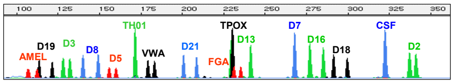


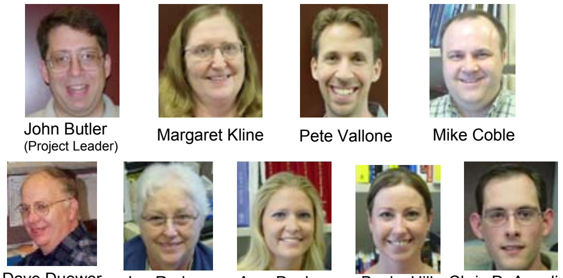
NIST On-Going Projects to Aid the Human Identity Testing Community

John Butler

Margaret Kline, Pete Vallone, Mike Coble
Jan Redman, Amy Decker, Becky Hill, Chris DeAngelis
Dave Duewer (NIST Analytical Chemistry Division)

NIJ DNA Grantees Meeting – June 29, 2005

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

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

Team Impact on Forensic Community

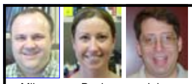
- **27 publications** since June 2004 (61 since 2000)
- **31 presentations** to the community since June 2004
- All NIST publications and presentations available on STRBase: http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
- Training materials: 2 workshops conducted with Bruce McCord
  - NEAFS (Sept 29-30, 2004)
  - Albany DNA Academy (June 13-14, 2005)
- *Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers*, 2nd Edition (John Butler)

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

Current Areas of NIST Research Effort

- **Resources for "Challenging Samples"** (miniSTRs)
- **Information on New Loci** (SNPs, Y-Chromosome, new STRs)
- **Standard Information Resources** (STRBase website, training materials/review articles, validation standardization)
- **Allele Sequencing and Interlaboratory Studies** (Real-time qPCR, mixture interpretation)

MIX05 Study Review

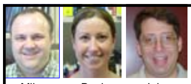


Mike Coble
Becky Hill
John Butler

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



Mike Coble
Becky Hill
John Butler

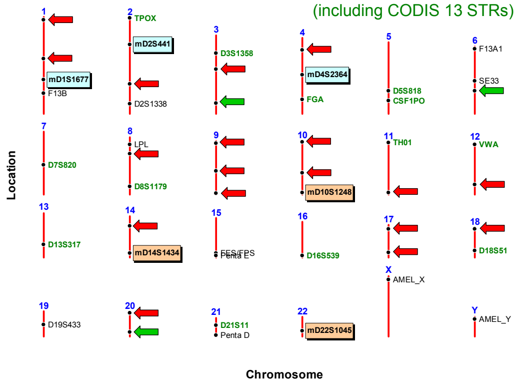
New Autosomal STR Loci

- 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
- 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 core loci

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

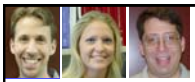
STR Loci Positions

(including CODIS 13 STRs)



Positions determined along May 2004 Human Genome Reference Sequence (NCBI Build 35)

Work with SNP Loci

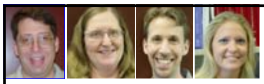


Pete Vallone
Amy Decker
John Butler

- 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
- 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

Work with Y-STRs

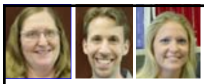


John Butler
Margaret Kline
Pete Vallone
Amy Decker

- Beta-testing of all commercial Y-STR kits
- 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
- Nomenclature defined for new loci
- 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

Evaluation of qPCR Assays

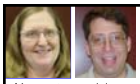


Margaret Kline
Pete Vallone
Amy Decker

- 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

STR Allele Sequencing and Characterization

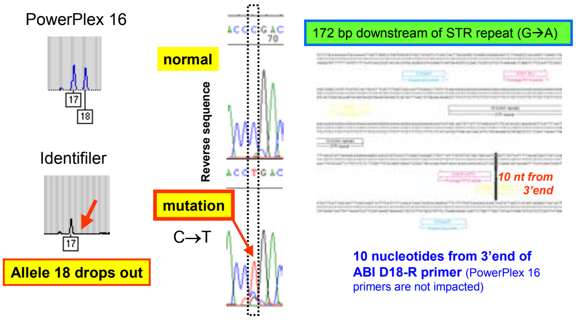


Margaret Kline
John Butler

- 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)
  - FGA allele 46.2 (Denver Crime Lab)
  - DYS392 allele "10.3" (AFDIL)
- Locus duplication or deletion
  - DYS390 (CFS Toronto)
  - DYS392 (MN BCA)
- **Send us your unusual STR alleles for sequence characterization**

