National Institute of Standards and Technology
. . .working with industry to develop and apply technology, measurements and standards

NIST Research Update

Margaret Kline

Peter M. Vallone, Michael D. Coble, Amy E. Decker, Janette W. Redman, David L. Duewer, John Butler

April 5, 2005

Present and Future Technological Advances in Human Identification

NIST Human Identity Project Team


John Butler
(Project Leader)


Pete Vallone


Margaret Kline


Jan Redman


Amy Decker


Mike Coble


Dave Duewer

Funding:
Interagency Agreement between National Institute of Justice and NIST Office of Law Enforcement Standards

Disclaimer

- This project was supported by NIJ **Grant Number 1999-IJ-R-A094 and 2003-IJ-R-029**, which is an interagency agreement between NIJ and the NIST Office of Law Enforcement Standards.
- Points of view in this document 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.

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

Current Areas of NIST Research Effort

- **Standard Information Resources** (STRBase information, training materials/review articles, validation standardization, calibration datasets)
- **Interlaboratory Studies** (Real-time PCR, mixture interpretation)
- **Resources for "Challenging Samples"** (miniSTRs for degraded DNA)
- **Information on New Loci** (Y-Chromosome, new STRs)

Standard Information Resources

STRBase, training materials, variant allele sequencing etc.

STRBase Updates
(since July 2004)

- Validation section
- miniSTR section
- Y-chromosome information (multiplexes & databases)
- Population data summary & OmniPop program download (courtesy of Brian Burritt)
- Reference Sequences for Commonly Used STR Markers

More minor additions

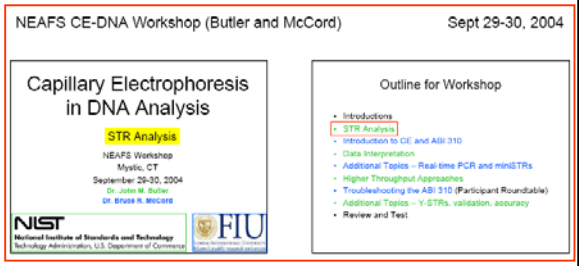
- Additional commercial STR kit schematics (Yfiler, PowerPlex Y)
- Published Promega primers (added PP16)
- Additional NIST publications/presentations (14 new talks, 12 new papers)
- Additional variant alleles & scientist addresses

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

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)


NEAFS CE-DNA Workshop (Butler and McCord) Sept 29-30, 2004

Capillary Electrophoresis in DNA Analysis

STR Analysis

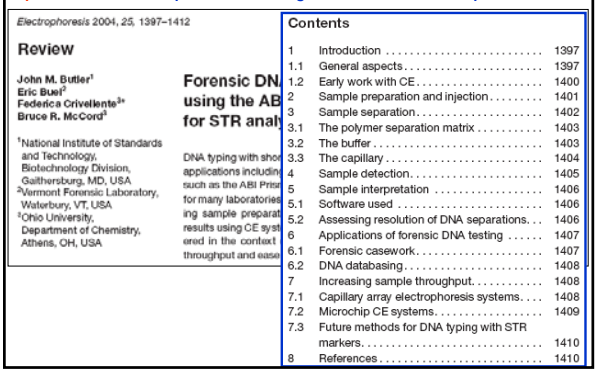
NEAFS Workshop
Mystic, CT
September 29-30, 2004
Dr. John M. Butler
Dr. Bruce R. McCord

Outline for Workshop

- Introductions
- STR Analysis
- Introduction to CE and ABI 310
- Data Interpretation
- Additional Topics – Real-time PCR and miniSTRs
- Higher Throughput Approaches
- Troubleshooting the ABI 310 (Participant Roundtable)
- Additional Topics – Y-STRs, validation, accuracy
- Review and Test

Review Article on STRs and CE

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


Electrophoresis 2004, *25*, 1397–1412

Review

John M. Butler[1]
Eric Buel[2]
Federica Crivellente[3]*
Bruce R. McCord[3]

[1]National Institute of Standards and Technology, Biotechnology Division, Gaithersburg, MD, USA
[2]Vermont Forensic Laboratory, Waterbury, VT, USA
[3]Ohio University, Department of Chemistry, Athens, OH, USA

Forensic DN using the AB for STR anal

DNA typing with shor applications includin such as the ABI Prism for many laboratories ing sample preparat results using CE syst ered in the context throughput and ease

Contents

Validation Standardization Efforts

Presentation at Promega meeting (October 2004)

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

We have been contacted by NFSTC with the desire to collaborate on validation standardization—they do not plan to produce validation kits now but rather have a workbook to help members of the community with validation…

Validation Standardization Questionnaire (conducted June-August 2004)

Can Validation be Standardized?

Statements from survey responders…

Over 86% (45/52) said yes

Those who responded "no" said

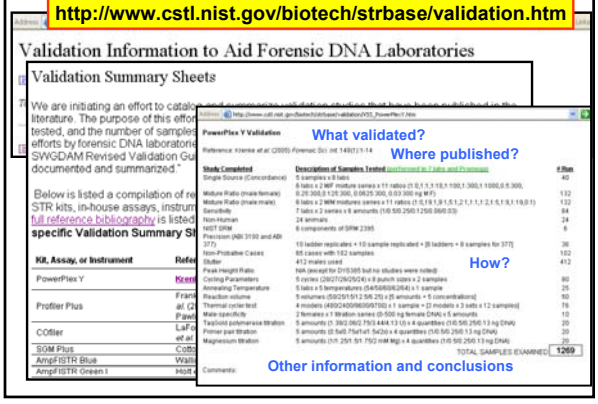
- "to some degree it can be, however, validation is specific to the platform, kits, …",
- "a start-up lab should do much more than an experienced lab…",
- "validation builds on previous work by lab or published data",
- "parts of it can be standardized; I don't think the non-probative cases could be", and
- "only in a general way, as with the SWGDAM guidelines. The uniqueness of each new procedure would make standardization difficult."

