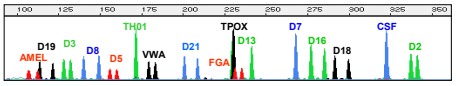


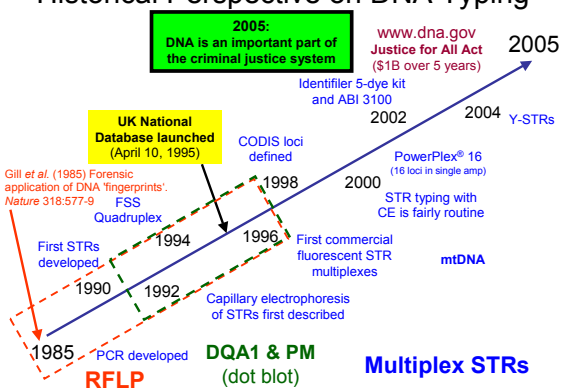
Setting Standards and Developing Technology to Aid the Human Identity Testing Community

John M. Butler
Peter M. Vallone, Michael D. Coble,
Janette W. Redman, Amy E. Decker, Carolyn R. Hill,
David L. Duewer, Margaret C. Kline

Human Identity Project Team, U.S. National Institute of Standards and Technology

ISFG Keynote Address – September 14, 2005

Historical Perspective on DNA Typing



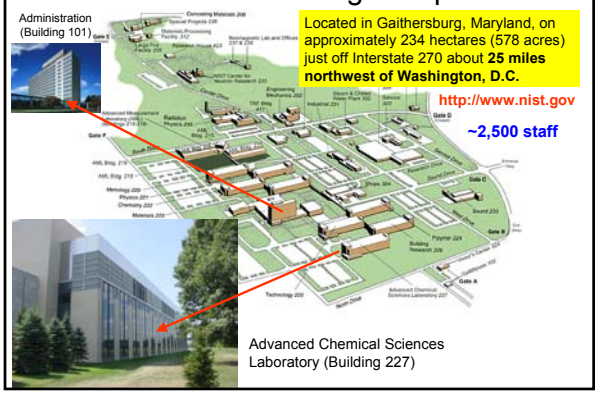
Presentation Overview

- Introduction to NIST and Project Team
- Standards
- Technology
- Training Materials

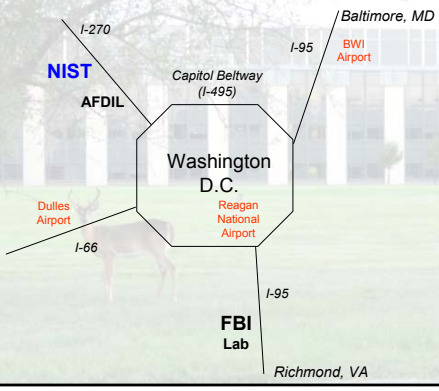
NIST Mission and History

- National Institute of Standards and Technology (NIST) was created in 1901 as the National Bureau of Standards (NBS). The name was changed to NIST in 1988.
- NIST is **part of the U.S. Department of Commerce** with a mission to develop and promote measurement, standards, and technology to enhance productivity, facilitate trade, and improve the quality of life.
- NIST supplies over 1,300 Standard Reference Materials (SRMs) for industry, academia, and government use in such areas as environmental analysis, health measurements, and industrial materials production and analysis.

NIST Gaithersburg Campus



General Location of NIST

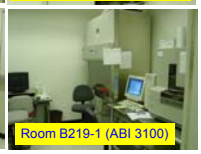


NIST Human Identity Project Team Laboratory Space
(within Building 227)


Room B257 (PCR Setup)


Room B223 (Setup/Analysis)

Room B261 (Post-PCR)


Room B219-1 (ABI 3100)

Room B219 (DNA extraction)

Room A230 (ABI 3130xl)

Equipment List

GeneAmp 9700
GeneAmp 9600
ABI 310
ABI 3100
ABI 3130xl (on order)
FMBIO III+
ABI 7500 real-time PCR

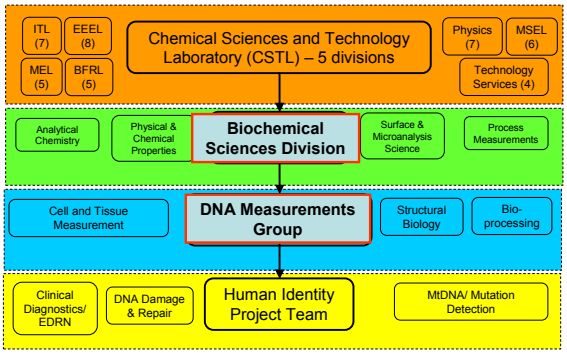
Agilent Bioanalyzer 2100
Varian UV spec
Varian HPLC
Bruker TOF-MS
MWG RoboAmp 4200
-80 ºC freezers

NIST and Forensic DNA

- Dennis Reeder, Margaret Kline, and Kristy Richie started DNA work at NIST in late 1980s
- The FBI's DNA Advisory Board required one member be from NIST, which was Dennis Reeder
- NIST has been an invited guest to TWGDAM/SWGDAM since 1988
- A number of Standard Reference Materials have been developed since 1992:
  SRM 2390 (RFLP), SRM 2391 (PCR/STRs), SRM 2392 (mtDNA),
  SRM 2395 (Y-STRs), SRM 2372 (DNA quant)
- STRBase website developed in 1996-1997 while John Butler was a postdoc
- Human Identity Project Team formed in 2000 with John Butler as project leader (when Dennis Reeder left NIST for Applied Biosystems)
- Numerous research areas now funded by the National Institute of Justice

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

Overview of NIST Organization



http://www.cstl.nist.gov/biotech/DNATechologies/Human_Identity.htm

NIST Human Identity Project Team


John Butler
(Project Leader)


Margaret Kline


Pete Vallone


Mike Coble


Dave Duewer
Anal. Chem. Division


Jan Redman


Amy Decker


Becky Hill


Chris DeAngelis
computer programmer

Funding: **Interagency Agreement 2003-IJ-R-029 between National Institute of Justice (NIJ) and NIST Office of Law Enforcement Standards (OLES)**

Team Productivity

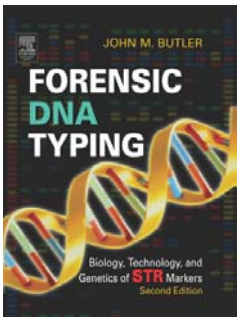
Since last ISFG meeting (Sept 2003):

- 43 publications
- 57 presentations
- Textbook: *Forensic DNA Typing*, 2nd Edition
- Widely-used website: hundreds of pages of new information added to STRBase

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

Forensic DNA Typing, 2nd Edition:
Biology, Technology, and Genetics of STR Markers

John Butler (not NIST)



double the size of the first edition
10 new chapters, >500 new references

Chapter 1 Overview & History of DNA Typing
Chapter 2 DNA Biology Review
Chapter 3 Sample Collection, Extraction, Quantitation
Chapter 4 PCR Amplification
Chapter 5 Common STRs and Commercial Kits
Chapter 6 Biology of STRs
Chapter 7 Forensic Issues
Chapter 8 **Single Nucleotide Polymorphisms**
Chapter 9 **Y-Chromosome DNA Tests**
Chapter 10 **Mitochondrial DNA Analysis**
Chapter 11 **Non-Human DNA and Microbial Forensics**
Chapter 12 DNA Separation Methods
Chapter 13 DNA Detection Methods
Chapter 14 Instrumentation for STR Typing: ABI 310, **3100**, FMBIO
Chapter 15 STR Genotyping Issues
Chapter 16 Lab Validation
Chapter 17 New Technologies, Automation, and **Expert Systems**
Chapter 18 CODIS and DNA Databases
Chapter 19 **Basic Genetic Principles and Statistics**
Chapter 20 **STR Database Analyses**
Chapter 21 **Profile Frequency Estimates**
Chapter 22 **Statistical Analysis of Mixtures and Degraded DNA**
Chapter 23 **Kinship and Paternity Testing**
Chapter 24 **Mass Disaster DNA Victim Identification**
Appendix I Reported STR Alleles
Appendix II **U.S. Population Data-STR Allele Frequencies**
Appendix III Suppliers of DNA Analysis Equipment
Appendix IV DAB QA Standards
Appendix V **DAB Recommendations on Statistics**
Appendix VI **Application of NRC II to STR Typing**
Appendix VII **Example DNA Cases**

