)	149 bp	238 bp	187 bp	$[TAGA]_{12}[TG]_5$	Lareu *et al.* (1998)
12 (b)	149 bp	238 bp	187 bp	$[TAGA]_{11}TAGG[TG]_5$	Lareu *et al.* (1998)
13 (a)	153 bp	242 bp	191 bp	$[TAGA]_{12}TAGG[TG]_5$	Lareu *et al.* (1998)
13 (b)	153 bp	242 bp	191 bp	$[TAGA]_{13}[TG]_5$	Phillips *et al.* (2010)
13.3	156 bp	245 bp	194 bp	$[TAGA]_1TGA[TAGA]_{11}TAGG[TG]_5$	Phillips *et al.* (2010)
14 (a)	157 bp	246 bp	195 bp	$[TAGA]_{13}TAGG[TG]_5$	Lareu *et al.* (1998)
14 (b)	157 bp	246 bp	195 bp	$[TAGA]_{14}[TG]_5$	Phillips *et al.* (2010)
14.3	160 bp	249 bp	198 bp	$[TAGA]_4TGA[TAGA]_9TAGG[TG]_5$	Phillips *et al.* (2010)
15	161 bp	250 bp	199 bp	$[TAGA]_{14}TAGG[TG]_5$	Lareu *et al.* (1998)
15.3	164 bp	253 bp	202 bp	$[TAGA]_4TGA[TAGA]_{10}TAGG[TG]_5$	Lareu *et al.* (1998)
16	165 bp	254 bp	203 bp	$[TAGA]_{15}TAGG[TG]_5$	Lareu *et al.* (1998)
16.3	168 bp	257 bp	206 bp	$[TAGA]_4TGA[TAGA]_{11}TAGG[TG]_5$	Lareu *et al.* (1998)
17	169 bp	258 bp	207 bp	$[TAGA]_{16}TAGG[TG]_5$	Lareu *et al.* (1998)
17.1	170 bp	259 bp	208 bp	Not published	Schröer *et al.* (2000)
17.3	172 bp	261 bp	210 bp	$[TAGA]_4TGA[TAGA]_{12}TAGG[TG]_5$	Lareu *et al.* (1998)
18	173 bp	262 bp	211 bp	$[TAGA]_{17}TAGG[TG]_5$	Phillips *et al.* (2010)
18.3	176 bp	265 bp	214 bp	$[TAGA]_4TGA[TAGA]_{13}TAGG[TG]_5$	Lareu *et al.* (1998)
19	177 bp	266 bp	215 bp	Not published	Asamura *et al.* (2008)
19.3	180 bp	269 bp	218 bp	$[TAGA]_4TGA[TAGA]_{14}TAGG[TG]_5$	Lareu *et al.* (1998)
20.3	184 bp	273 bp	222 bp	Not published	Gamero *et al.* (2000)

from Appendix 1, J.M. Butler (2011) Advanced Topics in Forensic DNA Typing: Methodology

NIST **U.S. Population Allele Frequencies**

D1S1656 (15 different alleles)

15 different alleles

Allele	African American (N = 341)	Caucasian (N = 361)	Hispanic (N = 236)
10	0.01433	0.00277	0.00630
11	0.04871	0.07756	0.02731
12	0.06304	0.11773	0.08824
13	0.10029	0.06648	0.11555
14	**0.25788**	0.07895	0.11765
14.3	0.00716	0.00277	0.00420
15	0.15616	**0.14820**	0.13866
15.3	0.03009	0.05817	0.05042
16	0.11032	0.13573	**0.17437**
16.3	0.10029	0.06094	0.05462
17	0.02865	0.04709	0.04202
17.3	0.05014	0.13296	0.14496
18	0.00287	0.00554	0.00630
18.3	0.02436	0.05125	0.02521
19.3	0.00573	0.01385	0.00420

N = 938

(only unrelated samples used; fathers removed from this sample set)

< 5/2N

D1S1656 Characteristics

- **15 alleles** observed
- **92 genotypes** observed
- **>89% heterozygotes** (heterozygosity = 0.8934)
- **0.0220 Probability of Identity** (P_I)

$$P_I = \sum (genotype\ frequencies)^2$$

These values have been calculated for all 24 STR loci across the U.S. population samples examined

Loci sorted on Probability of Identity (P_I) values

STR Locus	Alleles Observed	Genotypes Observed	Het. (obs)	P_I value N = 938
SE33	53	292	0.9360	0.0069
Penta E*	20	114	0.8799	0.0177
D2S1338	13	68	0.8785	0.0219
D1S1656	15	92	0.8934	0.0220
D18S51	21	91	0.8689	0.0256
D12S391	23	110	0.8795	0.0257
FGA	26	93	0.8742	0.0299
Penta D*	16	71	0.8754	0.0356
D21S11	25	81	0.8358	0.0410
D19S433	16	76	0.8124	0.0561
D8S1179	11	45	0.7878	0.0582
vWA	11	38	0.8060	0.0622
D7S820	11	32	0.8070	0.0734
TH01	8	24	0.7580	0.0784
D16S539	9	28	0.7825	0.0784
D13S317	8	29	0.7655	0.0812
D10S1248	12	39	0.7825	0.0837
D2S441	14	41	0.7772	0.0855
D3S1358	11	30	0.7569	0.0873
D22S1045	11	42	0.7697	0.0933
CSF1PO	9	30	0.7537	0.1071
D5S818	9	34	0.7164	0.1192
TPOX	9	28	0.6983	0.1283

23 STR Loci present in STR kits

rank ordered by their variability

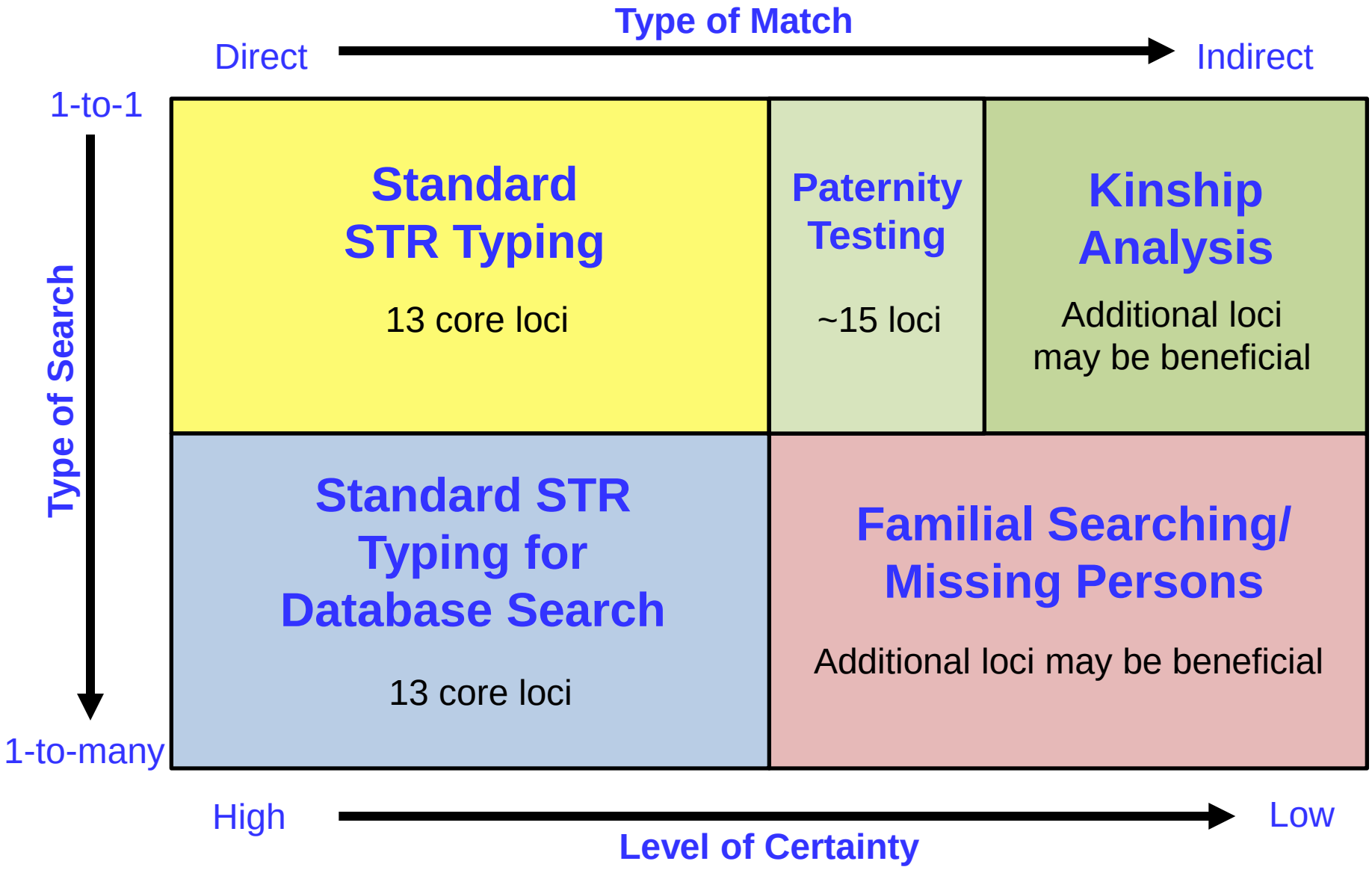
Better for mixtures (more alleles seen)

D6S1043
data not shown

There are several loci more polymorphic than the **current CODIS 13 STRs**

Better for kinship (low mutation rate)

Expanding the Forensic Core Competency



New STR Loci Characterized

Hill *et al.* (2008) *J. Forensic Sci*. 53(1):73-80

J Forensic Sci, January 2008, Vol. 53, No. 1
doi: 10.1111/j.1556-4029.2008.00595.x
Available online at: www.blackwell-synergy.com

Carolyn R. Hill, M.S.; Margaret C. Kline, M.S.; Michael D. Coble,† Ph.D.; and John M. Butler, Ph.D.