Our Conclusion…

to a certain extent it can…but everyone will always have a different comfort level…and **inflexible, absolute numbers for defined studies will not likely be widely accepted**

New Validation Homepage on STRBase

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

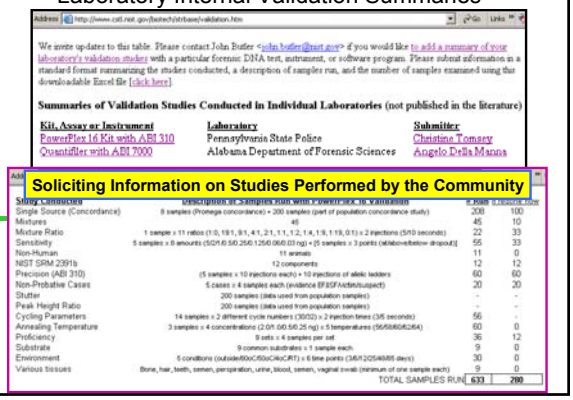


Validation Summary Sheet for PowerPlex Y

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
	TOTAL SAMPLES EXAMINED	1269

Krenke *et al.* (2005) *Forensic Sci. Int.* 148: 1-14

Laboratory Internal Validation Summaries



Goals of this Validation Standardization Project

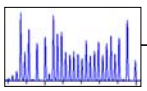
- To help the community gain a better understanding of the validation process and how others have implemented validation in their labs so that validation in one's own lab may be performed more quickly
- To help with establishing uniformity throughout the field to aid auditors in their inspections

NIST QA/QC Software

Tool being developed by Dave Duewer for STR Process Control

Tracks internal size standard in samples

Will be available soon for beta-testing; *still working on user's manual (and will need NCBI file conversion program to be more easily accessible)*

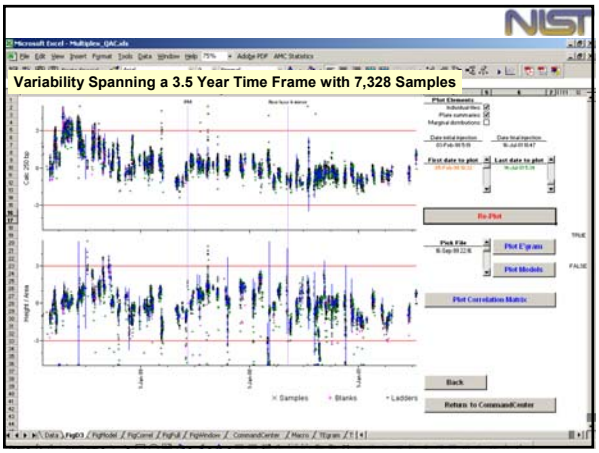


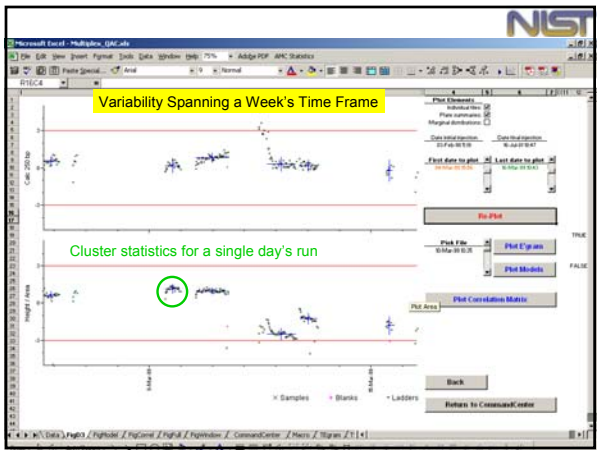
NCBI Program → Peak Height, Area, Size → NIST **Multiplex QAC** Software → Date vs Resolution (or Signal/Noise)

X | Y
103 | 436
104 | 569

We welcome data sets to help calibrate this software tool

This software does not perform genotyping.
It merely permits a view of analytical parameters over time.



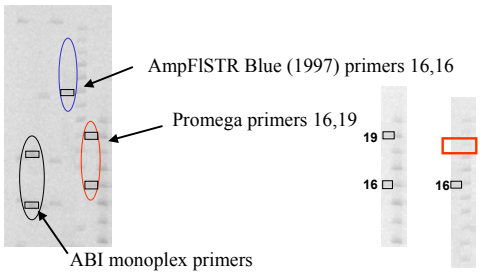


Variant Allele Sequencing

- Recent examples:
  - D18 null alleles
  - D18 large allele
  - DYS392 variant
- AAFS talk (Feb 26, 2005) by Margaret Kline on sequencing methods and applications
- We are happy to sequence unusual variant alleles for laboratories

vWA Allele Dropout Observed

What started our interest in sequencing variant alleles



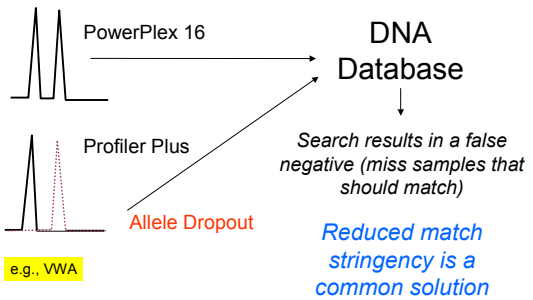
Kline, M.C., Jenkins, B. & Rodgers, S. (1998) Non-amplification of a vWA allele. *J Forensic Sci.*, 43(1), p250