National Institute of Justice
The Research, Development, and Evaluation Agency of the U.S. Department of Justice

Current Areas of NIST Effort with Forensic DNA

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

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

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

FBI's DNA Advisory Board
Quality Assurance Standards for Forensic DNA Testing Laboratories

(October 1, 1998)

STANDARD 9.5

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

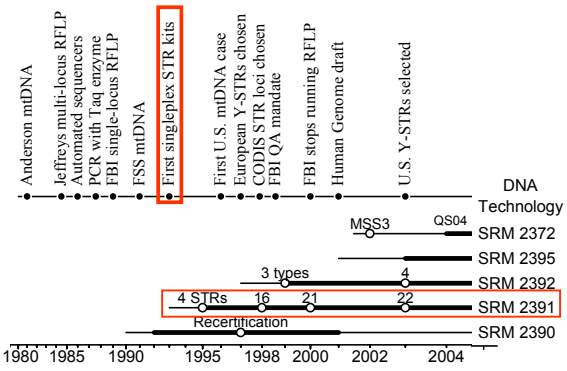
Margaret Kline Jan Redman

Standard Reference Materials

- Relevant Forensic DNA SRMs
  - SRM 2391b (DNA profiling – STRs, D1S80, DQA1/PM)
  - SRM 2392-I (mtDNA)
  - SRM 2395 (Y-chromosome)
  - SRM 2372 (Human DNA quantitation); in development
- Provides national/international traceability and compatibility (aids in ISO 17025 compliance)

http://www.nist.gov/srm

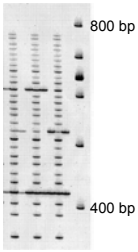
Forensic DNA Timeline with NIST SRM Development



1995: SRM 2391 PCR-based DNA Profiling Standard

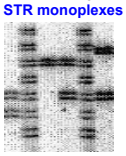
D1S80 Locus
Variable Number Tandem Repeat (VNTR)

DQ-alpha & PolyMarker reverse dot blot hybridization

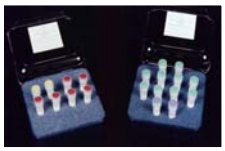


Samples originally selected to possess all DQA1 types

Silver stained gels for **STR monoplexes**

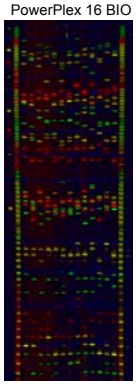


6 – Pre-amplified D1S80 samples
1 – D1S80 allelic ladder
1 – size standard
10 – Genome DNA Extracts
2 – Cell line cells spotted on 903 paper for extraction



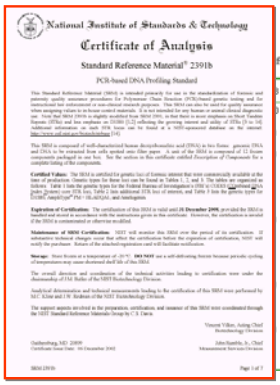
2002: SRM 2391b released

PowerPlex 16 BIO



- 22 STR Loci
- D1S80 and DQA1/PM (still certified)
- Many labs using 16plex STR kits

NIST SRM 2391b


National Institute of Standards & Technology

Certificate of Analysis

Standard Reference Material® 2391b

PCR-based DNA Profiling Standard

All current commercial kit STR loci are certified

fied Values for Additional STR Loci

	FES/FPS	LPL	Penta D	Penta E	D2S1338	D19S433
	12,12	10,11	10,15	7,12	17,23	13,16.2
	10,11	10,11	9,11	7,12	17,26	14,16
	11,12	11,12	11,12	13,14	20,24	12,14
	10,13	10,12	8,9	5,12	17,23	11,13
	11,13	10,12	10,13	7,13	17,19	12.2,14
	11,11	10,12	9,12	12,14	25,25	12,14
	11,11*	11,12	3.2,11	12,16	17,22	13,15.2
	10,1					.15
	10,1					.5
	11,1					.14
	10,1					.5
	11,11	10,12	8,12	11,11	23,23	13,14

22 autosomal STRs characterized across 12 DNA samples

STRBase:
A Standard Information Resource


John Butler Jan Redman

Primary updates performed monthly

- Summary of variant alleles and tri-allelic patterns
- List of STR references (Reference Manager database)
- NIST publications and presentations

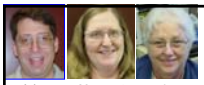
- **New content is being added regularly** to aid training and to support forensic DNA laboratories

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

Content of STRBase Website
http://www.cstl.nist.gov/biotech/strbase

- …/str_fact.htm STR Fact Sheets on Core Loci
- …/multiplx.htm Multiplex STR Kit Information
- …/y_strs.htm Y-Chromosome Information
- …/var_tab.htm Variant Alleles Reported
- …/mutation.htm Mutation Rates for Common STRs
- …/str_ref.htm Reference List with ~2,500 Papers
- …/training.htm Downloadable PowerPoints for Training
- …/validation.htm **Validation Information**
- …/miniSTR.htm miniSTR Information
- …/address.htm Addresses for Scientists
- …/NISTpub.htm Publications & Presentations from NIST

Validation Information


John Butler Margaret Kline Jan Redman

- Survey initiated at June 2004 NIJ meeting and conducted last summer resulted in 53 responses
- Talk at Promega meeting Oct 2004
- Validation summary sheets
- Validation website on STRBase
- **Workshop conducted August 2005** at NFSTC (DVD to be released as part of *President's DNA Initiative* training)

- We invite submission of your internal validation studies for inclusion in the NIST validation website

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

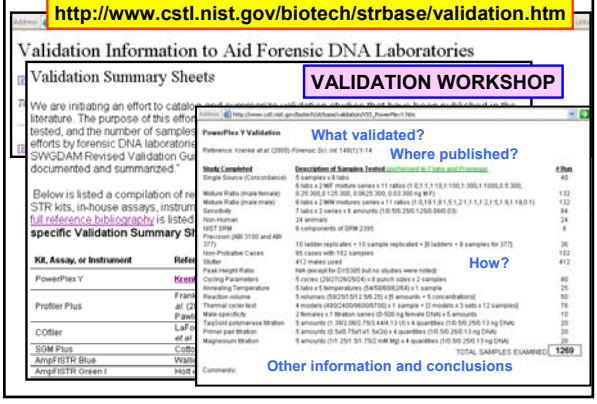
Pathway to Improved DNA Validation

- Collection of Current Philosophy on Validation
  - Community survey
  - Interviews
  - Literature summary
- Training
  - Auditors must be consistent in treatment of labs
- Providing Tools to Enable Improved Validation
  - Sample set(s)
  - Workbook – provide specific examples
  - Standard report form – documentation standardization
- Collection of Validation Data from Labs
  - NIJ-funded labs to submit data to STRBase validation website

VALIDATION WORKSHOP

New Validation Homepage on STRBase

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


Validation Information to Aid Forensic DNA Laboratories

Validation Summary Sheets

VALIDATION WORKSHOP

What validated?

Where published?

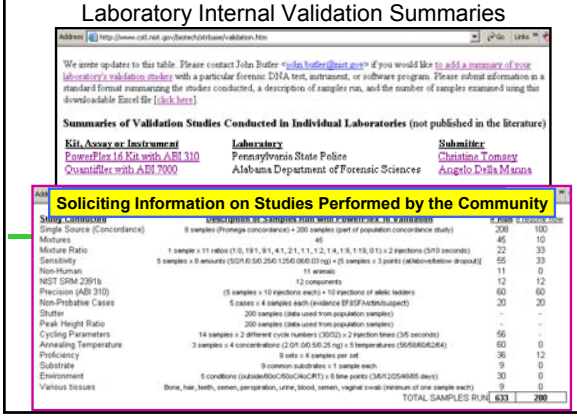
How?