Characterization of 26 MiniSTR Loci for Improved Analysis of Degraded DNA Samples

- Primer sequences (for miniplexes), GeneMapper bins and panels, genotypes on common samples, and allele frequency information **available on STRBase**

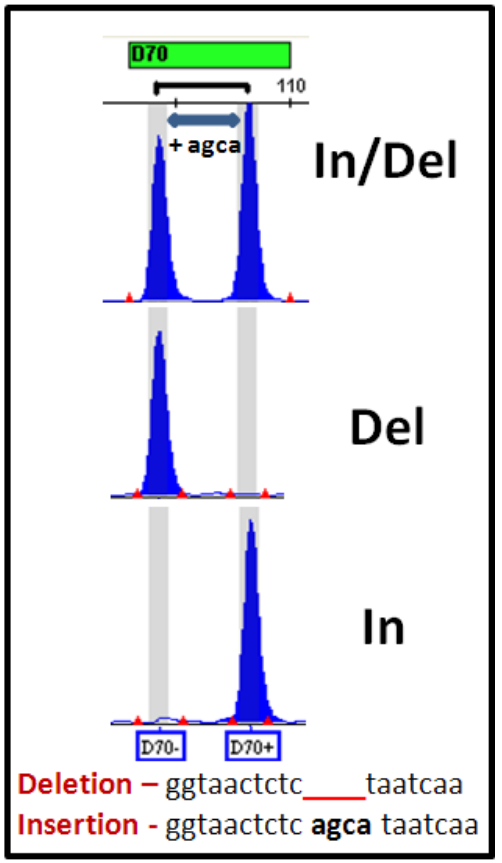
http://www.cstl.nist.gov/biotech/strbase/miniSTR.htm
http://www.cstl.nist.gov/biotech/strbase/miniSTR/miniSTR_NC_loci_types.htm
http://www.cstl.nist.gov/biotech/strbase/miniSTR/miniSTR_Panels_Panels.txt
http://www.cstl.nist.gov/biotech/strbase/miniSTR/miniSTR_Panels_NC_bins_bins.txt

Insertion/Deletion (InDel) Markers

Manuel **Fonde**vila 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



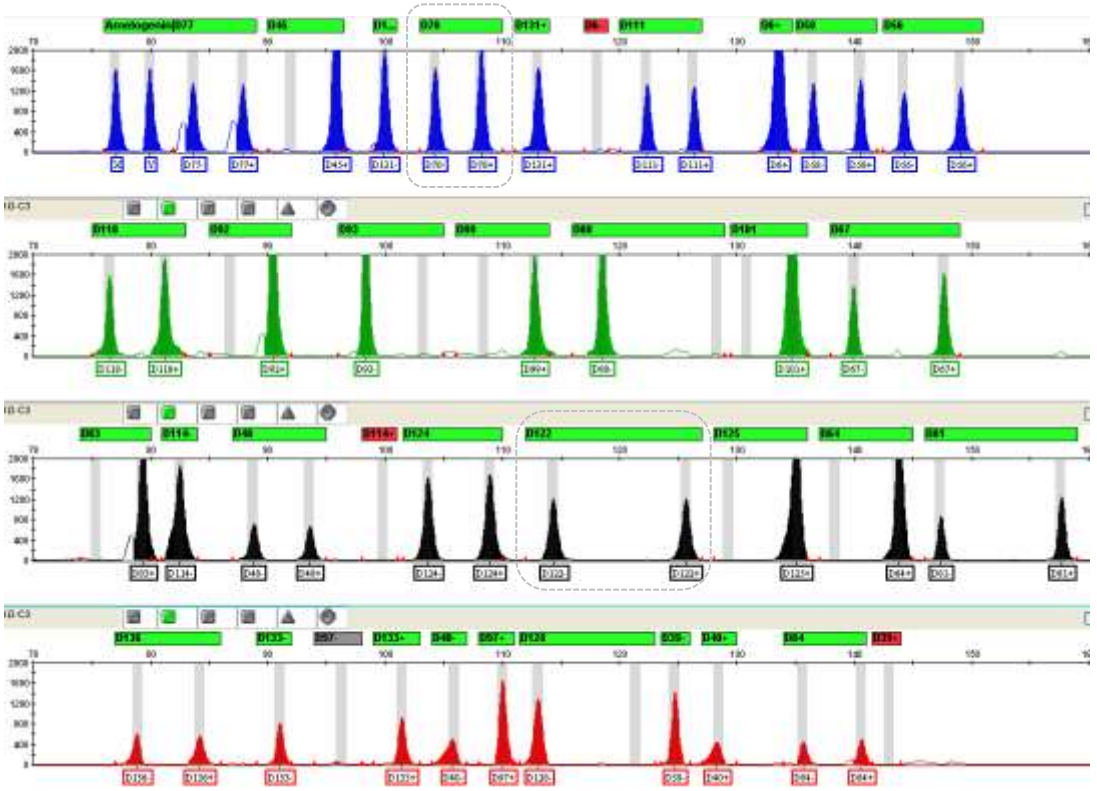
Presentations/Publications:

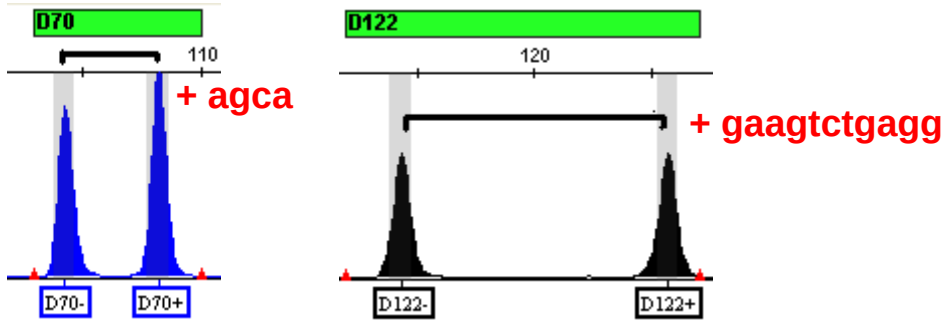
- FSI Genetics Suppl. Series 2011 article
- ISFG 2011 poster and ISHI 2011 presentation

DIPlex Insertion/Deletion Assay (Qiagen kit)

- Bi-allelic length polymorphisms with properties like SNPs
- PCR/CE detection properties like STRs
- 30 In/Dels (‘-’ or ‘+’ allele)
- Short amplicons (75-160 bp)
- Sensitive to ~100 pg with 30 cycle PCR
- Kits kindly provided by Qiagen

- Work performed by Manuel **Fonde**vila Alvarez (Santiago de Compostela, SPAIN), guest researcher at NIST