Null Alleles

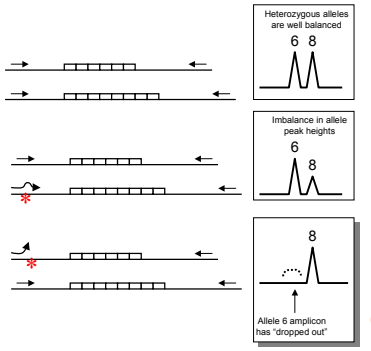
- Allele is present in the DNA sample but fails to be amplified due to a **nucleotide change in a primer binding site**
- Allele dropout is a problem because a heterozygous sample appears falsely as a homozygote
- Two PCR primer sets can yield different results on samples originating from the same source
- This phenomenon impacts DNA databases
- Large concordance studies are typically performed prior to use of new STR kits

For more information, see J.M. Butler (2005) *Forensic DNA Typing, 2nd Edition*, pp. 133-138

Concordance between STR primer sets is important for DNA databases

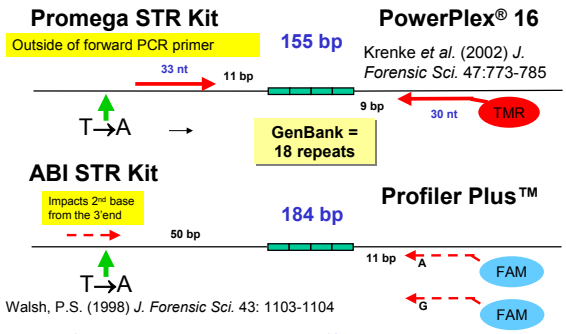


Impact of DNA Sequence Variation in the PCR Primer Binding Site



Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition*, Figure 6.9, ©Elsevier Academic Press

vWA Primer Position Comparisons



Walsh, P.S. (1998) *J. Forensic Sci.* 43: 1103-1104

In 2 out of 1,483 individuals tested = 0.067%

Lazaruk *et al.* (2001) *Forensic Sci Int.* 119:1-10

Apparent Null Alleles Observed During Concordance Studies

10/13 CODIS loci affected so far

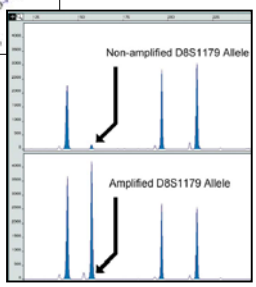
Locus	STR Kits/Assays Compared	Results	Reference
VWA	PP1.1 vs ProPlus	Loss of allele 19 with **ProPlus**; fine with PP1.1	Kline et al. (1998)
D5S818	PP16 vs ProPlus	Loss of alleles 10 and 11 with **PP16**; fine with ProPlus	Alves et al. (2003)
D13S317	Identifiler vs miniplexes	Shift of alleles 10 and 11 due to deletion outside of **miniplex assay**	Butler et al. (2003), Drabek et al. (2004)
D16S539	PP1.1 vs PP16 vs COfiler	Loss of alleles with **PP1.1**; fine with PP16 and COfiler	Nelson et al. (2002)
D8S1179	PP16 vs ProPlus	Loss of alleles 15, 16, 17, and 18 with **ProPlus**; fine with PP16	Budowle et al. (2001)
FGA	PP16 vs ProPlus	Loss of allele 22 with **ProPlus**; fine with PP16	Budowle and Sprecher (2001)
D18S51	SGM vs SGM Plus	Loss of alleles 17, 18, 19, and 20 with **SGM Plus**; fine with SGM	Clayton et al. (2004)
CSF1PO	PP16 vs COfiler	Loss of allele 14 with **COfiler**; fine with PP16	Budowle et al. (2001)
TH01	PP16 vs COfiler	Loss of allele 9 with **COfiler**; fine with PP16	Budowle et al. (2001)
D21S11	PP16 vs ProPlus	Loss of allele 32.2 with **PP16**; fine with ProPlus	Budowle et al. (2001)

From Table 6.2 in J.M. Butler (2005) *Forensic DNA Typing, 2nd Edition*, p. 136

Forensic Sci Int 2003;133:220-227

Identification of a D8S1179 primer binding site mutation and the validation of a primer designed to recover null alleles

Craig Leibelt[a,*], Bruce Budowle[b], Patrick Collins[a], Yasser Daoudi[a], Tamyra Moretti[a], Gary Nunn[a], Dennis Reeder[a], Rhonda Roby[a]

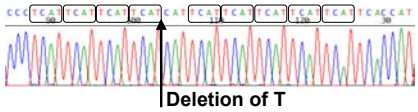


D8S1179-R

Extra D8S1179-R primer now present in Identifiler and Profiler Plus/*ID* kits

Microvariant "Off-Ladder" Alleles

- Defined as alleles that are not exact multiples of the basic repeat motif or sequence variants of the repeat motif or both
- Alleles with partial repeat units are designated by the number of full repeats and then a decimal point followed by the number of bases in the partial repeat
- Example: **TH01 9.3 allele**: $[TCAT]_4$ -CAT $[TCAT]_5$



Variation in the Flanking Region Can Cause Variant Alleles

D7S820 Example: commonly observed x .3 and x .1 alleles

Likely the result of a variation in the number of T's found in a poly(T) stretch 13 bases downstream of the core GATA repeat.