Other information and conclusions

Validation Summary Sheet for PowerPlex Y

Study Completed (17 studies done)	Description of Samples Tested (performed in 7 labs and Promega)	# Run
Single Source (Concordance)	5 samples x 8 labs	40
Mixture Ratio (male:female)	6 labs x 2 M/F mixture series x 11 ratios (1:0,1:1,1:10,1:100,1:300,1:1000,0.5:300, 0.25:300,0.125:300, 0.0625:300, 0.03:300 ng M:F)	132
Mixture Ratio (male:male)	6 labs x 2 M/M mixtures series x 11 ratios (1:0, 19:1, 9:1, 5:1, 2:1, 1:1, 1:2, 1:5, 1:9, 1:19, 0:1)	132
Sensitivity	7 labs x 2 series x 6 amounts (1/0.5/0.25/0.125/0.06/0.03)	84
Non-Human	24 animals	24
NIST SRM	6 components of SRM 2395	6
Precision (ABI 3100 and ABI 377)	10 ladder replicates + 10 sample replicated + [8 ladders + 8 samples for 377]	36
Non-Probative Cases	65 cases with 102 samples	102
Stutter	412 males used	412
Peak Height Ratio	N/A (except for DYS385 but no studies were noted)	
Cycling Parameters	5 cycles (28/27/26/25/24) x 8 punch sizes x 2 samples	80
Annealing Temperature	5 labs x 5 temperatures (54/58/60/62/64) x 1 sample	25
Reaction volume	5 volumes (50/25/15/12.5/6.25) x [5 amounts + 5 concentrations]	50
Thermal cycler test	4 models (480/2400/9600/9700) x 1 sample + [3 models x 3 sets x 12 samples]	76
Male-specificity	2 females x 1 titration series (0-500 ng female DNA) x 5 amounts each	10
TaqGold polymerase titration	5 amounts (1.38/2.06/2.75/3.44/4.13 U) x 4 quantities (1/0.5/0.25/0.13 ng DNA)	20
Primer pair titration	5 amounts (0.5x/0.75x/1x/1.5x/2x) x 4 quantities (1/0.5/0.25/0.13 ng DNA)	20
Magnesium titration	5 amounts (1/1.25/1.5/1.75/2 mM Mg) x 4 quantities (1/0.5/0.25/0.13 ng DNA)	20
Krenke *et al.* (2005) *Forensic Sci. Int.* 148:1-14	TOTAL SAMPLES EXAMINED	1269

Laboratory Internal Validation Summaries



Margaret Kline | Dave Duewer | Jan Redman | John Butler

Interlaboratory Studies

- DNA Quantitation Study (QS04)
  - 8 DNA samples supplied
  - 84 laboratories signed up (80 labs returned results)
  - 287 data sets using 19 different methods
  - 60 data sets with real-time qPCR (37 Quantifiler data sets)
  - Publication in May 2005: *J. Forensic Sci.* 50(3): 571-578
- Mixture Interpretation Study (MIX05)
  - 105 labs signed up (**71** labs returned data)
  - Interpretation requested of provided e-grams for 4 mock sexual assault cases
  - Data analysis is still on-going…

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

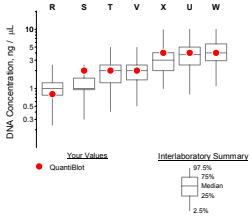
NIST Initiated Interlaboratory Studies

Studies involving STRs	# Labs	Publications
Evaluation of CSF1PO, TPOX, and TH01	34	Kline MC, Duewer DL, Newall P, Redman JW, Reeder DJ, Richard M. (1997) Interlaboratory evaluation of STR triplex CTT. *J. Forensic Sci.* 42: 897-906
Mixed Stain Studies #1 and #2 (Apr–Nov 1997 and Jan–May 1999)	45	Duewer DL, Kline MC, Redman JW, Newall PJ, Reeder DJ. (2001) NIST Mixed Stain Studies #1 and #2: interlaboratory comparison of DNA quantification practice and short tandem repeat multiplex performance with multiple-source samples. *J. Forensic Sci.* 46: 1199-1210
MSS3 Mixed Stain Study #3 (Oct 2000-May 2001)	74	Kline, M.C., Duewer, D.L., Redman, J.W., Butler, J.M. (2003) NIST mixed stain study 3: DNA quantitation accuracy and its influence on short tandem repeat multiplex signal intensity. *Anal. Chem.* 75: 2463-2469. Duewer, D.L., Kline, M.C., Redman, J.W., Butler, J.M. (2004) NIST Mixed Stain Study #3: signal intensity balance in commercial short tandem repeat multiplexes, *Anal. Chem.* 76: 6928-6934.
DNA Quantitation Study (Jan-Mar 2004) QS04	80	Kline, M.C., Duewer, D.L., Redman, J.W., Butler, J.M. (2005) Results from the NIST 2004 DNA Quantitation Study, *J. Forensic Sci.* 50(3):571-578
MIX05 Mixture Interpretation Study (Jan-Mar 2005)	**71**	***Data analysis currently on-going ...*** Will be presented at NIJ Grantees (June 2005), ISFG (Sept 2005), Promega (Sept 2005)

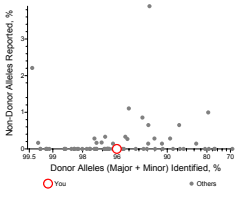
The National Institute of Standards and Technology
Gratefully Acknowledges the Participation of the

This feedback can be helpful to a laboratory to know where they stand relative to other labs to illustrate opportunities for improvement.

DNA Quantitation



Accuracy in STR Typing



Margaret C. Kline, Study Coordinator

See Kline, M.C., *et al.* (2003) *Anal. Chem.* 75: 2463-2469

NIST Quantitation Study 2004 (QS04)

Kline *et al.* (2005) *J. Forensic Sci.* 50(3):571-578

Consisted of:

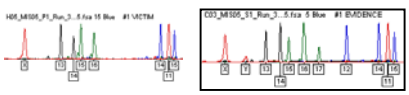
- 8 DNA extracts labeled A – H
- Shipped Dec 2003 –Jan 2004 to 84 laboratories for quantification; data received back by April 2004
- Labs were requested to use multiple methods / multiple analysts

We received data from 80 Labs (95%)
Total of 287 sets of data
Participants used 19 different quantification methods (primarily variations on Quantiblot and Real-time PCR)

Information from this interlab study is being used to help construct SRM 2372 (Human DNA Quantitation Standard)

Mixture Interpretation Interlab Study
(MIX05)

- **Only involves interpretation of data**
- **105 labs enrolled** for participation (20 from overseas)
- **71 labs have returned results**
- Four mock cases supplied with "victim" and "evidence" electropherograms (GeneScan .fsa files – that can be converted for Mac or GeneMapper; gel files made available to FMBIO labs)
- Data available with Profiler Plus, COfiler, SGM Plus, PowerPlex 16, Identifiler, PowerPlex 16 BIO (FMBIO) kits
- Summary of results with involve training materials to illustrate various approaches to solving mixtures



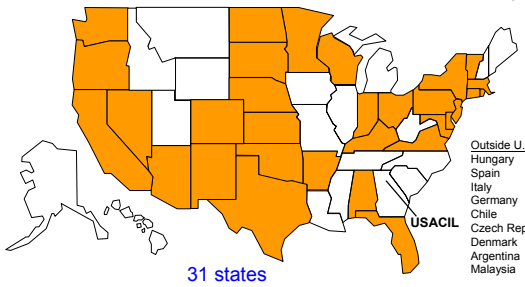
Perpetrator Profile(s) ??