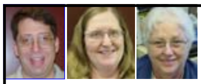
http://www.cstl.nist.gov/biotech/strbase/STRseq.htm

D18S51 Null Allele from Kuwait Samples with ABI Primers



10 nucleotides from 3'end of ABI D18-R primer (PowerPlex 16 primers are not impacted)

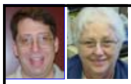
Clayton et al. (2004) **Primer binding site mutations affecting the typing of STR loci contained within the AMPFISTR SGM Plus kit.** *Forensic Sci Int.* 139(2-3): 255-259


John Butler Margaret Kline Jan Redman

Validation Standardization

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

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


John Butler Jan Redman

STRBase Updates

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 ~2,300 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


John Butler

Training Materials and Review Articles

- 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)
- PowerPoint slides from *Forensic DNA Typing*, 2nd Edition
- Review articles
  - ABI 310 and 3100 chemistry – Electrophoresis 2004, 25, 1397-1412
  - Forensic DNA analysis – Anal. Chem. 2005, 77, 3839-3860
  - STR core loci – J. Forensic Sci., *in press* (Nov 2005)

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


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

Part of R&D demonstrations on Tuesday, June 28, 2005

- NIST U.S. population database (internal Access database)

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


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)
  - 91 labs signed up (64 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

MIX05 Interlab Study Initial Review

- MIX05 study design and purpose
- Mixture selection, samples generation, and initial testing at NIST (case scenarios)
- Demographics of MIX05 study participants
- Value of this interlab study
- Opportunities for community improvement and standardization regarding mixture interpretation

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

MIX05 Study Design and Purpose

- **Goal is to understand the "lay of the land" regarding mixture analysis across the DNA typing community**
- Provide multiple case scenarios
- Permit a large number of forensic practioners to evaluate the same mixture data
- Generate data from multiple STR kits on the same mixture samples to compare performance in recovering potential mixture information

MIX05 Study Enrollment

- Announcements with handouts made at forensic meetings
  - CODIS User's Group (November 15, 2004)
  - Forensic Y User's Group (November 20, 2004)
  - SWGDAM (January 18, 2005)
- Emails to previous participants in NIST interlab studies such as Mixed Stain Study 3, DNA Quantitation Study 2004
- **70 labs initially enrolled** (28 states, 17 overseas)
- A second email push was made in January 2005
- **Total of 91 labs enrolled by March 2005**

Initial Requests for Kits and Data Format

STR kit requests	Analysis software requests
37 ProfilerPlus/COfiler	27 Genotyper Mac
16 PowerPlex 16	24 Genotyper NT
22 Identifiler	23 GeneMapper*ID*
1 SGM Plus	6 FMBIO Mac
8 FMBIO	2 FMBIO NT

Decision was made to supply all data for 5 different STR kits (ProPlus, COfiler, Identifiler, PP16, SGM Plus) **to all ABI kit labs**

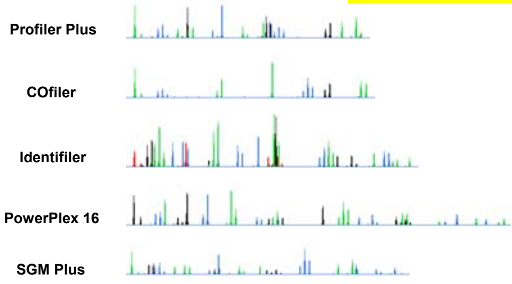
(FMBIO data supplied separately—generated in Pennsylvania State Police Lab and Arkansas State Crime Lab)