(Egyed, B. *et al. Forensic Sci. Int.* 2000, 113, 25-27).

http://www.cstl.nist.gov/biotech/strbase/var_d782.htm

Allele Designation	Allele Size	Instrument	Amp Kit*	Contributor	Verification/Conformation Method(s)	Notes	Frequency
12.1 [3]	281.65	ABI 310	PR	Kelly Duffy/R.Rubocki			
12.1 [4]	281.5	ABI 310	PR	Gintautas Svilpa	Observed both from suspect and crime scene stain		1
12.1 [5]	283.85	ABI 377	PS	Catherine Allor	Reamplified and Reanalyzed	Paternity samples only	1 in 11100
12.3	285.43	ABI 377	CO	Kelly Solis, Texas DPS	Re-extraction	Convicted offender	1 in 68000
13.1	286.8	ABI 310	PR, MP	Margaret Kline	Reamplified with two kits.		1 in 600 samples
13.1 [2]	287.58	ABI 377	CO, PS	Nicole Swinton	Re-extracted and Reamplified		

Allele Frequency at the time of reporting
12.3 frequency 1 in 68000
13.1 frequency 1 in 600

Variation in the Flanking Region Can Cause Variant Alleles

D7S820 Example: commonly observed x .3 and x .1 alleles

GATAGAACACTTGTCATAGTTTAGAACGAACTAACGATAGATAGATAGATAGATAGATAG
CTATCTTGTGAACAGTATCAAATCTTGCTTGATTGCTATCTATCTATCTATCTATCTATC

13 repeat units = $(GATA)_{13}$

8 T's TTTTTTTT x.3 → 12.3
10 T's TTTTTTTTTT x.1 → 13.1

9T's nominal "on ladder"

ATAGATAGATAGATAGATAGATAGATAGACAGATTGATAGTTTTTTTTTAATCTCACTAAA
TATCTATCTATCTATCTATCTATCTATCTGTCTAACTATCAAAAAAAAATTAGAGTGATTT

Variant Alleles Cataloged in STRBase

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

Off-Ladder Alleles

266 total variants reported as of 02/28/05

- CSF1PO (12)
- D2S1338 (3)
- D3S1358 (19)
- D5S818 (5)
- D7S820 (22)
- D8S1179 (4)
- D13S317 (10)
- D16S539 (11)
- D18S51 (30)
- D19S433 (11)
- D21S11 (23)
- FESFPS (1)
- FGA (68)
- HUMTH01 (7)
- Penta D (14)
- Penta E (13)
- TPOX (8)
- VWA (5)

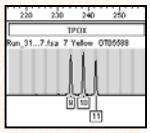
Currently **266** at 13/13 CODIS loci

Tri-Allelic Patterns

60 total patterns reported as of 03/01/05

- CSF1PO (2)
- D3S1358 (4)
- D5S818 (2)
- D7S820 (2)
- D8S1179 (5)
- D13S317 (3)
- D16S539 (1)
- D18S51 (6)
- D21S11 (4)
- FGA (9)
- HUMTH01 (1)
- TPOX (13)
- VWA (8)

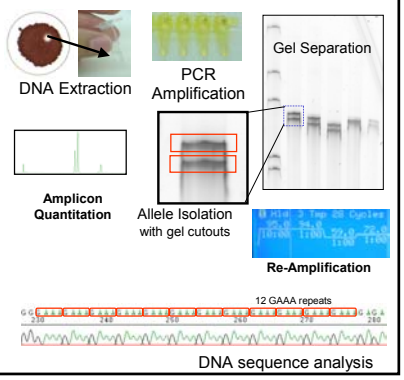
Currently **60** at 13/13 CODIS loci



Crouse, et al. report finding 18 TPOX three-banded patterns, and one CSF1PO three-banded pattern.

Steps in STR Allele Sequencing

- DNA Extraction
- Amplification with primers external to kit primers
- Gel Cutouts with Heterozygotes
- Re-Amplification
- Amplicon Quantitation
- ExoSAP
- Cycle Sequencing
- Dye Terminator Removal
- F/R Sequence Alignment to Reference Sequence



Steps in STR Allele Sequencing

DNA Extraction
Amplification with primers external to kit primers
Gel Cutouts with Heterozygotes
Re-Amplification
Amplicon Quantitation
ExoSAP
Cycle Sequencing
Dye Terminator Removal
F/R Sequence Alignment to Reference Sequence

PCR Amplification

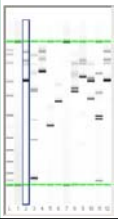
Sequencing primers outside commercial kit primers

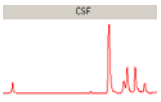
Steps in STR Allele Sequencing

DNA Extraction
Amplification with primers external to kit primers
Gel Cutouts with Heterozygotes
Re-Amplification
Amplicon Quantitation
ExoSAP
Cycle Sequencing
Dye Terminator Removal
F/R Sequence Alignment to Reference Sequence

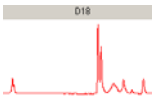
12 samples ~45 minutes

Agilent 2100 Bioanalyzer





CSF1PO
12,14
517 bp

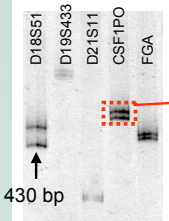


D18S51
15,22
431bp, 455bp

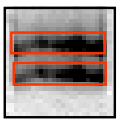
Steps in STR Allele Sequencing

DNA Extraction
Amplification with primers external to kit primers
Gel Cutouts with Heterozygotes
Re-Amplification
Amplicon Quantitation
ExoSAP
Cycle Sequencing
Dye Terminator Removal
F/R Sequence Alignment to Reference Sequence

Gel Separation



Allele Isolation with gel cutouts



Re-Amplification

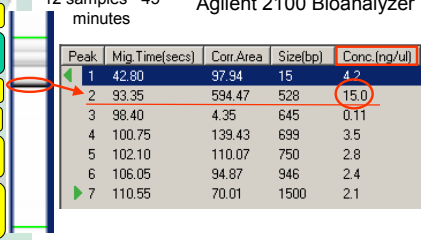
9%T 3%C
gel 32 cm
50 mMTris Formate in the Gel
Sodium Borate Running Buffer

Steps in STR Allele Sequencing

DNA Extraction
Amplification with primers external to kit primers
Gel Cutouts with Heterozygotes
Re-Amplification
Amplicon Quantitation
ExoSAP
Cycle Sequencing
Dye Terminator Removal
F/R Sequence Alignment to Reference Sequence