Along with reasons for making calls and any stats that would be reported

Participation in NIST Interlaboratory Study on Mixture Interpretation (MIX05)

Companies:
Myriad Genetics

20 labs outside of U.S. signed up



Outside U.S.:
Hungary
Spain
Italy
Germany
Chile
Czech Republic
Denmark
Argentina
Malaysia

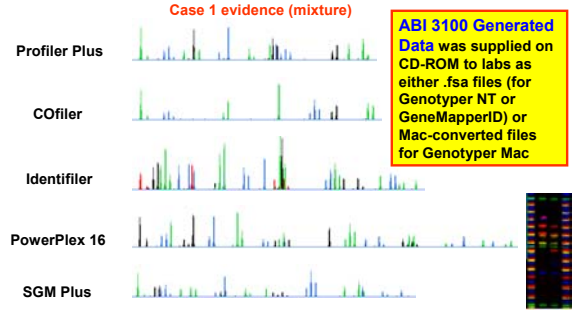
31 states

105 laboratories signed up for study (71 returned results so far)

MIX05 Results on Multiple Kits

http://www.cstl.nist.gov/biotech/strbase/interlab/MIX05.htm

Case 1 evidence (mixture)



ABI 3100 Generated Data was supplied on CD-ROM to labs as either .fsa files (for Genotyper NT or GeneMapperID) or Mac-converted files for Genotyper Mac

FMBIO data was also made available upon request

Purpose of MIX05 Study

- **Goal is to understand the "lay of the land" regarding mixture analysis across the DNA typing community**
- "If you show 10 colleagues a mixture, you will probably end up with 10 different answers"
  - Peter Gill, Human Identification E-Symposium, April 14, 2005
- One of the primary benefits we hope to gain from this study is **recommendations for a more uniform approach to mixture interpretation** and training tools to help educate the community

Value of the MIX05 Study

- Data sets exist with multiple mixture scenarios and a variety of STR kits that **can be used for training purposes**
- A wide variety of approaches to mixture interpretation have been applied on the **same data sets evaluated as part of a single study**
- **Interpretation guidelines from many laboratories are being compared to one another for the first time in an effort to determine challenges facing future efforts to develop "expert systems" for automated mixture interpretation**
- **We are exploring the challenges of supplying a common data set to a number of forensic laboratories** (e.g., if a standard reference data set was ever desired for evaluating expert systems)

Technology: Research Programs

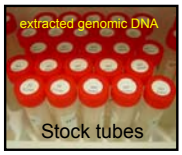
- miniSTRs
- Y-chromosome STRs
- mtDNA
- SNPs
- qPCR for DNA quantitation
- DNA stability studies
- Variant allele characterization and sequencing
- Software tools
- Assay development with collaborators

Standard U.S. Population Dataset

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

260 Caucasians, 260 African Americans, 140 Hispanics, 3 Asians = **663 males**

DNA extracted from whole blood (anonymous; self-identified ethnicities) received from Interstate Blood Bank (Memphis, TN) and Millennium Biotech Inc. (Ft. Lauderdale, FL)

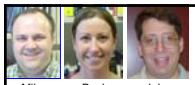


To date: (>100,000 allele calls)

Identifiler (15 autosomal markers + Amelogenin) (10,608)
Roche Linear Arrays (HV1/HV2 10 regions) (6,630)
Y STRs 22 loci—27 amplicons (17,388)
Y STRs 27 new loci (14,535)
Yfiler kit 17 loci (11,237)
Y SNPs 50 markers on sub-set of samples (11,498)
Orchid 70 autosomal SNPs on sub-set (13,230)
miniSTR testing-new loci and CODIS concordance (9,228)
New miniSTR loci – for 11 loci, 7,293 genotypes
mtDNA full control region sequences by AFDIL

Genotypes with various human identity testing markers

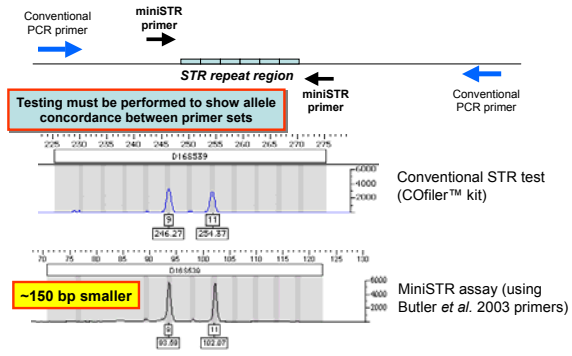
miniSTRs for Degraded DNA



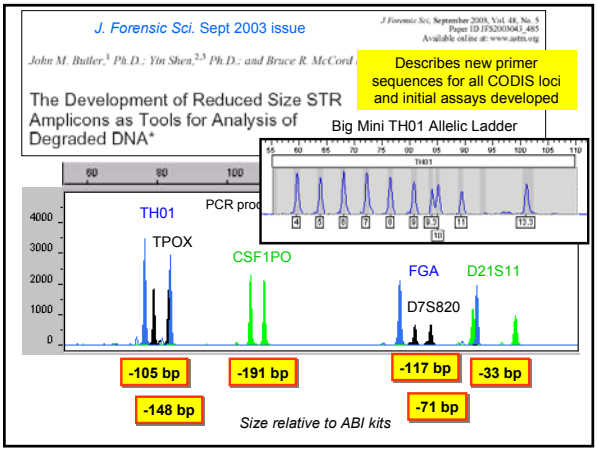
- Original miniSTR paper with CODIS loci, D2, D19, Penta D, Penta E
  - Butler *et al.* (2003) *J. Forensic Sci.* 48: 1054-1064
- Many CODIS loci are too big and make poor miniSTRs
- New miniSTRs and assays: NC01, NC02
  - Coble, M.D. and Butler, J.M. (2005) *J. Forensic Sci.* 50:43-53
- New miniSGM miniplex: AMEL, TH01, FGA, D18, D16, D2
- EDNAP/ENFSI degraded DNA study coordinated by Peter Gill
- Creation of miniSTR information on STRBase

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

A miniSTR is a reduced size STR amplicon that enables higher recovery of information from degraded DNA samples



Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition*, Figure 7.2, ©Elsevier Science/Academic Press



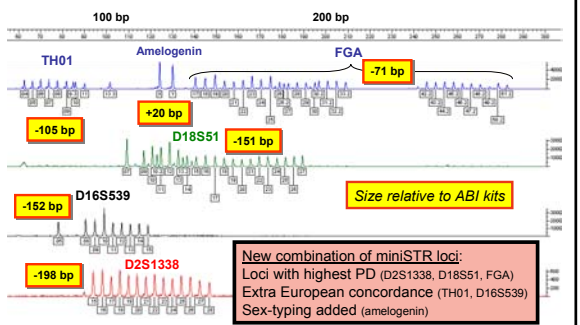
Recent Publications on miniSTRs

- Butler, J.M., Shen, Y., McCord, B.R. (2003) The development of reduced size STR amplicons as tools for analysis of degraded DNA. *J. Forensic Sci* 48(5): 1054-1064.
- Chung, D.T., Drabek, J., Opel, K.L., Butler, J.M., McCord, B.R. (2004) A study on the effects of degradation and template concentration on the efficiency of the STR miniplex primer sets. *J. Forensic Sci.* 49(4): 733-740.
- Drabek, J., Chung, D.T., Butler, J.M., McCord, B.R. (2004) Concordance study between miniplex STR assays and a commercial STR typing kit, *J. Forensic Sci.* 49(4): 859-860.
- Coble, M.D. and Butler, J.M. (2005) Characterization of new miniSTR loci to aid analysis of degraded DNA., *J. Forensic Sci.*, 50: 43-53.

http://www.cstl.nist.gov/biotech/strbase/miniSTR.htm
http://www.cstl.nist.gov/biotech/strbase/miniSTR/timeline.htm

New miniSGM miniplex assay

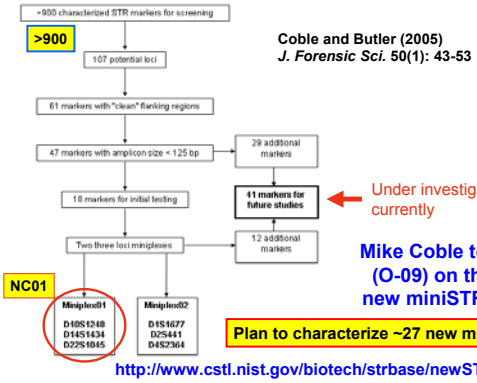
Provided to EDNAP/ENFSI group for degraded DNA study (Fall 2004)



http://www.cstl.nist.gov/biotech/strbase/miniSTR/miniSGM_Protocol.pdf

Retains same miniSTR primers from Butler et al. (2003) *J. Forensic Sci* 48(5): 1054-1064

Initial Testing Results with Potential miniSTR Loci



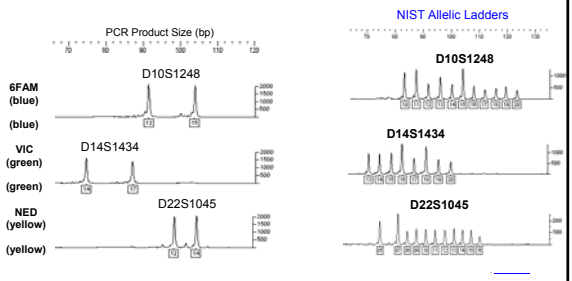
Coble and Butler (2005)
***J. Forensic Sci.* 50(1): 43-53**