Data was shipped in mid-January 2005

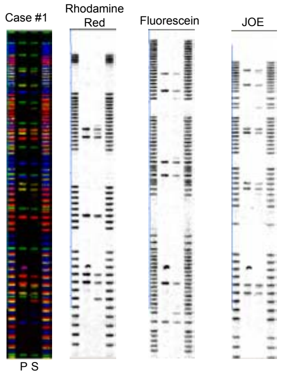
MIX05 Results on Multiple Kits

Case 1 evidence (mixture)

Data generated on ABI 3100 (also ABI 310 but not provided)



FMBIO Data Provided to Requesting Labs



PCR products generated at NIST using PP16 BIO

Pennsylvania State Police created **Mac compatible** FMBIO gel scans (Frank Krist and Chris Tomsey)

Arkansas State Crime Lab created **NT compatible** FMBIO gel scans (Mary Robnett and Kermit Channel)

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

Posted on January 13, 2005

Forensic SNP Information

- STRs101: Brief Introduction to STRs
- STR Fact Sheets (observed alleles and PCR product sizes)
- Sequence Information (annotated)
- Multiplex STR sets
- STR Training Materials

- Variant Allele Reports ◆
- Tri-Allelic Patterns ◆

- FBI CODIS Core STR Loci
- DNA Advisory Board Quality Assurance Standards
- NIST Standard Reference Material for PCR-Based Testing
- Chromosomal Locations
- Mutation Rates for Common Loci
- Published PCR primers
- Validation information Updated
- Interlaboratory Studies NEW NIST Interlab Study MIX05 Data
- Population data Updated
- Data from NIST U.S. Population Samples

- Y-chromosome STRs Updated
- miniSTRs (short amplicons) NEW
- Sex-typing markers
- Technology for resolving STR alleles

Interlab Study MIX05 Data Available for Download from STRBase

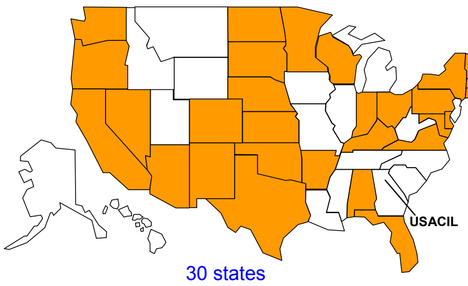
ABI 3100 Generated Data was also supplied on CD-ROM to all labs as either .fsa files (for Genotyper NT or GeneMapperID) or Mac-converted files for Genotyper Mac

---

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

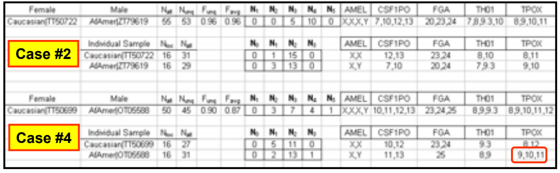
91 laboratories signed up for study (64 returned results so far)

Sample Selection Decisions

- Review of all possible allele combinations using **Virtual MixtureMaker** developed by Dave Duewer
- Compared 40 females against 660 males tested previously with Identifiler (Butler et al. JFS 2003;48(4):908-911)
- Scenarios to evaluate:
  - "Typical" ~1:3 mixture (Case #1)
  - Perpetrator is major contributor (Case #2)
  - "Balanced" ~1:1 mixture (Case #3)
  - "Extreme" ~1:10 mixture (Case #4)
- Supplied female "victim" and mixture "evidence" for each case (along with allelic ladder, pos. & neg. controls)

Sample Selection for MIX05

Output from Dave Duewer's Excel macro used to generate artificial mixtures



Female	Male	N_{all}	N_{unq}	F_{unq}	F_{avg}	N_1	N_2	N_3	N_4	N_5	AMEL	CSF1PO	FGA	TH01	TPOX
Caucasian|TT50722	AfAmer|ZT79619	55	53	0.96	0.96	0	0	5	10	0	X,X,X,Y	7,10,12,13	20,23,24	7,8,9.3,10	8,9,10,11