12 samples ~45 minutes

Agilent 2100 Bioanalyzer



Peak	Mig.Time(secs)	Corr.Area	Size(bp)	Conc.(ng/ul)
1	42.80	97.94	15	4.2
2	93.35	594.47	528	15.0
3	98.40	4.35	645	0.11
4	100.75	139.43	699	3.5
5	102.10	110.07	750	2.8
6	106.05	94.87	946	2.4
7	110.55	70.01	1500	2.1

Steps in STR Allele Sequencing

DNA Extraction
Amplification with primers external to kit primers
Gel Cutouts with Heterozygotes
Re-Amplification
Amplicon Quantitation
ExoSAP
Cycle Sequencing
Dye Terminator Removal
F/R Sequence Alignment to Reference Sequence

ExoSAP Treatment of PCR Products

Removes unconsumed dNTP's and primers

Target 7 ng of the PCR product for sequencing reactions

Cycle Sequencing (F primer only)	Cycle Sequencing (R primer only)

DYE Terminator removal

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

Forward & Reverse Sequence Alignment

EXAMPLES

(1) Deletions impacting miniSTRs with D13

(2) D18S51 point mutation causing null alleles in some Middle Eastern individuals

(3) Large D18S51 allele 40

(4) Small D18S51 allele “5.3”

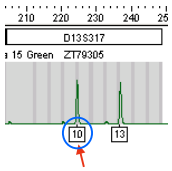
(5) Characterization of DYS635 allele 21.3 variant

Examination of Concordance:

African American sample ZT79305

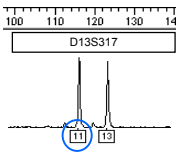
Drabek, J., Chung, D.T., Butler, J.M., McCord, B.R. (2004) Concordance study between miniplex STR assays and a commercial STR typing kit, *J. Forensic Sci.* 49(4): 859-860.

NIST Identifiler data



Really “11-1” allele

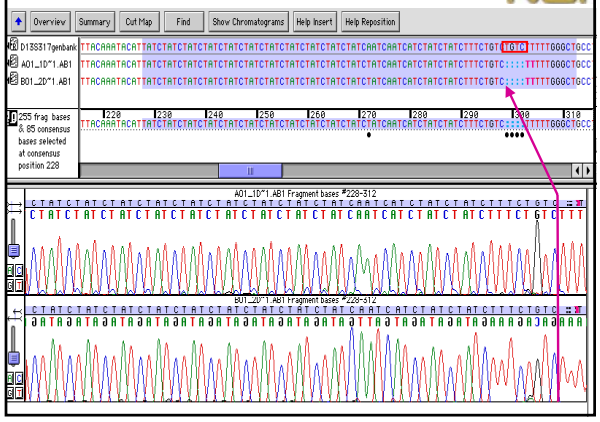
Ohio U miniSTR data



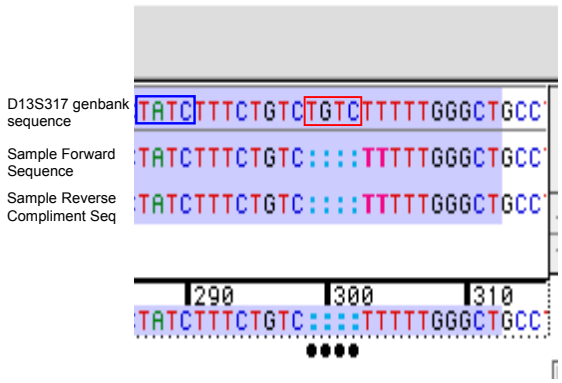
This problem has been seen multiple times by NYC OCME review of WTC BodePlex data

A deletion *outside* the miniSTR primers causes the commercial kit produced allele to appear one repeat smaller…