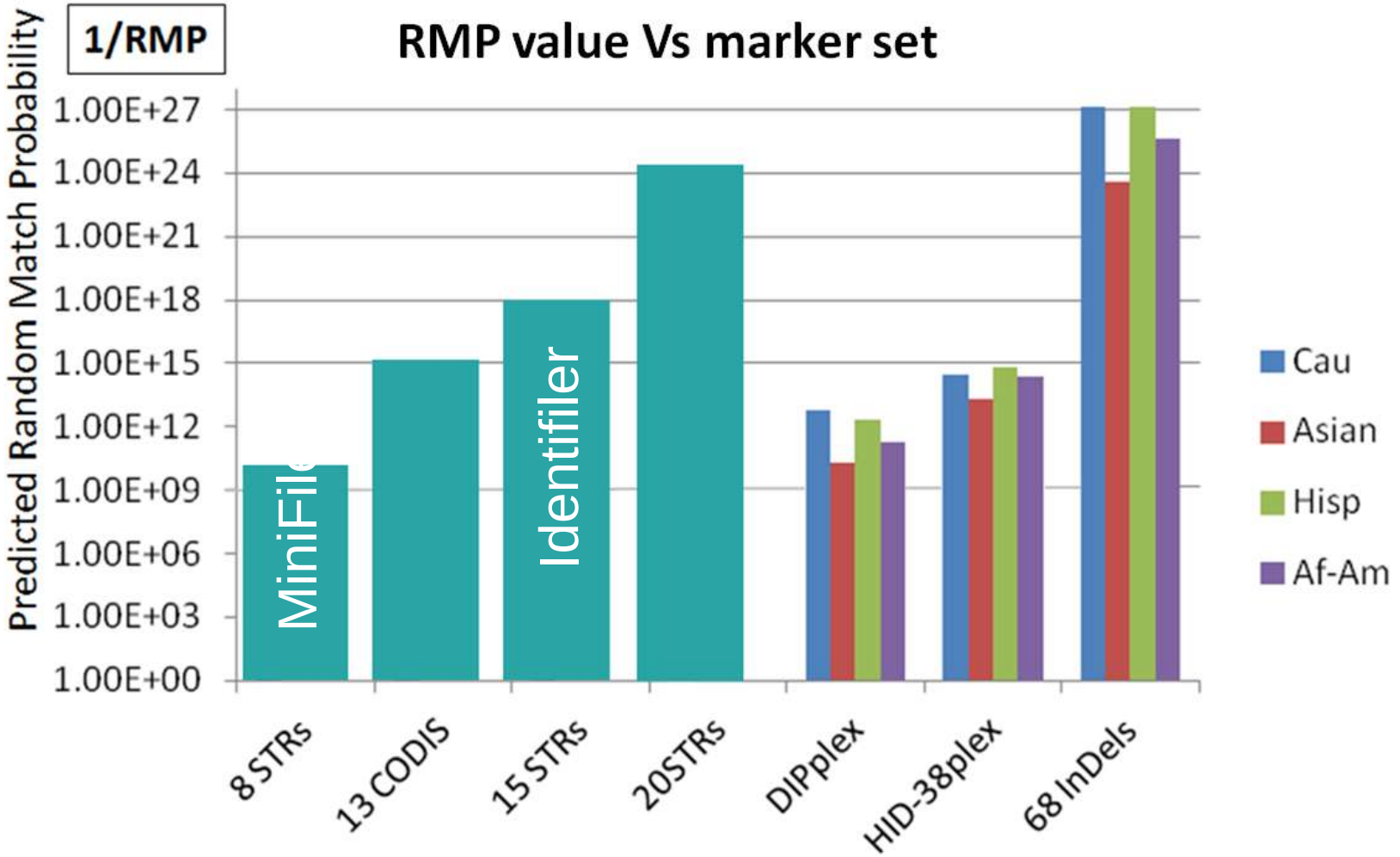




Heterozygous alleles with different insertion lengths

STR vs InDel Profile Frequency Comparisons

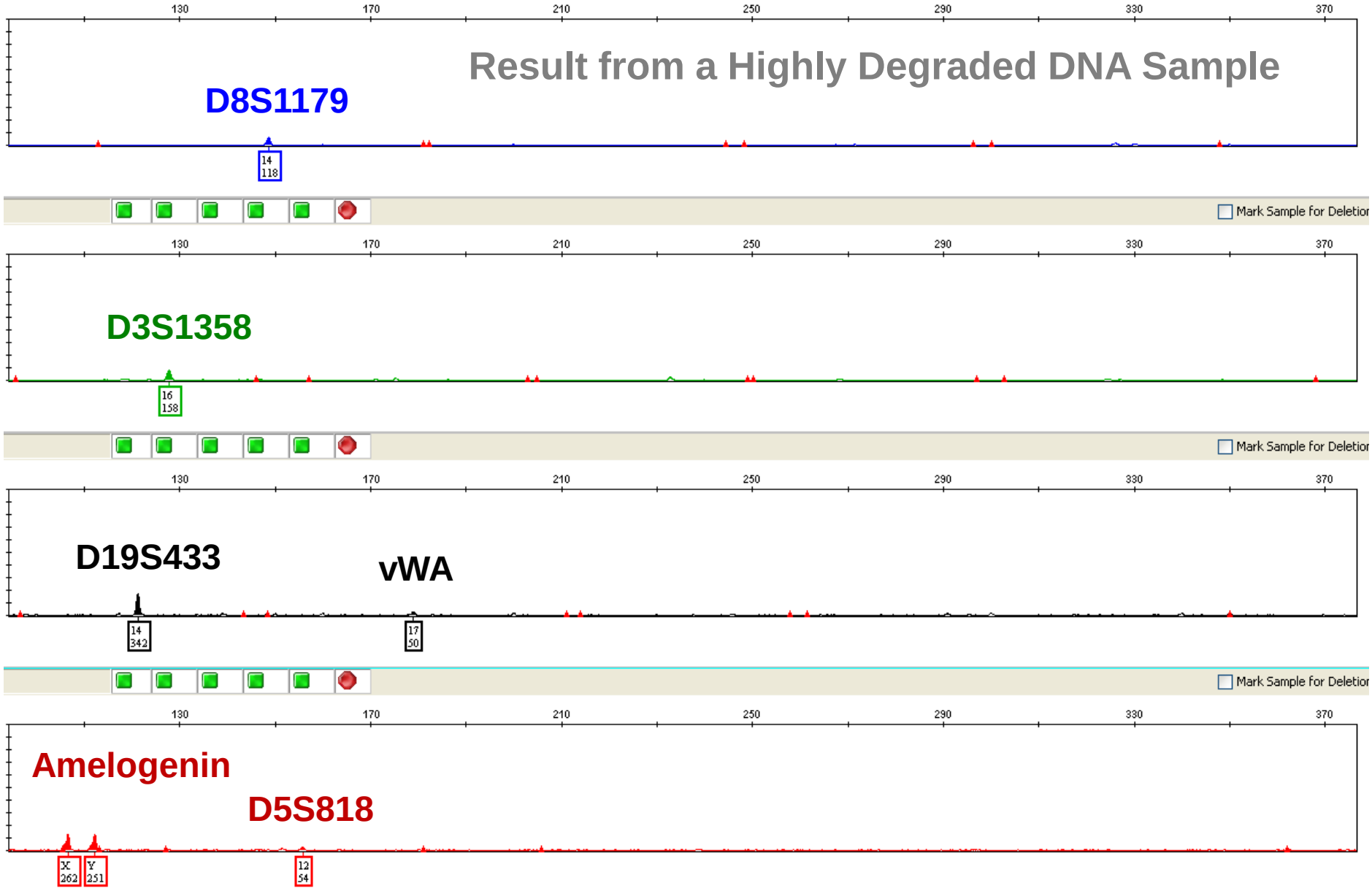
Each individual InDel assay supples an average RMP value that is lower than the 13 CODIS STRs while the two InDel assays together (68 InDel markers) supply a discrimination power higher than 20 STRs