Mike Coble to talk (O-09) on these new miniSTR loci

Plan to characterize ~27 new miniSTRs

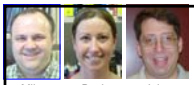
http://www.cstl.nist.gov/biotech/strbase/newSTRs.htm

Miniplex "NC01"



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

Coble and Butler (2005) Characterization of new miniSTR loci to aid analysis of degraded DNA *J. Forensic Sci.* 50(1): 43-53

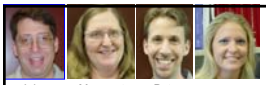


Mike Coble, Becky Hill, John Butler

New Autosomal STR Loci

- NC01 loci: **D10**S1248**, D14**S1434**, D22**S1045
- Peter Gill and the EDNAP/ENFSI group have recommended the NC01 loci as an extension of current European core loci following their demonstrated success in the recent EDNAP degraded DNA interlab study
- Being adopted/explored in multiple U.S. paternity testing labs
  (BRT Labs, Orchid Cellmark East Lansing, DNA Diagnostics Center)
- Population data, locus characterization, and allelic ladders for **27 new autosomal STRs** under development as new miniSTRs
- All new STR loci are physically unlinked to CODIS and ENFSI core loci

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



John Butler, Margaret Kline, Pete Vallone, Amy Decker

Work with Y-STRs

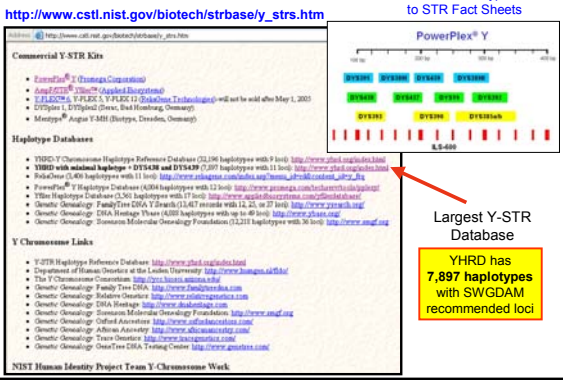
- Beta-testing of all commercial Y-STR kits
- U.S. population data supplied to Yfiler haplotype database
- **49 Y-STR loci evaluated with ~650 U.S. samples**
- New Y-chromosome information on STRBase linking to all available haplotype databases
- Human Y-Chromosome DNA Profiling Standard Reference Material (SRM 2395) – updates with DYS635 for Yfiler
- **Separation of two brothers with 47 Y-STRs**

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

New Y-Chromosome Information Resources on STRBase

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

Locus boxes are hyperlinked to STR Fact Sheets



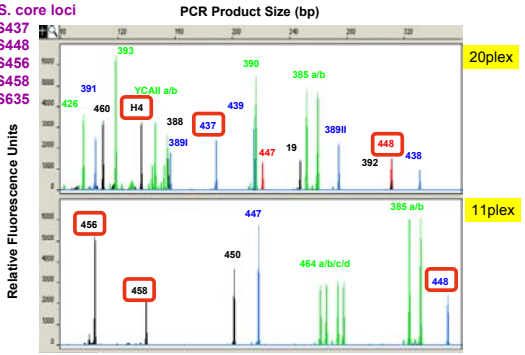
Largest Y-STR Database

YHRD has **7,897 haplotypes** with SWGDAM recommended loci

NIST Multiplexes for High-Throughput Y STR Typing

Laid the Groundwork for Applied Biosystems' Yfiler Y-STR Kit

11 U.S. core loci
+ DYS437
+ DYS448
+ DYS456
+ DYS458
+ DYS635
+ H4



Schoske *et al.* (2004) High-throughput Y-STR typing of U.S. populations…, *Forensic Sci. Int.*, 139:107-121

Yfiler Haplotype Database

http://www.appliedbiosystems.com/yfilerdatabase/

Population	# Haplotypes
African American	985
Asian	330
Caucasian	1276
Filipino	105
Hispanic	597
Native American	106
Sub-saharan African	59
Vietnamese	103
All	3561

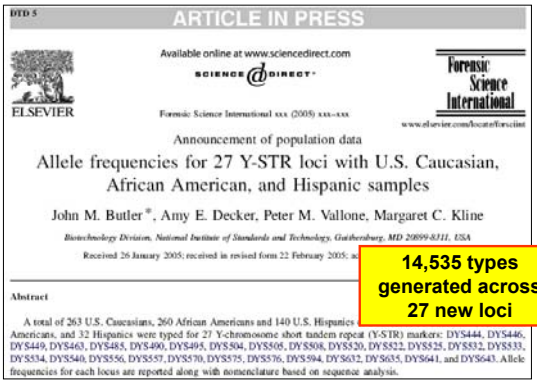
#Samples Contributed by NIST

259 African Americans
3 Asians
262 Caucasians
139 Hispanics

Data provided by NIST

663/3561 = **18.6%**

NIST Work with New Y-STR Loci



DTD 5
ARTICLE IN PRESS

Available online at www.sciencedirect.com
SCIENCE DIRECT

ELSEVIER

Forensic Science International

Announcement of population data

Allele frequencies for 27 Y-STR loci with U.S. Caucasian, African American, and Hispanic samples

John M. Butler *, Amy E. Decker, Peter M. Vallone, Margaret C. Kline

Abstract

A total of 263 U.S. Caucasians, 260 African Americans and 140 U.S. Hispanics ... Americans, and 32 Hispanics were typed for 27 Y-chromosome short tandem repeat (Y-STR) markers: DYS444, DYS446, DYS449, DYS463, DYS485, DYS490, DYS495, DYS504, DYS505, DYS508, DYS520, DYS522, DYS525, DYS532, DYS533, DYS534, DYS540, DYS556, DYS557, DYS570, DYS575, DYS576, DYS594, DYS632, DYS635, DYS641, and DYS643. Allele frequencies for each locus are reported along with nomenclature based on sequence analysis.

14,535 types generated across 27 new loci

Subdividing Common Types with More Loci

657 males from 3 U.S. populations

Most common type
DYS19 – 14
DYS389I – 13
DYS389II – 29
DYS390 – 24
DYS391 – 11
DYS392 – 13
DYS393 – 13
DYS385 a/b – 11,14

Minimal haplotype
(19, 389I/II, 390, 391, 392, 393, 385 a/b)

U.S. Haplotype (26)

SWGDAM recommended loci (+ 438,439)

(7) (15) (4)

PowerPlex Y (+437)

(2) (4) (1) (3) (12) (1) (3)

Yfiler (+448,456,458,635,H4)

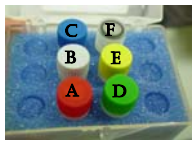
(1)(1) (1)(1)(1)(1) (1)(1) (1)(1)(1) (1)(1)(1)(1)(1)(1)(1)(3)(1)(1) (1) (1)(1)(1)

Identical: DYS…444,446,485,495,505,508,534,540,556

Subdivide into two groups (2)(1): DYS…449,463,520,532,533,557,570,594,643
Subdivide into three groups (1)(1)(1): **DYS522, DYS576**

Y-Chromosome Standard NIST SRM 2395

Available since July 2003



Human Y-Chromosome DNA Profiling Standard

- **5 male samples** + 1 female sample (neg. control)
- **100 ng of each (50 µL at ~2 ng/µL)**
- **22 Y STR markers sequenced**
- **9 additional Y STR markers typed**
- **42 Y SNPs typed** with Marligen kit

Certified for all loci in commercial Y-STR kits:

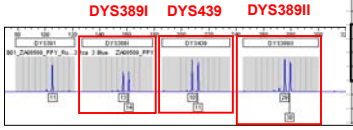
Y-PLEX 6
Y-PLEX 5
Y-PLEX 12
PowerPlex Y

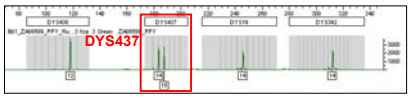
SWGDAM recommended loci:
DYS19, DYS385 a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439

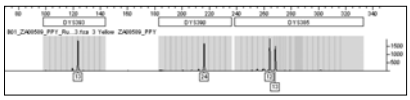
Y-filer - adds DYS635 (C4); now sequenced