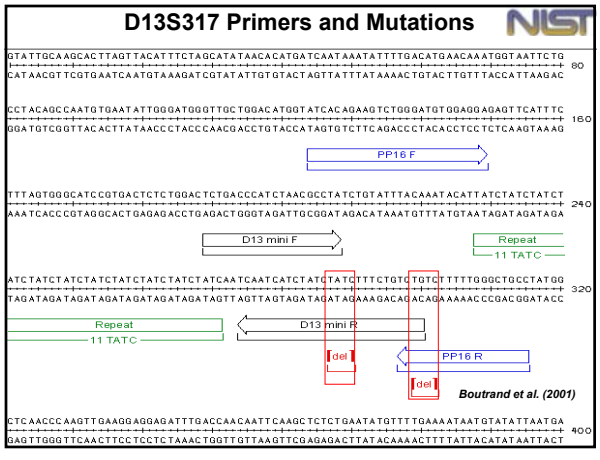
ZT79305: D13S317 TGTC deletion

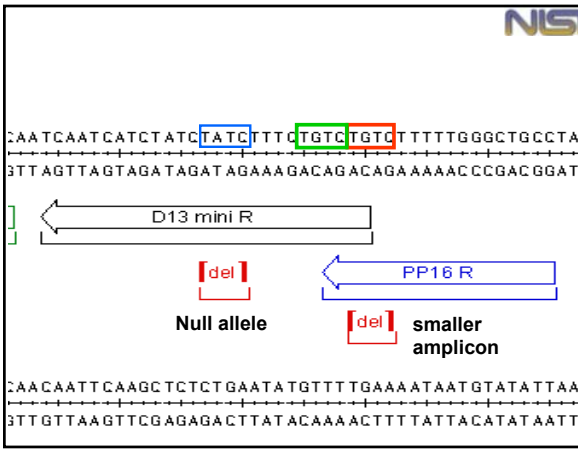


ZT79305: D13S317 TGTC deletion



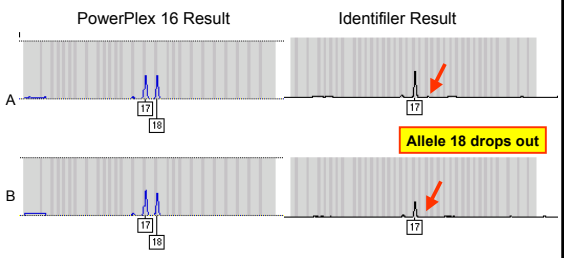
D13S317 Primers and Mutations





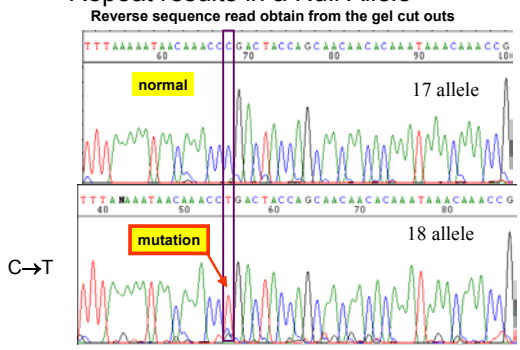
D18S51 Variant Allele

Kuwait Samples



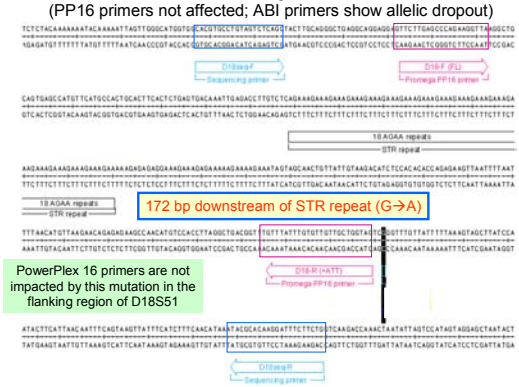
Clayton et al. (2004) **Primer binding site mutations affecting the typing of STR loci contained within the AMPFISTR SGM Plus kit.** *Forensic Sci Int.* 139(2-3): 255-259

Base change ~ 172 bases from D18S51 Repeat results in a Null Allele

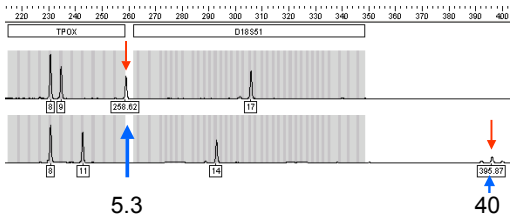


D18S51 Polymorphism

(PP16 primers not affected; ABI primers show allelic dropout)

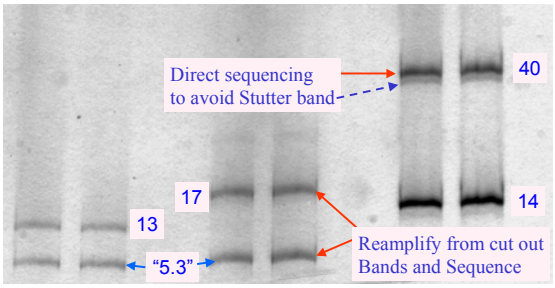


Extreme D18S51 Variants Range 258 – 396 bp

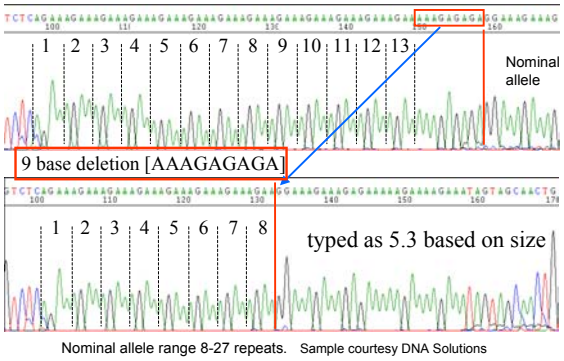


Samples amplified with Identifiler.Nominal range 8 - 28
Small D18 allele just outside of the TPOX allele range.
Large D18 is 46 bp larger than the D18 allele range.

Gel Separation Of D18 Variant Alleles



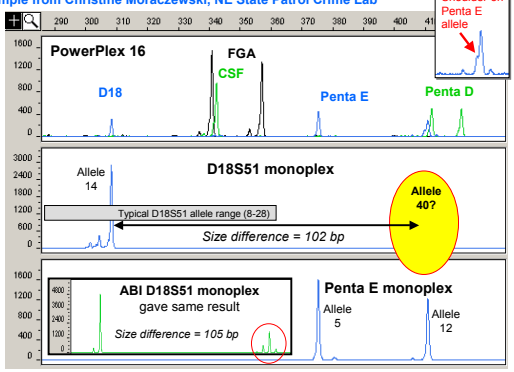
D18S51 sequencing Sample 188



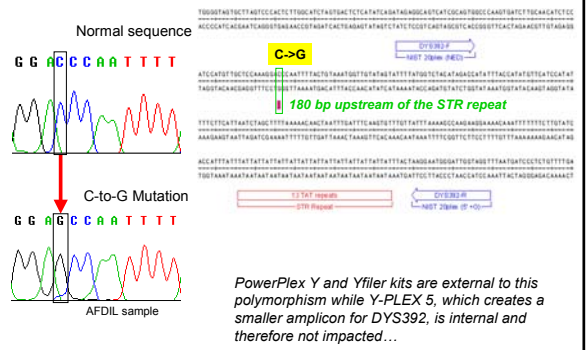
Nominal allele range 8-27 repeats. Sample courtesy DNA Solutions

Large D18S51 Allele Characterized at NIST

Sample from Christine Moraczewski, NE State Patrol Crime Lab



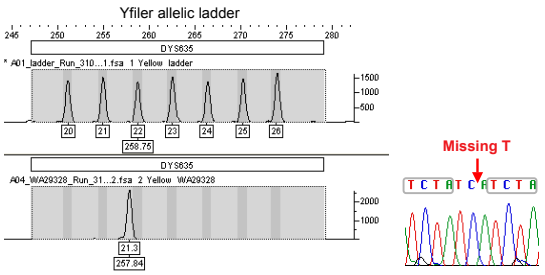
DYS392 Flanking Region Polymorphism



PowerPlex Y and Yfiler kits are external to this polymorphism while Y-PLEX 5, which creates a smaller amplicon for DYS392, is internal and therefore not impacted…