All profiles shown scaled to 2000 RFUs

28 PCR cycles

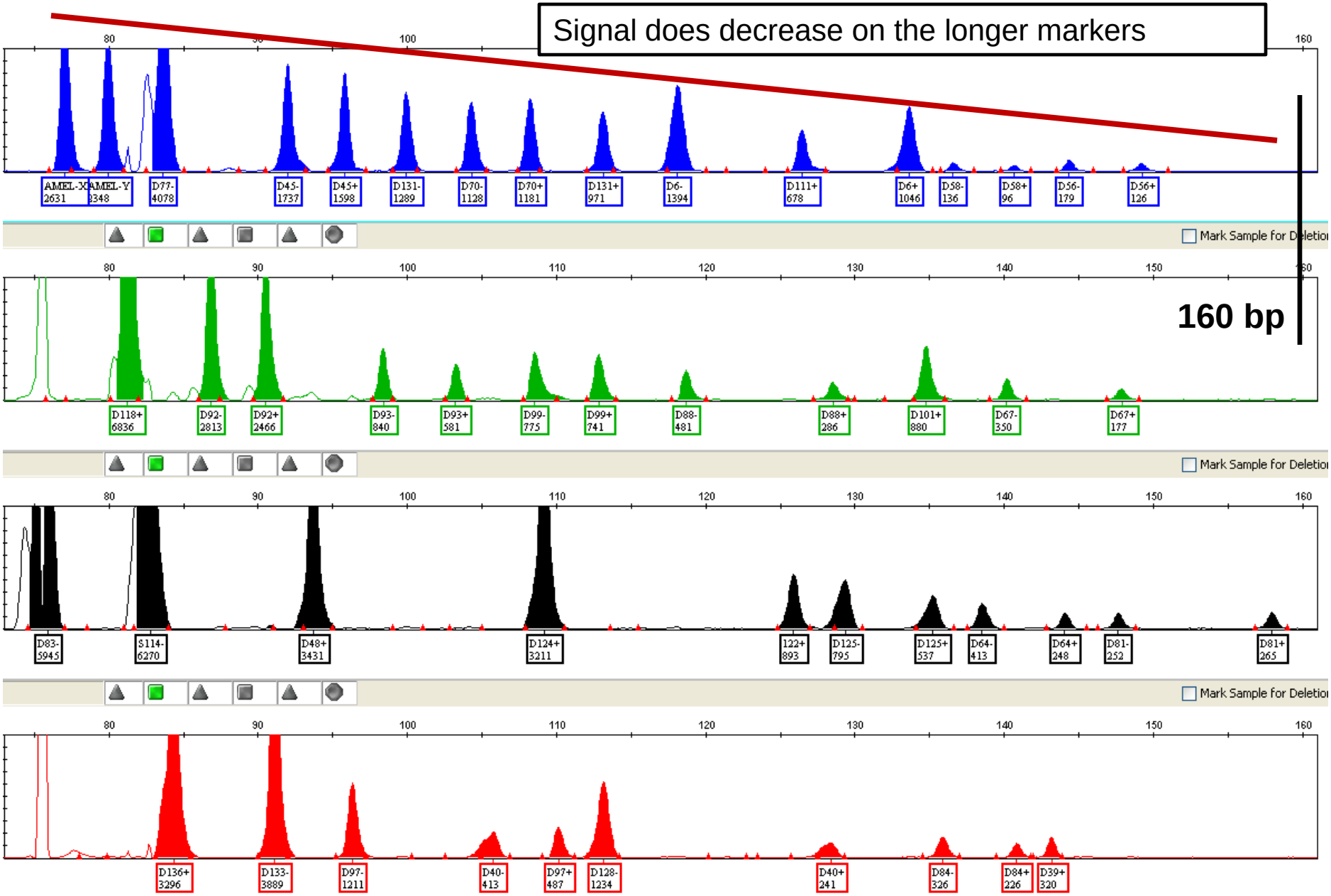
Identifiler – 7 alleles detected



All profiles shown scaled to 2000 RFUs

30 PCR cycles

DIPplex – 49 alleles detected



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

Common Thermal Cycling Times

Can we reduce PCR cycling times? What are the effects or limitations?

Thermal Cycling Times for Current STR Typing Kits

Year	Run on a 9700 thermal cycler	Hot start	Time per cycle	Cycles	Post soak	Total time
1997/98	Profiler Plus/Cofiler	11 min	3 min	28	60 min	2:52
1999	SGM Plus	11 min	3 min	28	45 min	2:53
2000	PowerPlex 16	12 min	1 min 45 s	32	30 min	3:00
2001	Identifiler	11 min	3 min	28	60 min	2:58
2003	PowerPlex Y	12 min	1 min 45 s	32	30 min	3:18
2004	Yfiler	11 min	3 min	30	80 min	2:45
2007	PowerPlex S5	2 min	4 min	30	45 min	3:21
2007	minifiler	11 min	3 min 20 s	30	45 min	3:16
2009	ESI 16, 17 ESX 16,17	2 min	4 min	30	45 min	3:22
2009	PowerPlex 16 HS	2 min	1 min 45 s	32	30 min	2:42
2009	NGM	11 min	3 min 20 s	29	10 min	2:33
2009	Identifler Direct	11 min	3 min	26	25 min	2:34
2010	Idenfiler Plus	11 min	3 min 20 s	28	10 min	2:18
2011	PowerPlex 18D	2 min	1 min 10s	27	20 min	1:25

Thermal Cyclers

1. GeneAmp 9700 (Applied Biosystems)
2. Mastercycler Pro S (Eppendorf)
   - Peltier based
3. Rotor-Gene Q (Qiagen)
   - Air heated and cooled
4. SmartCycler (Cepheid)
   - Hot plates for heating, fans for cooling

Intended for real-time PCR

- Cycling for most STR kits is run in the
- '9600 emulation mode' (1°C/s)

1

2

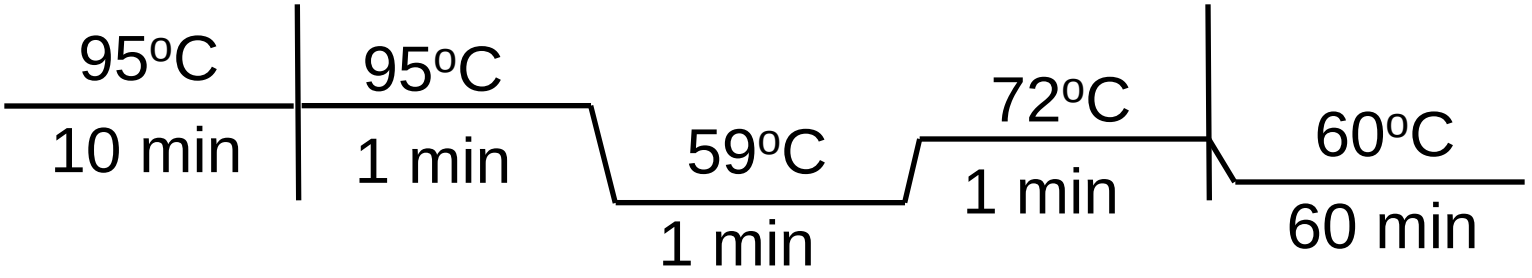
3

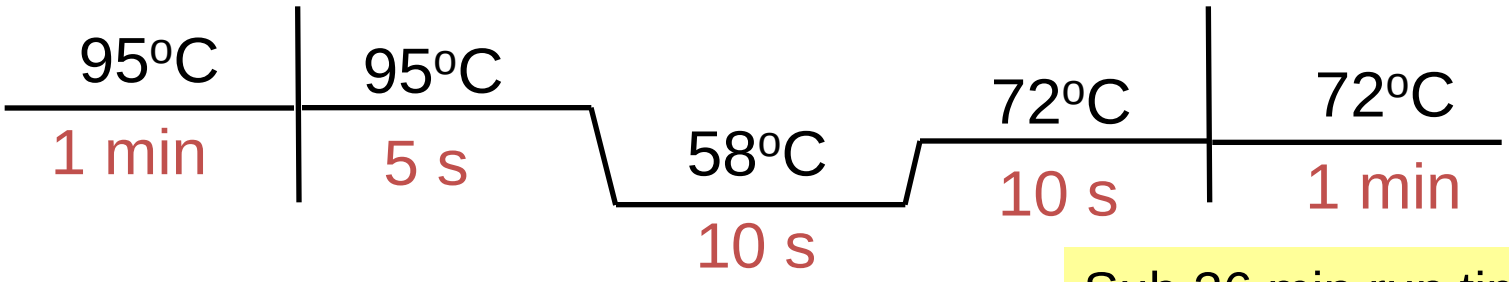
4

PCR Thermal Cycling Profile

Identifiler STR kit

28 cycles of PCR





Sub 36 min run time

Maximum heating/cooling rate of ~2 to 6°C/s (cycler dependent)

