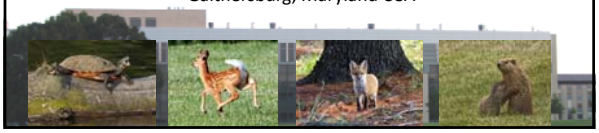
EDNAP and 36th ENFSI DNA WG Meeting
April 24-26, 2012 – Linköping, Sweden

NIST Update

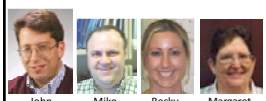
Peter M. Vallone

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 | Mike Coble | Becky Hill | Margaret Kline

STRBase, Workshops & Textbooks | Mixtures, mtDNA & Y | Concordance & LT-DNA | SRM work, variant alleles & Cell Line ID

Office Manager Patti Rohmiller

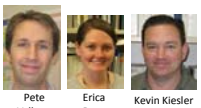
Guest Researcher

Manuel **Fonde**vila Álvarez

Data Analysis Support

Dave Duewer

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



Pete Vallone | Erica Butts | Kevin Kiesler

Rapid PCR, Direct PCR & Biometrics | ABI 3500 & DNA Extraction | PLEX-ID & NGS Exploration

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

Presentation Topics

- **Group Research Overview**
  - Release of Standard Reference Material SRM 2391c
  - STR kit concordance testing
  - Variant allele sequencing
  - STRBase website updates
  - Insertion/Deletion markers
  - PlexID (Mass spectrometry)
  - Rapid DNA (PCR, cyclers, instrumentation)
  - Mixture interpretation training & TrueAllele evaluation
  - Rapidly Mutating Y-STRs
  - SRM 2372 (recertification)
  - 3500 Genetic Analyzer (validation)
  - Advanced Topics in Forensic DNA Typing: Interpretation

NIST SRM 2391c

Margaret Kline | Becky Hill

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

- **http://www.nist.gov/srm**
- **SRM 2391c released Aug 2011**

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

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

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)
- **24 autosomal STRs, 17 Y-STRs, and Amel** with **certified values**
- 26 additional autosomal STRs with **reference values**
- 1 STR (Penta C) with **informational values**
- **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)

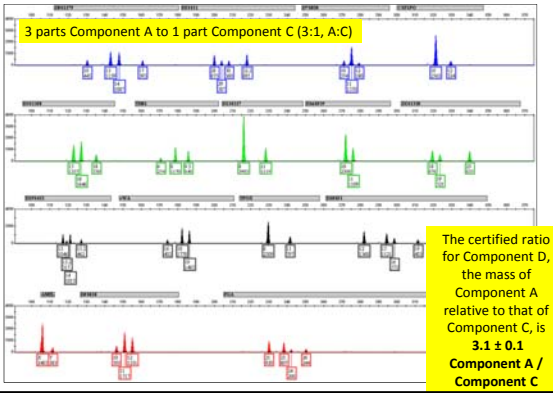
Description of Components in SRM 2391c

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.

http://www.cstl.nist.gov/biotech/strbase/pub_pres/Vallone_EDNAP_2012.pdf

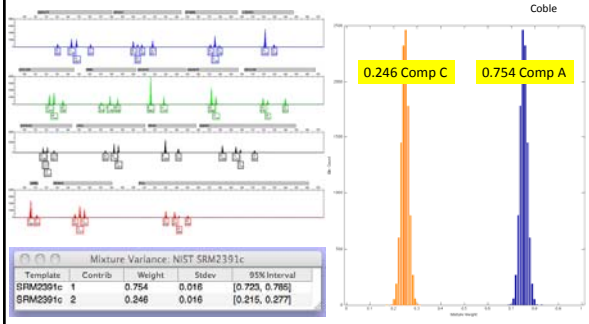
SRM 2391c Component D



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**

Estimate of Each Component Using *True Allele* Software


Mike Coble



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

If a primer binding site mutation exists

Presentations/Publications:

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

Kit Concordance Comparisons

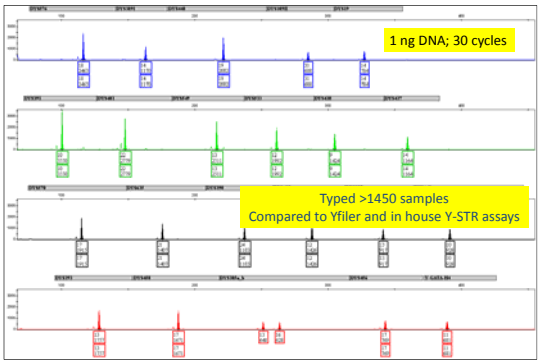

Becky Hill

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				
ID-MiniFiler	1308				
SGM-NGM	1436				
ID-NGM	1449				
ProPlus-NGM	1427				
SGM-ESI	1436				
ProPlus-ESX	1427				
ESI-ESX	1455				
ESI-ESSplex	1445				
ESX-ESSplex	1445				
ESI-NGMSElect	715				

128 kit-to-kit comparisons
1,104,031 allele comparisons
1224 differences observed
~99.9% concordance
(many corrected now)
Includes PP21
Becky just completed PPY23

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

PPY23 Example Profile (2800M)



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)


Forensic Science International: Genetics
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

Presentations/Publications:

- FSI Genetics article (Aug 2011) and numerous talks

http://www.cstl.nist.gov/biotech/strbase/pub_pres/Vallone_EDNAP_2012.pdf

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
- Mixture Interpretation
- Kinship Analysis
- miniSTRs (short amplicons)
- Null Alleles - discordance observed between STR kits
- STR Reference List - *now 3644 references*

Cataloged as of Mar 2012
632 variant alleles
310 tri-allelic patterns

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

NYC DNA Training Workshop


Mike Coble


John Butler


New York DNA Training Workshop
Topics and Techniques for Forensic DNA Analysis
Continuing Education Seminar
Wednesday, April 18, 2012
Agenda
9:00 – 9:20 a.m. Introductions
9:20 – 10:30 a.m. Data Interpretation & Statistical Analysis Overview
10:30 – 10:45 a.m. BREAK
10:45 a.m. – 12:15 p.m. Mixture Interpretation
12:15 – 1:30 p.m. LUNCH
1:30 – 3:00 p.m. STR Markers and CE Instrumentation
3:00 – 3:15 p.m. BREAK
3:15 – 4:45 p.m. Y-STRs, mtDNA, and the Romanov Case
4:45 – 5:00 p.m. Q&A

http://www.cstl.nist.gov/biotech/strbase/training/NewYork-April2012-Workshop.htm

Insertion/Deletion (InDel) Markers

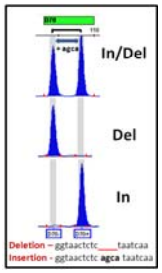

Manuel Fondevila Álvarez
Guest Researcher from Spain

Main Points:

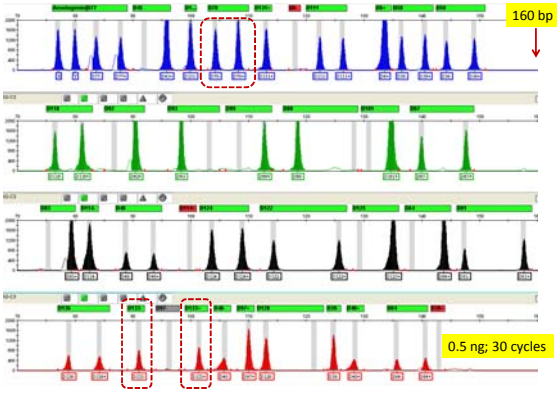
- InDels (insertion-deletion) or DIPs (deletion-insertion polymorphisms) are short length polymorphisms, consisting of the presence or absence of a short (typically 1-50 bp) sequence
- Like SNPs, InDels have low mutation rate (value to kinship analysis), small amplicon target sizes (value with degraded DNA), and can be highly multiplexed
- Can be separated and detected on CE instruments (similar to STR analysis)
- Typed a **commercial 30plex** (Qiagen DIPplex) and an **in house 38plex (IPATIMUP)** for over 700 U.S. population samples