DYS635 Variant Allele 21.3

(Y-GATA-C4)

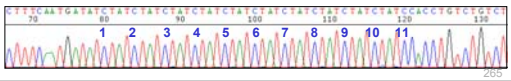


$[TCTA]_4(TGTA)_2[TCTA]_2(TGTA)_2[TCTA]_2(TGTA)_2\,[TCTA]_5\,TC\text{-}A\,[TCTA]_2$

Analysis of Common STR Variant Alleles

- We have monoplex primers for all common STR loci and kits
- We have sequencing primers that bind outside of STR kit primer sequence positions to enable view of polymorphic nucleotides that cause primer binding site mutations
- NIJ has funded us to characterize STR variants for the forensic DNA community

D16S539 (bottom strand)



Interlaboratory Studies

DNA Quantitation (2004),
Mixture Interpretation (2005)

NIST Quantitation Study 2004 (QS04)

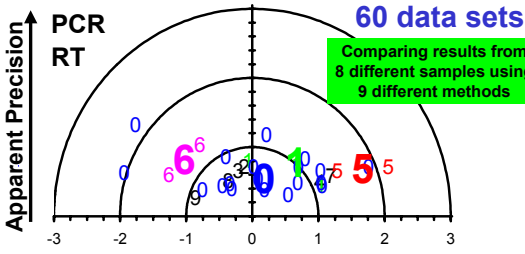
Consisted of:
- 8 DNA extracts labeled A – H
- Shipped Dec 2003 –Jan 2004 to 84 laboratories for quantification; data received back by April 2004
- Labs were requested to use multiple methods / multiple analysts

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)

Interlaboratory Comparisons

Laboratory Performances with Real-Time PCR Methods



60 data sets

Comparing results from 8 different samples using 9 different methods

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

Kline, *et al., J. Forensic Sci., in press (May 2005)*

Table 2. The percent success rate reported for a sample.

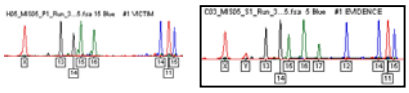
		% Quantitative Results*							
Target [DNA] ng/μL		1.5	0.5	0.5	0.16	0.16	0.05	0.05	0.05
Method	N_{total}	A	B	E	C	F	D	G	H
Quantifiler	37	100	100	100	100	100	100	100	100
Other RT-PCR	23	100	100	100	100	100	100	100	100
"ACES"	14	100	100	100	100	100	100	100	100
AluQuant	13	100	100	100	100	100	100	100	100
PicoGreen	12	100	100	92	100	100	92	83	83
ECL	75	100	99	99	93	95	84	77	87
TMB	98	100	100	99	93	94	59	62	63
Yield gel	14	57	0	0	0	0	0	0	0
	286								

a Quantitative results are those that were reported as values, values reported as the range between contiguous calibration standards, values reported as less-than the lowest calibration standard if smaller than the target [DNA], or values reported as greater-than the highest calibration standard if larger than the target[DNA].

Kline, *et al., J. Forensic Sci., in press (May 2005)*

Mixture Interpretation Interlab Study (MIX05)

- **Only involves interpretation of data**
- As of early March, ~97 labs are enrolled for participation (22 from overseas)
- Four mock cases supplied with "victim" and "evidence" electropherograms (GeneScan .fsa files – that can be converted for Mac or GeneMapper; gel files MAC & NT 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

Plans for Dissemination of MIX05 Results

- Data shipped & on STRbase in mid-January 2005 through February
- Responses were due on March 15, 2005. **BUT we are still taking results!**
- Goal is to understand the "lay of the land" regarding mixture analysis across the DNA typing community
- Results will be discussed at NIJ DNA Grantees Meeting (June 2005), SWGDAM (June 2005), and ISFG (Sept 2005)
- We plan to develop training materials to aid in mixture interpretation with available software tools and to help in standardizing reports involving mixture analysis

Resources for Challenging Samples

Degraded DNA and Mixtures

Degraded DNA work

- ENFSI study participation
  - compared STRs, miniSTRs, and autosomal SNPs on same set of degraded DNA samples provided by Peter Gill
- miniSTR website
  - http://www.cstl.nist.gov/biotech/strbase/miniSTR.htm
- New miniSTR loci published
  - http://www.cstl.nist.gov/biotech/strbase/pub_pres/Coble2005miniSTR.pdf
- SNP markers and assays
  - http://www.cstl.nist.gov/biotech/strbase/SNP.htm
- Performance of miniSTRs on shed hairs
  - Mike Coble spoke at AAFS (Feb 25, 2005)

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., in press*. (January 2005 issue)

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

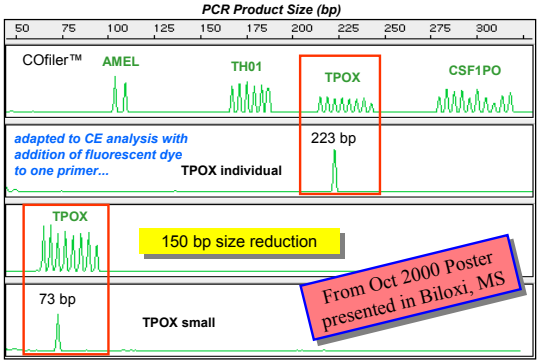
STR Size Reduction

Through Moving Primer Positions Closer to Repeat

Primer positions define PCR product size
Repeat information is independent of amplicon size