Rapid PCR Conditions

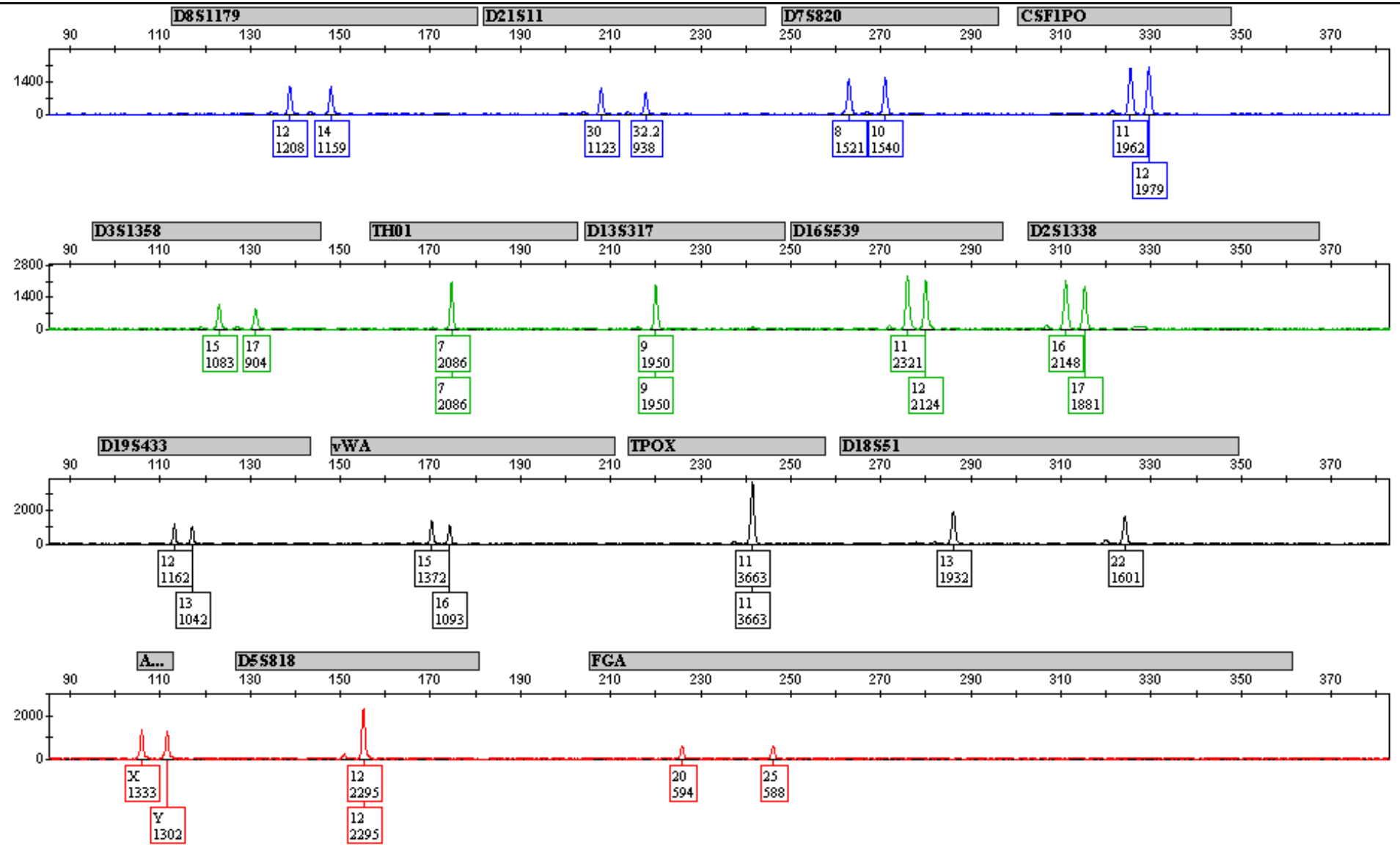
- 1 X Takara PCR mastermix, 1 U SpeedStar polymerase
  - *Premix Ex Taq*™ (Perfect Real Time)
- 10 μL total reaction in a thin walled tube (8-strip)
- 2 μL of Identifiler PCR primer mix
- ~1 ng of template DNA

- Utilize maximum ramp rate on thermal cyclers
  - GeneAmp 9700 = 1.6ºC/s (36 min)
  - Rotor-Gene Q = 1.6ºC/s (36 min)
  - SmartCycler = 5.8ºC/s (20 min)
  - Mastercycler Pro S = 6.8ºC/s (19 min)

Effective heating/cooling rates

Identifiler STR Profile

Mastercycler Pro S - 19 min PCR



Peltier Cyclers: 9700 and Mastercycler Pro S

Standard Injection

9700 36 min	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	AMEL	D5S818	FGA
1000 pg																
750 pg																
500 pg																
250 pg																
125 pg																
62.5 pg									51							
									180							
31.3 pg						85	71				60			134		56
		61		77		50	75	140						54		

full profile	
full drop out	
one allele	*number corresponds to the height of the peak that is called
drop in/high stutter	

Standard Injection

Mastercycler 19 min	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	AMEL	D5S818	FGA
1000 pg																
750 pg																
500 pg																
250 pg									59							
125 pg							65		165							
							50						54			
62.5 pg																
		67				63	56	94	127		93					51
31.3 pg																
									64				62			

full profile	
full drop out	
one allele	*number corresponds to the height of the peak that is called
drop in/high stutter	

Recent Training Workshops


John Butler


Mike Coble

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

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

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

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

Mixture Workshop (Promega ISHI 2010)

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

Handout >200 pages

Literature list of >100 articles

13 Modules Presented

Introductions (Robin)
SWGDAM Guidelines (John)
Analytical thresholds (Catherine)
Stutter (Mike)
Stochastic effects (Robin)
Peak height ratios (Charlotte)
Number of contributors (John)
Mixture ratios (John)
Mixture principles (Charlotte)
Statistics (Mike)

Case Example 1 (Robin)
Case Example 2 (Charlotte)
Case Example 3 (John)

Catherine Grgicak
Boston U.

Mike Coble
NIST

Robin Cotton
Boston U.

John Butler
NIST

Charlotte Word
Consultant

NIJ Grant to Boston University
funded ~150 state & local lab analysts to attend

AAFS 2011 Mixture Workshop

February 22, 2011 (Chicago, IL)

DNA Mixture Analysis: Principles and Practice of Mixture Interpretation and Statistical Analysis Using the SWGDAM STR Interpretation Guidelines

Topics (Speakers)

SWGDAM Guidelines (John Butler)
Mixture Fundamentals (Mike Adamowicz)
Validation & Thresholds (Joanne Sgueglia)
Mixture Statistics (Todd Bille)
Case Summary Analysis (John Butler)
Worked Case Example (Mike Coble)
Complex Mixtures (Gary Shutler)
Software Survey (Mike Coble)
Updating Protocols (Jennifer Gombos)
Training Staff (Ray Wickenheiser)

~220 people attended

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

TrueAllele Mixture Software Evaluation


Mike Coble

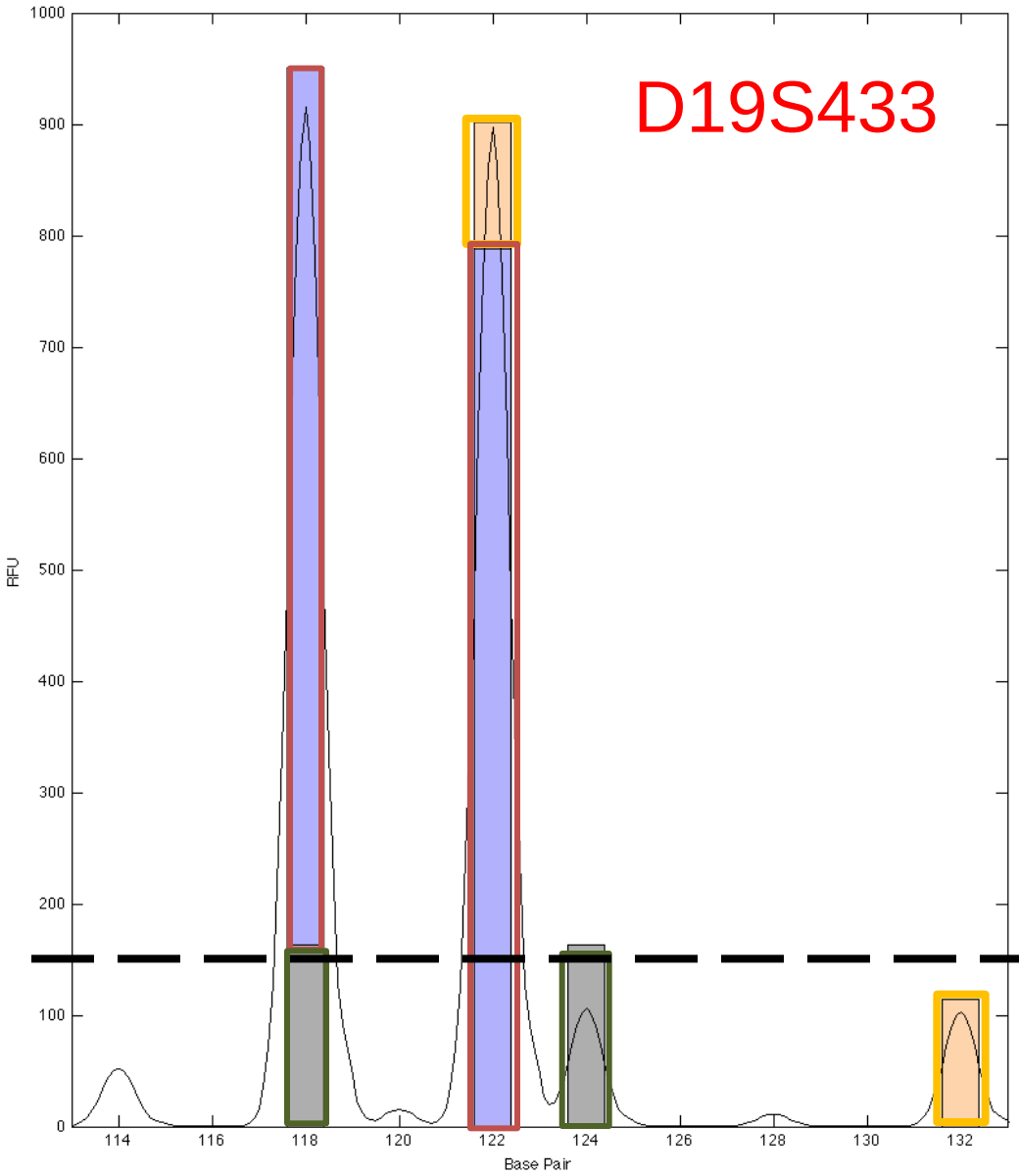
Main Points:

- 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)

Presentations/Publications:

- ISFG 2011 presentation
- ISHI 2011 mixture workshop

Review of One Replicate (of 50K)

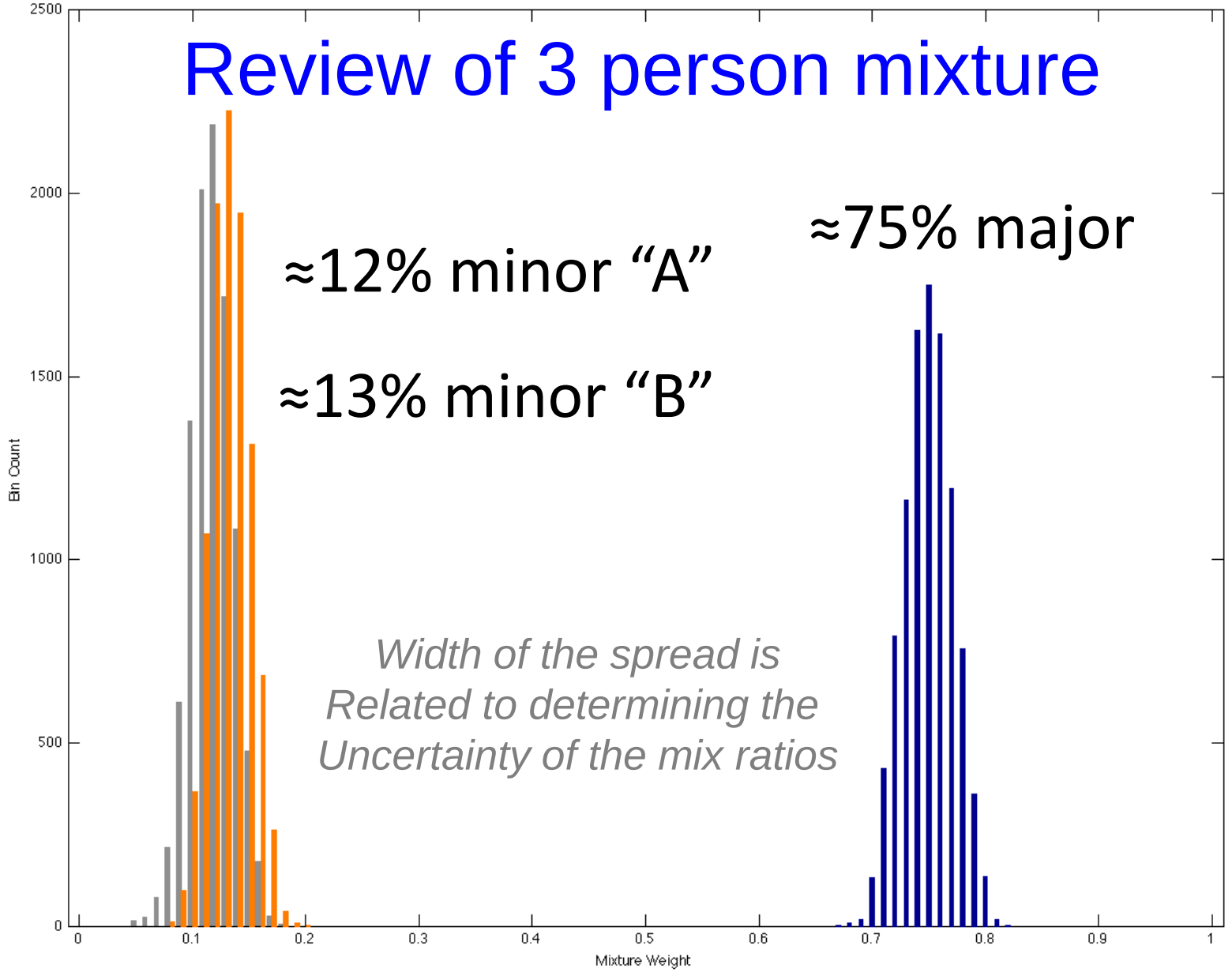


3P mixture,
2 Unknowns,

Conditioned on the Victim (major)

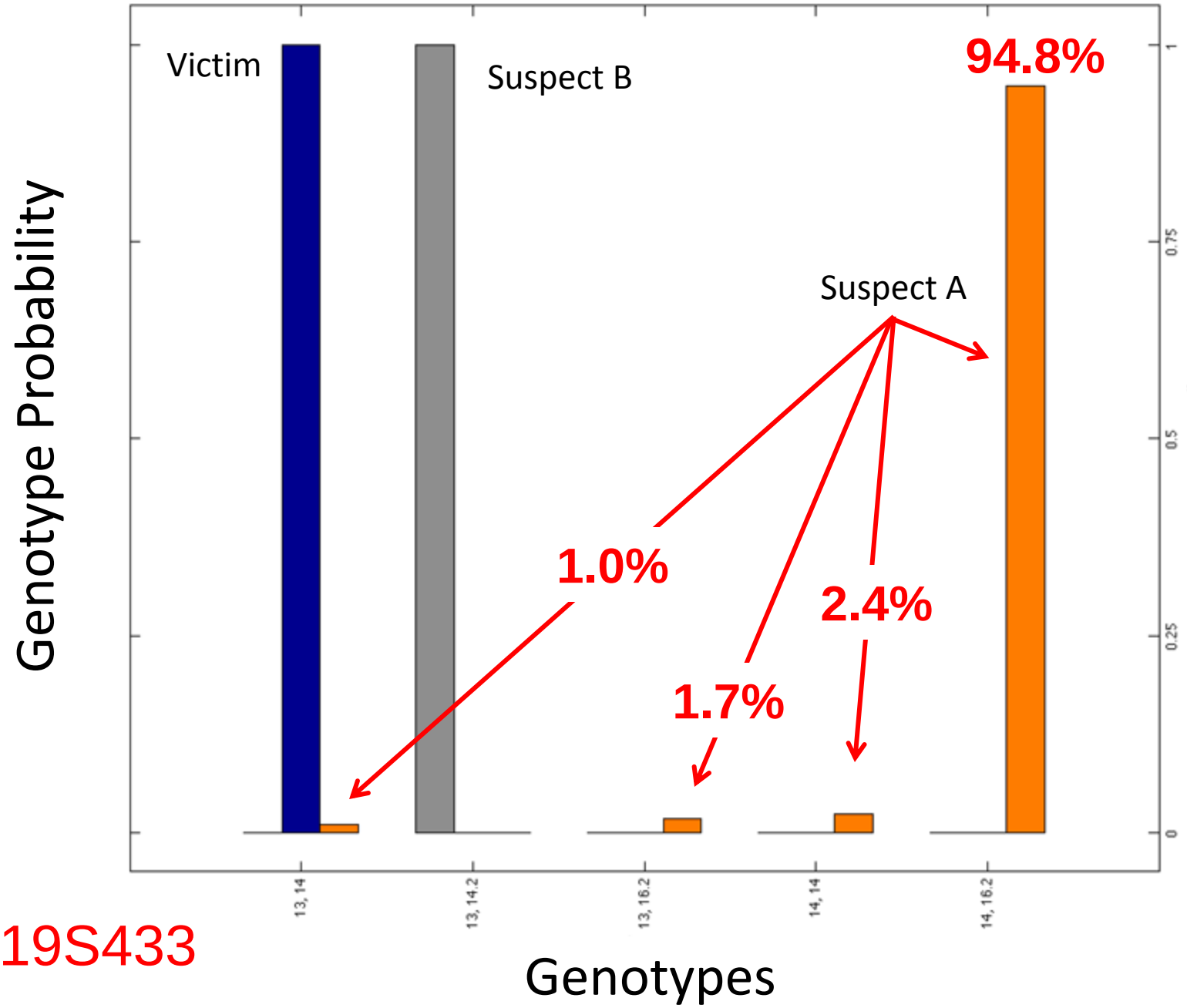
Good fit of the data to the model

Review of 3 person mixture



Bin Count

Mixture Weight


Victim
Suspect B
94.8%
Suspect A
1.0%
2.4%
1.7%
Genotype Probability
0
0.25
0.5
0.75
1
13, 14
13, 14.2
13, 16.2
14, 14
14, 16.2
Genotypes