Helps meet FBI Standard 9.5 (and ISO 17025)…traceability to a national standard

Duplication at Multiple Loci with Single-Source Sample







Y-chromosome mapping

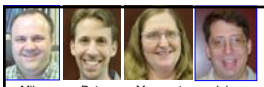
q-arm

Y STR Marker	Position (Mb)
DYS391	13.413
DYS635 (C4)	13.690
DYS434	13.777
DYS437	13.778
DYS435	13.807
DYS439	13.826
DYS389 I/II	13.923
DYS388	14.057
DYS442	14.071
DYS438	14.248

Entire region of Y-chromosome has likely been duplicated and then diverged

Most duplications have a single repeat spread in allele patterns

Butler *et al.* (2005) *J. Forensic Sci.* 50(4):853-859



Mike Coble | Pete Vallone | Margaret Kline | John Butler

Work with mtDNA

- **Collaboration with Armed Forces DNA Identification Laboratory** to develop multiplex mtDNA SNP assays for coding region polymorphisms
- Beta-testing and automation of Roche LINEAR ARRAY HVI/HVII probes
- Population study performed with LINEAR ARRAY HVI/HVII probes
- Exploration of effective strategies for forensic analysis in the mitochondrial DNA coding region

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

Control Region (16024-576) SNPs

- 1,122 nucleotide positions; typically only 610 bases analyzed (HVI: 16024-16365; HVII: 73-340)
- Challenges with typing closely spaced SNPs
  - Probes are disrupted by neighboring polymorphism(s)

Coding Region (577-16023) SNPs

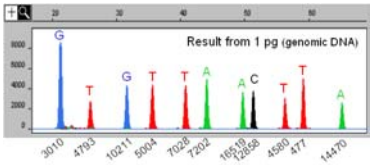
- 15,446 nucleotide positions
- Challenges with typing widely spaced SNPs
  - Multiplex PCR required
- Polymorphisms may have medical significance

NIST mtDNA Work

Roche Linear Arrays
(probes for HVI/HVII)

J. Forensic Sci. 2005, 50(2): 377-385

Automated washing/ Population Study



Coding Region mtSNP 11plex
(minisequencing assay)

Developed with AFDIL to resolve mtDNA most common types

Int. J. Legal Med., 2004; 118: 147-157

Typing frequencies for 666 NIST population samples

# in Group	Freq	% Types	% People
1	185	65.6	27.8
2	46	16.3	13.8
3	18	6.4	8.1
4	4	1.4	2.4
5	3	1.1	2.3
6	4	1.4	3.6
7	1	0.4	1.1
8	9	3.2	10.8
9	2	0.7	2.7
10	4	1.4	6.0
11	1	0.4	1.7
12	1	0.4	1.8
18	1	0.4	2.7
23	1	0.4	3.5
28	1	0.4	4.2
51	1	0.4	7.7

Summary of Our Population Typing with Roche mtDNA LINEAR ARRAYS

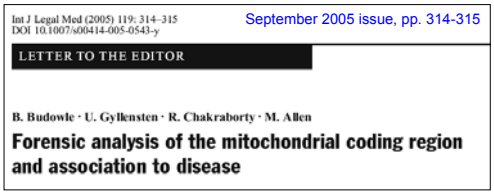
LINEAR ARRAY summary
- 282 different types
- 185 were unique (occurred only once)
- 51 samples had "Most Common Type"

HV1/HV2 sequencing summary
- 518 different types
- 454 were unique (occurred only once)
- 17 samples had "Most Common Type"

"Most Common Type" evaluated further with mtDNA coding region SNP assay

Kline *et al.* (2005) *J. Forensic Sci.* 50(2): 377-385

Analysis of mtDNA coding region


Int J Legal Med (2005) 119: 314–315
DOI 10.1007/s00414-005-0543-y

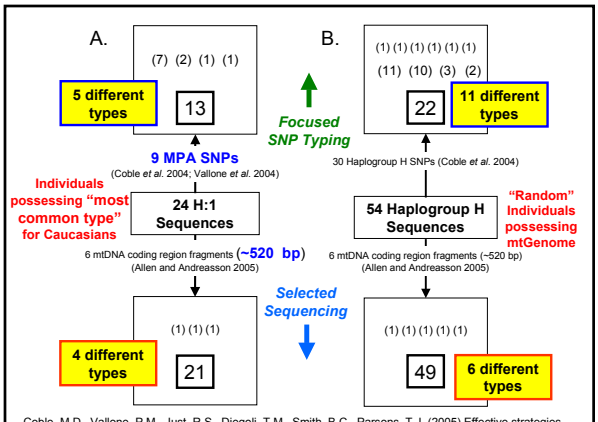
September 2005 issue, pp. 314-315

LETTER TO THE EDITOR

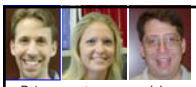
B. Budowle · U. Gyllensten · R. Chakraborty · M. Allen

Forensic analysis of the mitochondrial coding region and association to disease

- They propose that sequencing stretches of the coding region is more informative than evaluating selected SNP sites…



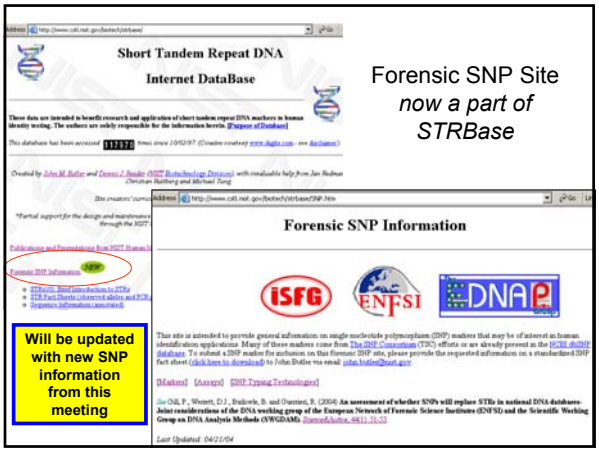
Coble, M.D., Vallone, P.M., Just, R.S., Diegoli, T.M., Smith, B.C., Parsons, T.J. (2005) Effective strategies for forensic analysis in the mitochondrial DNA coding region. *Int. J. Legal. Med., in press*


Pete Vallone
Amy Decker
John Butler

Work with SNP Loci

- U.S. population frequencies with 70 autosomal SNPs
  - Vallone *et al.* (2005) *Forensic Sci. Int.* 149: 279-286
- U.S. population information with 50 Y-SNPs
  - Vallone *et al.* (2004) *J. Forensic Sci.* 49: 723-732
- Construction of 12plex autosomal SNP assay
  - Vallone *et al.* (2005) **Poster P-296 at ISFG**
- Creation of Forensic SNP Information website on STRBase
  - see Gill et al. *Science&Justice* 44(1): 51-53

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



Forensic SNP Site
now a part of STRBase

Allele Frequencies for 70 SNP Loci in U.S. Populations

Hispanic N = 44	1
CC	0.455
TT	0.068
CT	0.477
He	0.425
p	0.723
	13
CC	0.068
TT	0.568
CT	0.364
He	0.375
p	1.000
	25
CC	0.318
TT	0.114
CT	0.568
He	0.479
p	0.326
	37
CC	0.523
TT	0.045
CT	0.432
He	0.386
p	0.894
	49
CC	0.068
TT	0.636
CT	0.295
He	0.339
p	0.381
	61
CC	0.068
TT	0.455
CT	0.477
He	0.525
p	0.721

African American N = 71	1
CC	0.648
TT	0.070
CT	0.282
He	0.333
p	0.278
	13
CC	0.141
TT	0.451
CT	0.408
He	0.452
p	0.296
	25
CC	0.099
TT	0.394
CT	0.507
He	0.456
p	0.310
	37
CC	0.324
TT	0.183
CT	0.493
He	0.490
p	1.000
	49
CC	0.085
TT	0.549
CT	0.366
He	0.392
p	0.369
	61
CC	0.310
TT	0.225
CT	0.465
He	0.496
p	0.634