Case #2

Individual Sample	N_{loc}	N_{all}	N_0	N_1	N_2	N_3	AMEL	CSF1PO	FGA	TH01	TPOX
Caucasian|TT50722	16	31	0	1	15	0	X,X	12,13	23,24	8,10	8,11
AfAmer|ZT79619	16	29	0	3	13	0	X,Y	7,10	20,24	7,9.3	9,10

Female	Male	N_{all}	N_{unq}	F_{unq}	F_{avg}	N_1	N_2	N_3	N_4	N_5	AMEL	CSF1PO	FGA	TH01	TPOX
Caucasian|TT50699	AfAmer|OT05588	50	45	0.90	0.87	0	3	7	4	1	X,X,X,Y	10,11,12,13	23,24,25	8,9,9.3	8,9,10,11,12

Case #4

Individual Sample	N_{loc}	N_{all}	N_0	N_1	N_2	N_3	AMEL	CSF1PO	FGA	TH01	TPOX
Caucasian|TT50699	16	27	0	5	11	0	X,X	10,12	23,24	9.3	8,12
AfAmer|OT05588	16	31	0	2	13	1	X,Y	11,13	25	8,9	9,10,11

- Once the particular male and female samples were chosen to reflect a variety of allele combination scenarios, different ratios were decided upon and generated in the laboratory by mixing genomic DNA samples
- In all cases, 1 female was mixed with 1 male

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 set(s) and 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)

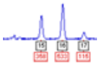
All samples amplified from the same DNA mixture and on the same thermal cycler using manufacturer protocols

Kit-to-Kit Variation:

D3S1358 Case1 Evidence Mixture from All 5 STR Kits

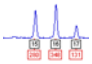
Peak Height Ratios

Profiler Plus



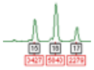
3.1 : 4.6 : 1

COfiler



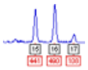
2.1 : 2.7 : 1

Identifiler



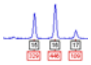
1.5 : 2.6 : 1

PowerPlex 16



4.1 : 4.5 : 1

SGM Plus



3.0 : 4.1 : 1

The Need for a More Uniform Approach to Mixture Interpretation

- "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

Comments from Some MIX05 Participants

- I'm anxious to see the paper with everyone's results, it should be interesting. I think it will show **we need some consistent guidelines for mixtures**.
- Thank you for letting us participate in this NIST study. **As always, it is a rewarding educational experience**.
- Thank you for allowing us to participate in the NIST studies. **We always find them very interesting, thought-provoking, and useful**. We look forward to seeing the results from this study and to participating in more studies in the future.

Concerns with Failure to Even Attempt Mixture Interpretation

- **A number of laboratories chose not to report anything in this study citing that without a suspect, mixtures are not examined**
- Why does a National DNA Database such as CODIS exist and how can it be helpful and reach its full potential if casework mixtures are not examined?
- Are cases containing mixtures being put aside because they are being viewed as "too difficult"? Are labs being too conservative and therefore not interpreting anything?
- Will improved training information and software tools aid in mixture interpretation (or will lab policies prevent examination of these cases no matter what tools are brought to bear on this problem)?

Additional Comments

- The data review is taking longer than initially anticipated but is proving very interesting and should be quite helpful to the community.
- All MIX05 participants will receive a certificate of participation and a summary of findings (prior to any publications).
- Please contact Margaret Kline (301-975-3134; margaret.kline@nist.gov) **if you are still interested in participating in MIX05**.

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

Acknowledgments

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

- Margaret Kline (sample prep, running study)
- John Butler (study design and data review)
- Becky Hill (GeneMapper*ID* data review)
- Jan Redman (Access database entry)
- Dave Duewer (*Virtual MixtureMaker* to aid sample selection)
- Chris Tomsey & Frank Krist (FMBIO Mac data)
- Kermit Channel & Mary Robnett (FMBIO NT data)

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

Disclaimers and Collaborations

Funding: Interagency Agreement 2003-IJ-R-029 between the National Institute of Justice and NIST Office of Law Enforcement Standards

Points of view are those of the authors and do not necessarily represent the official position or policies of the US 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.

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

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)