D19S433

ABI 3500 Validation Summary

Erica Butts

- The 3500 has proven to be reliable, reproducible and robust
  - Out of 498 samples between Identifiler and Identifiler Plus only 5 required reinjection
- **Dye-specific analytical thresholds** resulted in less allelic and full locus dropout than applying one analytical threshold to all dyes
- Stochastic thresholds are linked to analytical thresholds
  - If the analytical threshold is adjusted, the stochastic threshold should be reevaluated along with expected peak height ratios
    - Requires consideration for overall interpretation workflow which we are still evaluating
- RFID tracking decreases flexibility in our research experience

DNA Community Moving to ABI 3500s

Advantages

- Smaller footprint and 110V power requirement
- Better polymer delivery and temperature control
  - Improved success rates?
- New capabilities
  - between instrument normalization
  - 6-dye detection (bigger kits with more loci)
- Simpler software

Disadvantages

- Up-front cost of new instruments
  - In the U.S., federal government (NIJ) will likely be expected to foot the bill
- Generates .hid files
  - Requires new analysis software
- Validation down-time
  - New RFU thresholds
- Higher per run cost with RFID tags & limited expiration
  - many labs cannot purchase reagents rapidly throughout the year
- Creating technicians not scientists
  - Plug and play approach leading to loss of understanding for process
  - Less flexible *(impacts research with it)*

Cost for the Forensic DNA Community to Switch from ABI 3100s to 3500s

1. **Instrument up-front cost**
   - Within the U.S. funding requests will likely come from federal grants

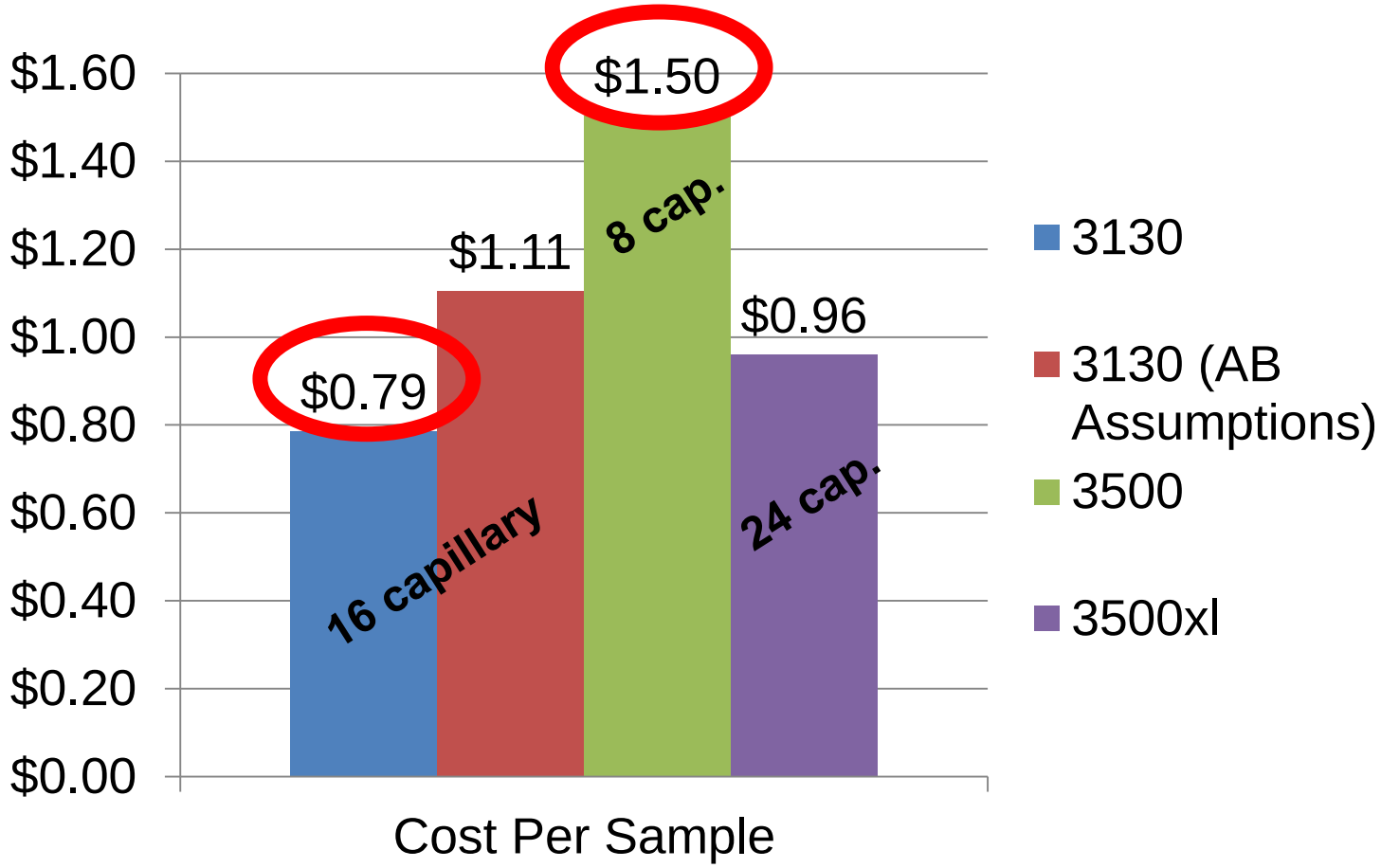
2. **New software purchase**
   - Will likely be requested from federal grant funds (NIJ)
   - new .hid file format will not work on current software (GMIDv3.2)
   - 3500 will not create .fsa files with 36cm arrays (HID applications)

3. **Validation time & expense**
   - Relative fluorescent scales are completely different…

4. **Operational cost**
   - ABI claims that the running costs are equivalent to 3130s…

NIST Calculated Cost per Sample for ABI 3130xl vs. 3500 and 3500xl Reagents

Running two plates per day (10 plates per week)



Consumable RFID Tracking Limits

	RFID Hard Stops	Usage Comments From a Research Laboratory Standpoint
Array	None	1. Very easy to change between HID and sequencing 2. Array from validation was stored at least twice and reinstalled on 3500 during validation
Buffer	Expiration Date 7 Days on Instrument # Injections	1. Can no longer use in-house buffer 2. Very easy to change on the instrument (snap-and-go)
Polymer	Expiration Date # Samples # Injections	1. Hard stop with the expiration date has caused us to discard unused polymer we would have otherwise kept on the instrument 2. ~50% of total polymer remains in the pouch after "consumption" 3. Expiration dates have changed purchasing strategy (smaller batches, based on ongoing project needs)

ABI 3500 Genetic Analyzer

Status Update on Open Letter to Applied Biosystems

Open Letter to Applied Biosystems on Concerns with ABI 3500

- **3/14/11 - emailed ~900 forensic DNA scientists** (SWGDAM, forens-dna, ENFSI, EDNAP) inviting them to sign onto a letter that will be sent to Applied Biosystems expressing concern with ABI 3500
- **Very positive response with 101 who agreed to sign the letter**
- Letter was sent March 31 to the president of ABI and scientists involved with the ABI 3500
- **Community will be notified of ABI's response**

A Sampling of Feedback I Received…

- **People did not just sign the letter but many have an opinion about the issues or concern about ABI customer support (I have received >100 emails – often with some very strong thoughts)**

- "I think that the AB3500 related issues most likely represent the beginning of a sea of problems, against which every independent lab must take arms. **It is not up to the manufacturer of a machine to decide the basic procedures of a lab - it is up to the lab**" (4/29/11)

- "I greatly appreciate your advocacy on behalf of our community. **Hopefully we will be heard**." (4/1/11)

Concerns Expressed in Open Letter

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

Hopefully a change will result…

A desire for greater communication with the community – the 3500 FAQ sheet is a good start but does not directly address all of the concerns raised

What was learned from ABI visit to NIST on May 11, 2011

- RFID over-ride is possible (their R&D lab has instrument that can use “expired” reagents)
- New software is required for 3500 .hid or .fsa files due to new file structure
- They do not have ANY data to support short shelf life of 3500 reagents
  - hard stops keep labs from having failures that lead to ABI having to replace arrays
- ABI 31xx instruments have a 4X signal reduction

Recent Decision to Reduce Stringency on Polymer Expiration

- New collected and collated data from Applied Biosystems on their ABI 3500 reagent expiration studies were shared with NIST on September 21, 2011