Presentations/Publications:

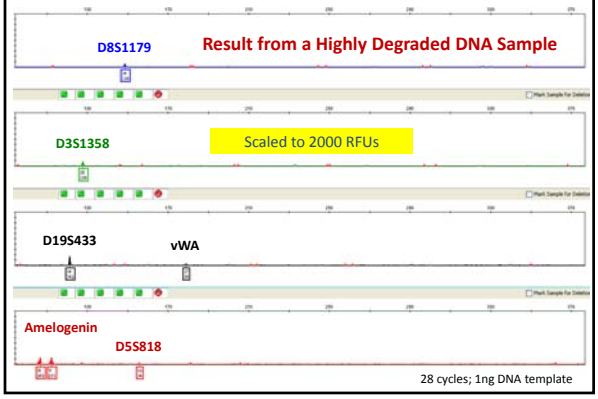
- FSI Genetics Suppl. Series 2011 article
- ISFG 2011 poster and ISHI 2011 presentation
- **Manuscript submitted to IJLM (allele frequencies, mapping loci, sequencing of null or imbalanced alleles)**



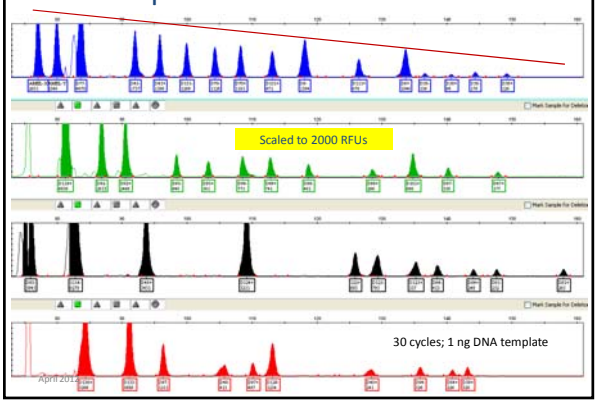
DIPplex Insertion/Deletion Assay (Qiagen kit)



Identifiler – 7 alleles detected



DIPplex – 49 alleles detected



http://www.cstl.nist.gov/biotech/strbase/pub_pres/Vallone_EDNAP_2012.pdf

Additional and Current work


Manuel Fondevila Álvarez

- Completed typing of a revised ancestry 34plex on NIST U.S. population samples
  - Submitted for publication
- Collaboration with The George Washington University on typing population samples
- Will be finalizing work at NIST in July 2012

We are open to future (funded) guest researcher collaborations

Performance Assessment of PLEX-ID Mass Spectrometer


Kevin Kiesler

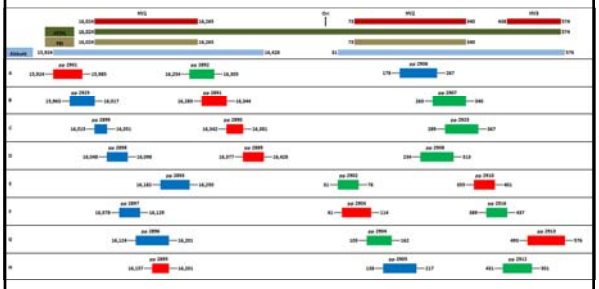

Pete Vallone

Abbott Ibis Biosciences PLEX-ID System

- FBI collaboration
- Assessment of ESI-TOF mass spectrometer for mtDNA analysis
- Base composition of the control region determined from 8 triplex PCRs
- Started operating the PLEX-ID platform mid-October 2011
- Scheduled to complete experiments in April 2012
- Concordance with Sanger sequencing, sensitivity, contamination, mixtures, robustness of assay/platform

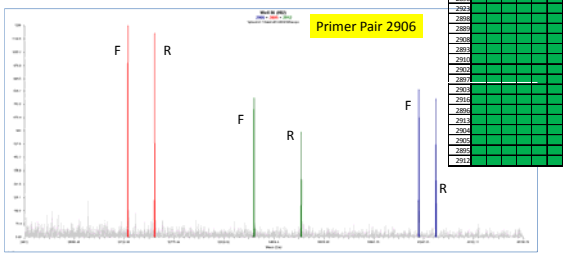
PLEX-ID mtDNA 2.0 Assay

Coverage (24 amplicons) 8 triplex amplifications
Positions 15,924 through 16,428 and 31 through 576