Caucasian N = 74	1	2	3	4	5	6	7	8	9	10	11	12
CC	0.243	0.405	0.068	0.581	0.311	0.149	0.486	0.108	0.203	0.068	0.257	0.054
TT	0.243	0.135	0.514	0.135	0.189	0.338	0.122	0.378	0.284	0.459	0.216	0.365
CT	0.514	0.459	0.419	0.284	0.500	0.514	0.392	0.514	0.514	0.473	0.527	0.581
He	0.500	0.463	0.497	0.401	0.493	0.482	0.433	0.463	0.497	0.423	0.499	0.452
p	0.816	1.000	1.000	**0.008**	0.816	0.475	0.413	0.305	1.000	0.269	0.818	**0.021**
	13	14	15	16	17	18	19	20	21	22	23	24
CC	0.068	0.243	0.392	0.446	0.243	0.162	0.473	0.365	0.270	0.108	0.270	0.203
TT	0.514	0.311	0.122	0.149	0.284	0.324	0.054	0.203	0.189	0.432	0.176	0.270
CT	0.419	0.446	0.486	0.405	0.473	0.514	0.473	0.432	0.541	0.459	0.554	0.527
He	0.401	0.498	0.463	0.496	0.499	0.497	0.412	0.487	0.497	0.447	0.498	0.498
p	1.000	0.220	0.805	0.183	0.663	0.818	0.163	0.348	0.485	1.000	0.362	0.659
	25	26	27	28	29	30	31	32	33	34	35	36
CC	0.243	0.176	0.162	0.068	0.257	0.432	0.419	0.527	0.122	0.581	0.257	0.108
TT	0.203	0.446	0.432	0.689	0.284	0.149	0.122	0.122	0.311	0.068	0.203	0.243
CT	0.554	0.378	0.405	0.243	0.459	0.419	0.459	0.351	0.568	0.351	0.541	0.649
He	0.499	0.463	0.463	0.307	0.500	0.460	0.456	0.418	0.482	0.368	0.499	0.491
p	0.480	0.135	0.327	0.119	0.350	0.302	0.797	0.092	0.088	0.538	0.642	**0.008**
	37	38	39	40	41							
CC	0.378	0.095	0.378	0.149	0.257							
TT	0.122	0.473	0.149	0.514	0.216							
CT	0.500	0.432	0.473	0.338	0.486							
He	0.467	0.428	0.474	0.433	0.497							
p	0.444	0.790	0.806	0.060	0.827							
	49	50	51	52	53							
CC	0.137	0.205	0.178	0.356	0.096							
TT	0.582	0.370	0.247	0.178	0.534							
CT	0.301	0.425	0.575	0.466	0.370							
He	0.410	0.486	0.498	0.484	0.404	0.495	0.452	0.452	0.470	0.331	0.401	0.489
p	**0.040**	0.230	0.226	0.630	0.401	0.639	0.472	0.612	0.618	0.068	0.785	0.331
	61	62	63	64	65	66	67	68	69	70		
CC	0.068	0.473	0.189	0.162	0.233	0.378	0.486	0.324	0.216	0.284		
TT	0.608	0.027	0.486	0.284	0.219	0.162	0.054	0.108	0.351	0.149		
CT	0.324	0.500	0.324	0.554	0.548	0.459	0.459	0.568	0.432	0.568		
He	0.354	0.401	0.458	0.493	0.500	0.477	0.407	0.477	0.491	0.491		
p	0.326	**0.043**	**0.011**	0.232	0.500	0.805	0.401	0.142	0.346	0.263		

Vallone *et al.* (2005) *Forensic Sci. Int.* 149:279-286

These are the Orchid SNP markers used in their WTC testing

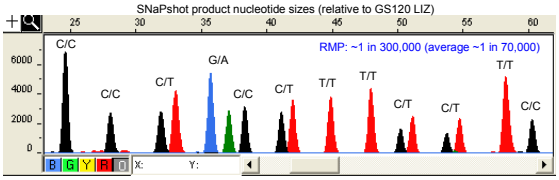
See Peter Vallone poster #296 (session II)

Autosomal 12plex SNP Assay

12plex PCR and 12plex ASPE from 0.5 ng DNA template

all PCR products <90 bp

SNaPshot product nucleotide sizes (relative to GS120 LIZ)



locus:	1	2	3	4	5	6	7	8	9	10	11	12
chr:	5	15	10	1	17	13	17	1	6	11	20	11

"Best" 12 SNPs selected from 70 originally tested;
>0.45 heterozygosity in U.S. Caucasians, Hispanics, and African Americans

Vallone, P.M., Decker, A.E., Butler, J.M. (2005) Allele frequencies for 70 autosomal SNP loci with U.S. Caucasian, African American, and Hispanic Samples. *Forensic Sci. Int.* 149:279-286.

Population data has been collected on >1,000 samples (662 U.S. and 375 world samples) **from 10 different population groups**

Publication on U.S. Groups with Y-SNPs

J. Forensic Sci. 2004; 49(4): 723-732

J Forensic Sci, July 2004, Vol. 49, No. 4
Paper ID JFS2003303
Available online at: www.astm.org

Peter M. Vallone,[1] Ph.D. and John M. Butler,[1] Ph.D.

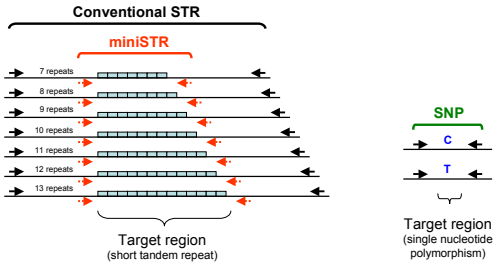
Y-SNP Typing of U.S. African American and Caucasian Samples Using Allele-Specific Hybridization and Primer Extension*

Different technologies yield the same Y-SNP type
Full concordance was observed between hybridization and primer extension technologies on 18 different Y-SNPs (>3,800 allele calls)

Y-SNPs will have limited value for individualizing a sample
18 different types observed in 229 individuals

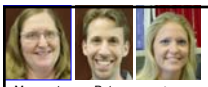
Current Y-SNPs appear to have limited value for ethnic differentiation in U.S. populations (with the exception of M2 that is only found in African Americans and not in Caucasians)

Comparison of STRs and SNPs



Larger target region (miniSTR targets same region)
More possible variants than SNPs
Only need a moderate number of STR markers
Range of sizes examined (e.g., 28 bp spread if 4 bp/repeat)
Mixtures easier to detect and decipher

Smaller target region
Fewer possible variants
Need more SNP markers
Constant size examined



Margaret Kline | Pete Vallone | Amy Decker

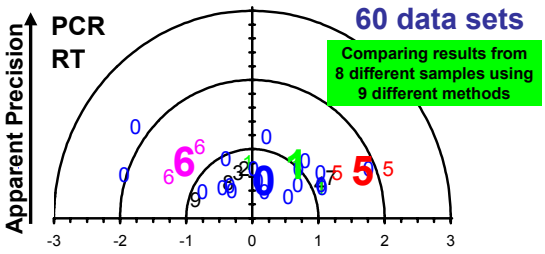
Evaluation of qPCR Assays

- Evaluation of published assays on same samples
- Characterization of Quantifiler lot-to-lot performance
- Additional studies under way utilizing qPCR:
  - Examining the challenge of multiplexing qPCR assays
  - Studies to track DNA recovery from various types of tubes
  - **Characterizing potential SRM 2372 components** (Human DNA Quantitation Standard)

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

Interlaboratory Comparisons

Laboratory Performances with Real-Time PCR Methods



0 = Quantifiler
1 = Alu RT-PCR
5 = BRCA1
6 = CFS-HUMRT

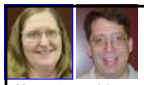
Kline, *et al.* (2005) *J. Forensic Sci.* 50(3):571-578