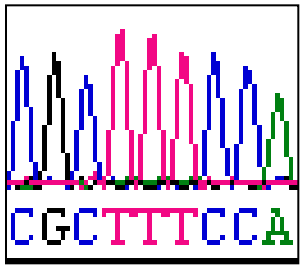
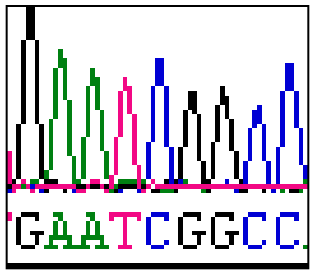
- At the Promega ISHI meeting in early October, ABI shared a poster stating that **polymer expiration dates will no longer be a hard stop** but only a warning with the future Windows 7 software upgrade

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
- 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

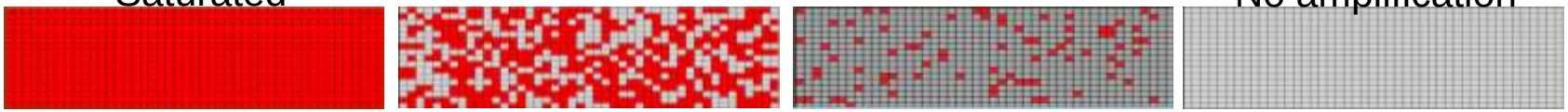
Comparison of Measurement Techniques

and ability to resolve two 9-base sequences

	CGCTTTCCA	GAATCGGCC
(a) Electrophoresis (fragment migration)	**≈9 nucleotides** (compared to size standard)	**≈9 nucleotides** (compared to size standard)
(b) Mass spectrometry (base composition)	**2566 Da** $A_1G_1C_4T_3$	**2640 Da** $A_2G_3C_3T_1$
(c) DNA sequencing (base position)		

NIST Digital PCR Instrument

Binary Detection



- Digital PCR performs hundreds of qPCR amplifications in very small volume wells with only 1-2 starting DNA target molecules per well
- Based on the number of wells that exceed a threshold (red squares), starting copy numbers can be determined mathematically

Support to the Community

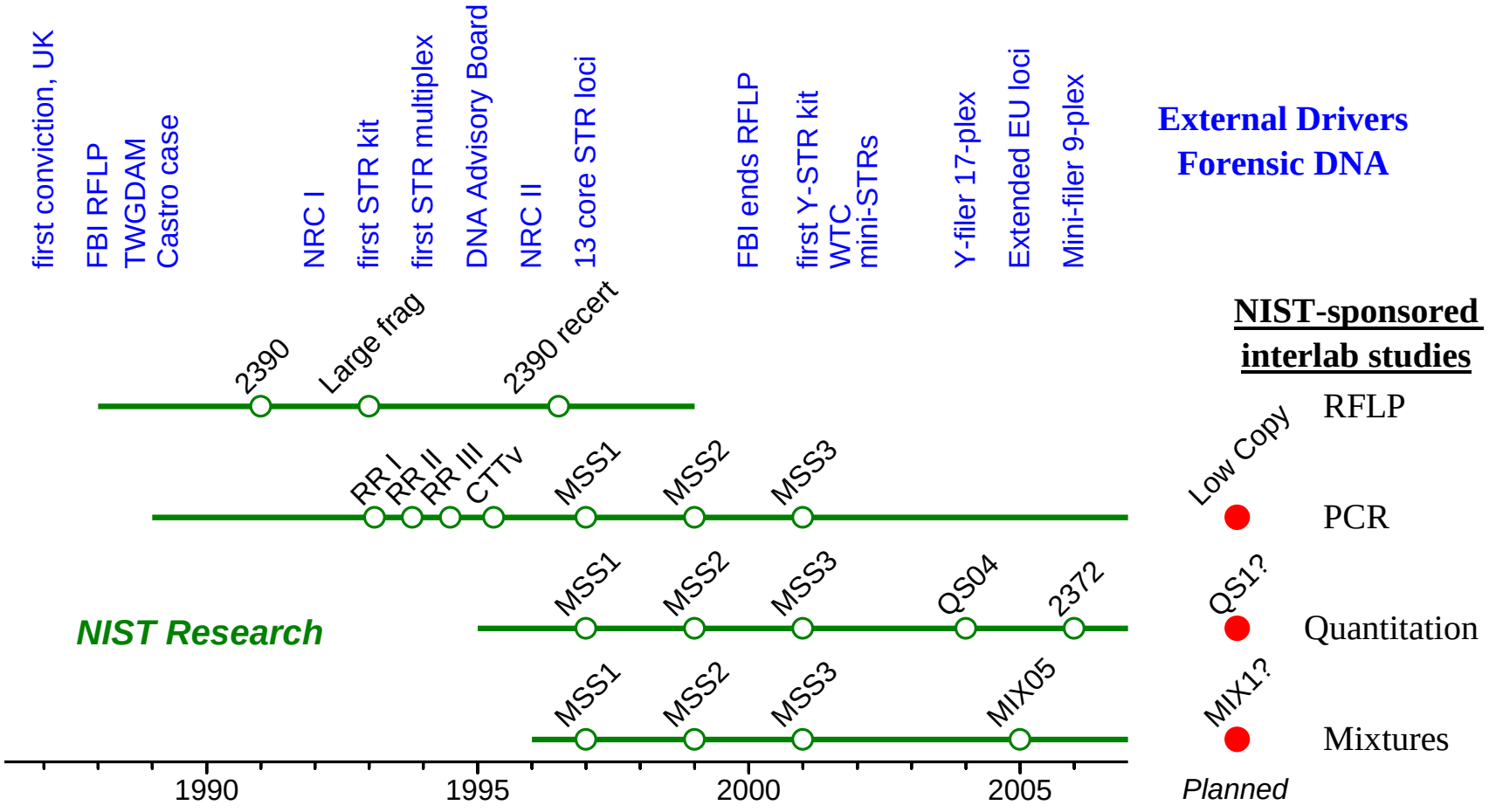
…Bringing traceability and technology to the scales of justice…

- Conduct interlaboratory studies
- Perform beta-testing of new human identity testing products
- **We collaborate with other NIJ grantees**
- We provide input to (or have aided):
  - Scientific Working Group on DNA Analysis Methods (**SWGDAM**)
  - Department of Defense Quality Assurance Oversight Committee for DNA Analysis
  - Virginia DFS Science Advisory Committee
  - American Prosecutor's Research Institute (**APRI**) DNA Forensics Program "Course-in-a-Box" for training lawyers
  - WTC Kinship and Data Analysis Panel (**KADAP**) and Hurricane Katrina efforts
  - NIJ Expert System Testbed (**NEST**) Project

NIST-Sponsored Interlab Studies


Margaret Kline

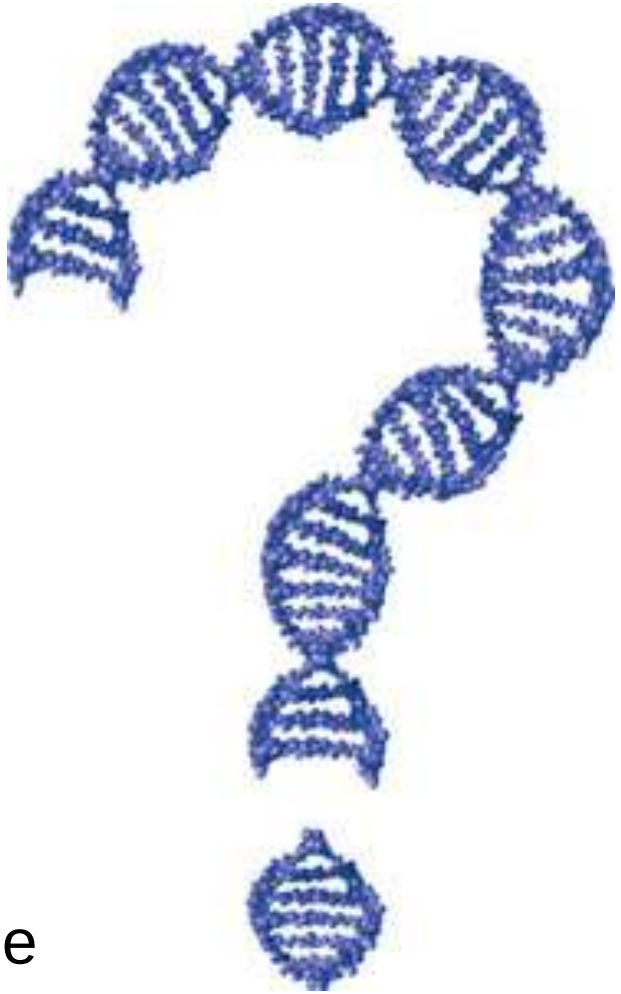

Dave Duewer



13 interlaboratory studies conducted over the past 20 years

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