Sensitivity

Full profile generated at 2.5 pg per reaction (nuclear quantitation)



PLEX-ID Experiments

Experiment	Number of Plates	Number of Unique Samples
Mixtures	21	3
Concordance	61	458
Sensitivity / Limit of Detection	17	3
Contamination	15	1
Total	**114**	**10,994 wells examined**

- Mixtures can be detected with minor component present at 20-25%
- Concordance with Sanger sequencing (99.2%) (n=248)
- Full profiles generated at ≈ 2.5 to 5.0 pg/well

Rapid PCR and Rapid DNA Testing


Pete Vallone


Erica Butts

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 (R-DNA) typing devices**
  - NIST will be examining rapid DNA instruments in collaboration with the FBI

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

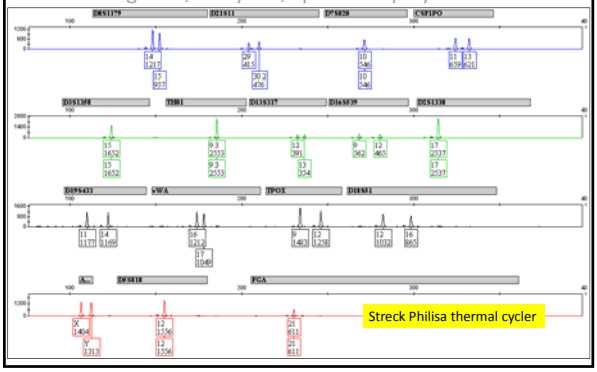
http://www.cstl.nist.gov/biotech/strbase/pub_pres/Vallone_EDNAP_2012.pdf

Rapid Thermal Cyclers

Streck Philisa	Thermo Scientific Piko	Ahram Palm PCR
		
Heating: 15°C/s Cooling: 12°C/s	Heating: 5°C/s Cooling: 4.5°C/s	Heating: ?°C/s Cooling: ?°C/s

Identifiler Profile - 14 min PCR Cycling

1 ng DNA; 28 cycles; SpeedStar polymerase



Rapid-DNA Instrumentation


Pete Vallone


Erica Butts

- Multiple commercial efforts for a *fully integrated* rapid DNA typing instrument
  - Buccal swab (input) – STR profile (output)
- In collaboration with the FBI, NIST is developing a testing plan to assist in the assessment as the platforms become available


→ **R-DNA instrument** DNA extraction, PCR, separation/detection, and interpretation → 1-2 h → CMF/STR Profile

NIST Testing of R-DNA Platforms

(general prioritization)

- **'Burn in' samples for baseline performance**
- **NIST testing samples** (anonymous buccal swabs and cultured cells)

- Different operator(s)
- Negatives for contamination
- Store chips at 37 & 65°C
- Add extracted DNA to Biochip
- Swabbing cells surfaces
- Test DNA mixtures
- Operating instrument in varying locations
- Other?

- Concordance (correct genotype)
- Reproducibility
- Expert v manual interpretation
- CE and PCR metrics
  - Base pair sizing precision
  - Dye-specific S/N
  - Stutter ratios
  - Heterozygote peak balance
  - Adenylation peaks
  - Artifacts observed

TrueAllele Mixture Software Evaluation


Mike Coble

Main Points:

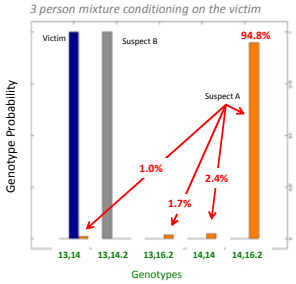
- Exploring the capabilities and limitations of a probabilistic genotyping approach
- Studying TrueAllele software with a number of different types of mixtures (including low-level and 3-4 person mixtures)
- Work being performed at NIST independently of Cybergenetics

Presentations/Publications:

- ISFG 2011 presentation
- ISHI 2011 mixture workshop

D19S433 result from one replicate of 50,000 simulations

3 person mixture conditioning on the victim



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

Promega ISHI 2012 Mixture Workshop

- **John Butler**, Ph.D., NIST, Gaithersburg, MD
- **Michael Coble**, Ph.D., NIST, Gaithersburg, MD
- **Robin Cotton**, Ph.D., Boston University, Boston, MA
- **Catherine Grgicak**, Ph.D., Boston University, Boston, MA
- **Charlotte J. Word**, Ph.D., Gaithersburg, MD

This workshop is for analysts, technical reviewers and technical leaders performing and interpreting validation studies and/or interpreting and reviewing STR data, particularly more difficult mixtures. Various DNA profiles will be analyzed and interpreted using selected analytical thresholds and stochastic thresholds to demonstrate the impact of those values on the profiles amplified with low-template DNA vs. higher amounts of DNA. Different statistical approaches and conclusions suitable for the profiles will be presented.

Rapid Mutating Y-STRs

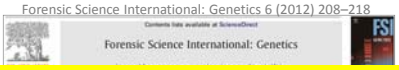

Becky Hill


Mike Coble

Locus
(average mutation rate)

DYS449 (1.2%)
DYS518 (1.8%)
DYS547 (2.4%)

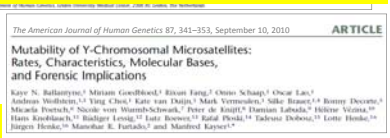
Forensic Science International: Genetics 6 (2012) 208–218


Forensic Science International: Genetics

- Running 3 RM YSTR multiplexes
- NIST population samples (665) + Father/Son pairs (802) - completed
- Part of Manfred Kayser's multicenter study
- Analysis….will be completed by the May 15 deadline!

DYF387S1 (1.6%)
DYF399S1 (7.7%)
DYF403S1 a/b (3.1/1.2%)
DYF404S1 (1.3%)
DYS526 a/b (1.3%)

DYS458 (0.64%) is highest in Yfiler loci where average is ~0.2%


The American Journal of Human Genetics 87, 341–353, September 10, 2010 ARTICLE

Mutability of Y-Chromosomal Microsatellites: Rates, Characteristics, Molecular Bases, and Forensic Implications

Status of SRM 2372


Margaret Kline

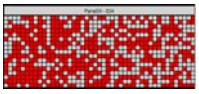
- NIST SRM 2372 Human DNA Quantitation Standard – **currently not for sale** (as of March 2012)
- Certified based on absorbance values
  - 1 OD ≈ 50 ng/µL
  - Used to calibrate qPCR standards
  - The absorbance of the components is drifting (increasing by 10-15%)
  - Possibly due to double strand to single strand dissociation? (not evaporation)

Should not affect qPCR use of materials

Status of SRM 2372


Margaret Kline

- Current status
- qPCR evaluations are currently underway
- Recertify based on:
- Updated absorbance measurements
- Digital PCR – independent copy number determination
  - Initial experiments are underway on the Fluidigm digital PCR system



Goal is to have the SRM back by fall of 2012

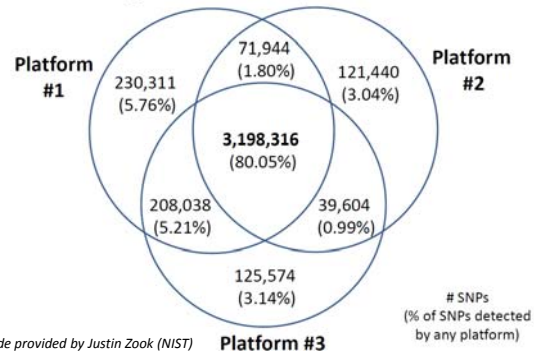
Next Generation Sequencing


Kevin Kiesler


Pete Vallone

- Interagency Workshop on the use of Next-Generation DNA Sequencing for Human Identification and Characterization (Jan 31 2012)
- Discussion of forensic applications of NGS (NIST, DoD, FBI, DHS) – materials can be found at:
  - http://www.nist.gov/mml/biochemical/genetics/ngs_hid_workshop.cfm
- We are in the process of looking at platforms to characterize forensic markers (mitochondrial, STRs, SNPs)
- Evaluate accuracy, reproducibility, identify initial requirements for a NGS forensic reference material