Margaret Kline Dave Duewer

DNA Stability Studies

- Comparison of DNA recovery from FTA, IsoCode, S&S 903, Whatman papers with bloodstains from the same individuals
- Recovery of DNA from aged bloodstains on untreated paper stored at room temperature

http://www.cstl.nist.gov/biotech/strbase/NIJ/DNArecovery.htm



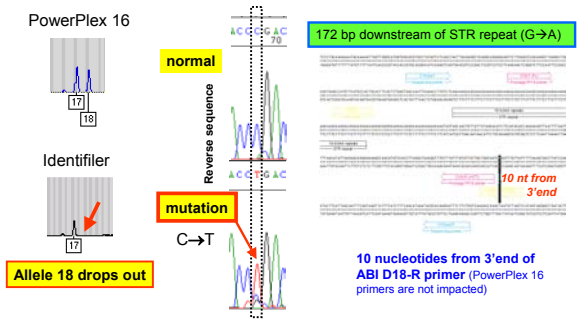
Margaret Kline John Butler

STR Allele Sequencing and Characterization

- Variant characterization
  - TPOX 10.3 (Maryland State Police)
  - **D18S51 null alleles** (FSS and Kuwait govt)
  - D18S51 allele 40 (Nebraska State Crime Lab)
  - D18S51 allele "5.3" (DNA Solutions)
  - DYS392 allele "10.3" (AFDIL)
  - VWA allele "15.1" (Peter de Knijff)
- Locus duplication or deletion
  - DYS390 deletion (CFS Toronto)
  - DYS392 deletion (MN BCA)
- **Forensic labs are sending us unusual STR alleles for sequence characterization**

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

D18S51 Null Allele from Kuwait Samples with ABI Primers



Clayton et al. (2004) **Primer binding site mutations affecting the typing of STR loci contained within the AMPFISTR SGM Plus kit.** *Forensic Sci Int.* 139(2-3): 255-259

Pete Vallone Dave Duewer Chris DeAngelis

Software Tools

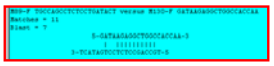
- AutoDimer – multiplex PCR primer screening tool

http://www.cstl.nist.gov/biotech/strbase/AutoDimerHomepage/AutoDimerProgramHomepage.htm

- mixSTR – mixture component resolution tool
- Multiplex_QA – quality assessment tool for monitoring instrument performance over time
- NIST U.S. population database (internal Access database)

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

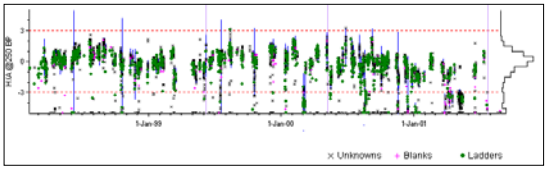
AutoDimer Overview



- **Evaluates potential primer dimers** and hairpins in batch mode.
- Requires input of primer list (already chosen in singleplex fashion with Primer3 or some other program). Runs as stand-alone Visual Basic program.
- Used in development of all NIST multiplex PCR/STR/SNP assays.

Available for download from STRBase:
http://www.cstl.nist.gov/biotech/strbase/software.htm

Multiplex_QA Overview

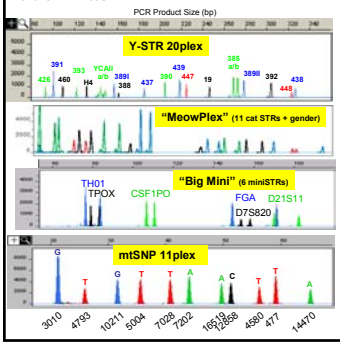


- **Research tool** that provides quality metrics to review instrument performance over time (e.g., examines resolution on internal size standard peaks)
- Runs with Microsoft Excel macros. Requires STR data to be converted with NCBI's BatchExtract program into numerical form.

Available for download from STRBase:
http://www.cstl.nist.gov/biotech/strbase/software.htm

Assay Development with Collaborators


Pete Vallone John Butler



- **Y-STR 20plex** (Mike Hammer and Alan Redd)
  - Butler et al. (2002) FSI
- **Cat STRs** (Marilyn Raymond and Victor David)
  - Butler et al. (2002) Profiles in DNA
  - Raymond et al. (2005) JFS
- **miniSTRs** (Bruce McCord)
  - Butler et al. (2003) JFS
- **mtDNA coding region SNPs** (AFDIL)
  - Vallone et al. (2004) IJLM
- Others…

Training Materials and Review Articles


John Butler

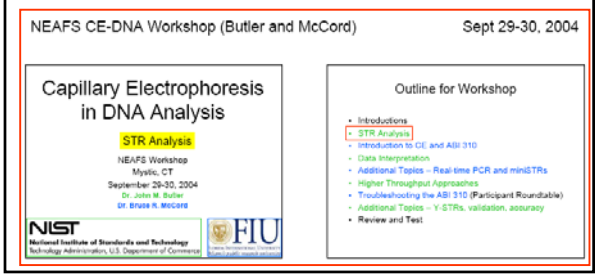
- Workshops on STRs and CE (ABI 310/3100)
  - Taught with Bruce McCord (Florida Int. Univ.)
  - NEAFS (Sept 29-30, 2004)
  - U. Albany DNA Academy (June 13-14, 2005)
- Validation Workshop
  - Taught with Robyn Ragsdale (FDLE) at NFSTC (August 24-26, 2005)
- PowerPoint slides from *Forensic DNA Typing*, 2nd Edition
  - >150 slides available now (~1,000 planned) for download
  - **http://www.cstl.nist.gov/biotech/strbase/FDT2e.htm**
- Review articles
  - ABI 310 and 3100 chemistry – Electrophoresis 2004, 25, 1397-1412

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

NEAFS Workshop Slide Handouts

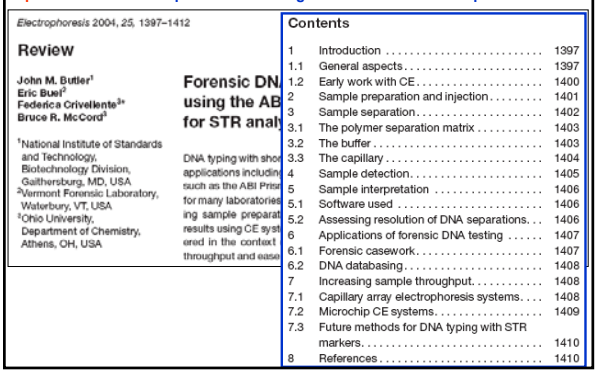
Handouts available as downloadable pdf files from
http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm#NEAFSworkshop

2 day workshop with **>500 slides** describing STRs and CE (ABI 310 and ABI 3100)



Review Article on STRs and CE

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



Analytical Chemistry Application Review

June 15, 2005 issue of *Analytical Chemistry*

Forensic Science

T. A. Brettell*

Office of Forensic Sciences, New Jersey State Police, New Jersey Forensic Science and Technology Complex, 1200 Negron Road, Horizon Center, Hamilton, New Jersey 08691

J. M. Butler

National Institute of Standards and Technology, Gaithersburg, Maryland 20899-8311

R. Saferstein

Box 1334, Mount Laurel, New Jersey 08054

Review Contents
- Forensic DNA Analysis
  - Collection, Characterization, Preservation, Extraction, and Quantitation of Biological Material
  - Short Tandem Repeats
  - Single-Nucleotide Polymorphisms
  - Y-STR Typing, Gender Identification, and X-Chromosome Analysis
  - Mitochondrial DNA Typing
  - Nonhuman DNA Typing Systems and Microbial Forensics
  - DNA Databases
  - Interpretation and Statistical Weight of DNA Typing Results
  - General Reviews

250 articles referenced covering forensic DNA analysis during 2003-2004

Acknowledgments

Funding from interagency agreement 2003-IJ-R-029 between NIJ and the NIST Office of Law Enforcement Standards

John Butler, Margaret Kline, Pete Vallone, Mike Coble, Jan Redman, Amy Decker, Becky Hill, Chris DeAngelis, Dave Duewer

The many forensic scientists and their supervisors who took time out of their busy schedules to examine the MIX05 data provided as part of this interlaboratory study

Past and Present Collaborators (also funded by NIJ):
Mike Hammer and Alan Redd (U. AZ) for Y-chromosome studies
Tom Parsons, Rebecca Just, Jodi Irwin (AFDIL) for mtDNA coding SNP work
Sandy Calloway (Roche) for mtDNA LINEAR ARRAYs
Bruce McCord and students (FL Int. U.) for miniSTR work
Marilyn Raymond and Victor David (NCI-Frederick) for cat STR work
Artie Eisenberg and John Planz (U. North Texas) for miniSTR testing on bones
Murray Brilliant (U. AZ) for phenotype markers
Ken Kidd (Yale U.) for SNP typing population samples

Thank you for your attention…

Questions?

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

Contact Information:
john.butler@nist.gov
001-301-975-4049
http://www.cstl.nist.gov/biotech/strbase