Whole genome sequencing technologies disagree about 100,000's of SNPs



Slide provided by Justin Zook (NIST)

ABI 3500 Validation Studies


Erica Butts

Main Points:

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

Presentations/Publications:

- MAAFS talk (May 2011)
- ABI road show talks (July & Aug 2011)
- **ISFG presentation (Sept 2011)**
- *Forensic News* (Spring 2012)


HID in Action
3500 Genetic Analyzer: Validation Studies

http://marketing.appliedbiosystems.com/mk/get/FORENSICNEWS_HIDINACTION#article5

http://www.cstl.nist.gov/biotech/strbase/pub_pres/Vallone_EDNAP_2012.pdf

ABI 3500 Open Letter Update

Concerns Expressed in 3/31/11 Open Letter

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

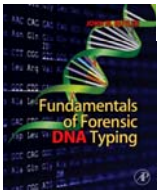
At the Promega ISHI meeting (Oct 2011), ABI described data for studies around reagent expiration through a poster at their booth.
Sailus, Wheaton, Fisher, Calandro. "Understanding the Consumables on the 3500 Genetic Analyzers in the context of a Human Identification (HID) Laboratory"

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

Forensic DNA Typing Textbook

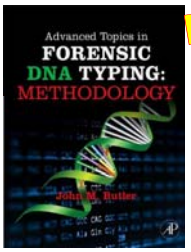
3rd Edition is Three Volumes

For beginning students, general public, & lawyers



Sept 2009

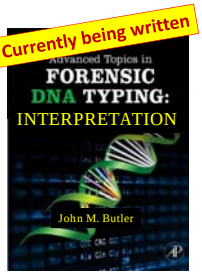
~500 pages



August 2011

~700 pages

Currently being written



Fall 2012

~500 pages

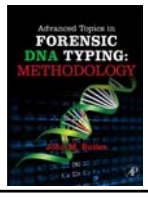
Steps Involved in Process of Forensic DNA Typing

1) Data Interpretation
2) Statistical Interpretation

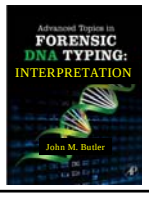
Gathering the Data | **Understanding the Data**

Collection/Storage/Characterization → Extraction/Quantitation → Amplification/Marker Sets → Separation/Detection → **Interpretation** → **Report**

Advanced Topics: Methodology



Advanced Topics: Interpretation



Advanced Topics in Forensic DNA Typing: INTERPRETATION

Chapter	Topic (current planned chapters)
	Introduction
1	Data interpretation overview
2	Thresholds
3	STR alleles & artifacts
4	STR genotypes & dropout
5	STR profiles
6	**Mixture interpretation**
7	Low-level DNA and complex mixtures
8	CE troubleshooting
9	Statistical interpretation overview
10	STR population data analysis
11	Profile frequency estimates
12	**Mixture statistics**
13	Coping with potential missing alleles
14	**Kinship and parentage analysis**
15	Lineage marker statistics
16	Drawing conclusions & **report writing**
	Glossary
App 1	U.S. Population Data (24 loci with N=938)
App 2	Revised Forensic DNA QAS (Sept 2011)
App 3	DAB Recommendations on Stats (Feb 2000)
App 4	NRC II Recommendations (1996)
App 5	SWGDAM STR Interp Guidelines (Jan 2010)

Features in New Book
(planned for Spring 2013 release)

- Explanations of SWGDAM interpretation guidelines
- Interviews on report writing from multiple perspectives
- Mixture interpretation
- Kinship analysis
- CE troubleshooting
- Standard U.S. pop data
- Numerous D.N.A. Boxes (**D**ata, **N**otes, & **A**pplications)
  - Worked examples to show relevance of equations
  - "Better know a statistician"

Thank you for your attention!

Questions?
peter.vallone@nist.gov
1+301-975-4872

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