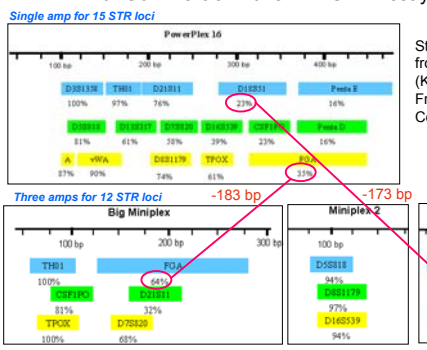
Advantages of Approach:
Size reduction enhances success rate with degraded DNA
Retains same marker information (database compatibility)
Uses highly polymorphic STR loci (high discriminatory power)

Development of miniSTRs: Past Work



New primer sets are intended to aid with typing degraded DNA samples as well as future microchip CE and mass spectrometry applications…

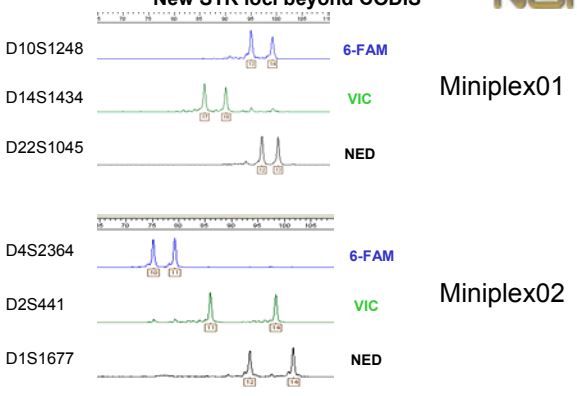
Comparison of PCR Amplification Success Rates with Commercial Kit vs. miniSTR Assays



Study with 31 bones from the "Body Farm" (Knoxville, TN) and Franklin County Coroner's Office (OH)

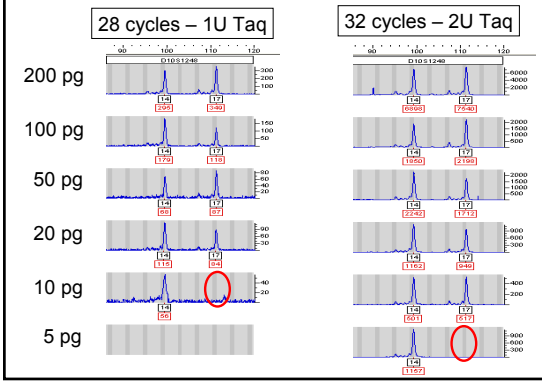
Chung, et al., The application of miniplex primer sets in the DNA profiling of human skeletal remains, submitted

New STR loci beyond CODIS



Coble & Butler (2005) JFS 50:43-53

miniSTR Assay Sensitivity (D10S1248)



Information on New Loci

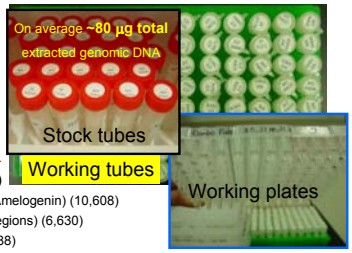
Autosomal SNPs, Y-Chromosome

NIST U.S. Population Samples

As of 06/2003 663 males (anonymous; self-identified ethnicities)

260 Caucasians
260 African Americans
140 Hispanics
3 Asians

Whole blood received from Interstate Blood Bank (Memphis, TN)



Working tubes/plates 1 ng/uL
To date: (~85,000 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)
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)
mtDNA full control region sequences by AFDIL

Samples supplied to **OhioU** for miniSTR typing and **AFDIL** for whole mtGenome sequencing

Standard U.S. Population Dataset

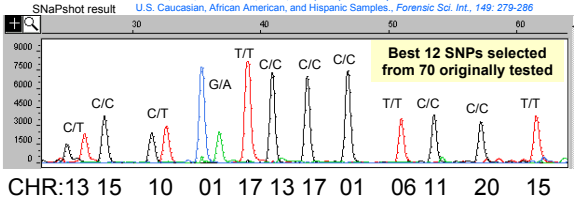
http://www.cstl.nist.gov/biotech/strbase/NISTpop.htm

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

Genetic Markers	Loci Examined	Publications
Common STRs	**D2S1338 and D19S433 information has been provided to the FBI for inclusion in PopStats to aid statistical calculations**	Butler et al. (2003) JFS
miniSTRs		Drabek et al. (2004) JFS
New autosomal STRs		Coble et al. (2005) JFS
Autosomal SNPs	**70 C/T SNPs** (Orchid panel)	Vallone et al. (2004) FSI
Common Y-STRs	22 loci (27 regions)	Schoske et al. (2004) FSI
	Yfiler concordance study	***Data in ABI Yfiler database***
New Y-STRs	**27 additional loci**	**Butler et al., in press FSI**
Y-SNPs	50 loci spanning haplogroups A-R	Vallone et al. (2004) JFS
mtDNA	LINEAR ARRAY and coding mtSNPs	Kline et al. (2005) JFS
	Full control regions by **AFDIL**	***inclusion in EMPOP***

NIST Autosomal 12plex SNP Assay

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*



CHR:13 15 10 01 17 13 17 01 06 11 20 15

12plex PCR followed by 12-plex ASPE
Fragments separated on a ABI 3100 in 35 minutes
A Genotyper macro has been developed to type data
The 12plex assay has been run on over 600 samples
Works well on 1-2 ng of template
Sensitivity studies are underway along with degraded DNA

New Y-STR Loci, Issues, and Assays

- Updates on Y-chromosome information
  - http://www.cstl.nist.gov/biotech/strbase/y_strs.htm
- Testing on 27 new Y-STR loci
  - Butler, J.M., Decker, A.E., Vallone, P.M., Kline, M.C. (2005) Allele frequencies for 27 Y-STR loci with U.S. Caucasian, African American, and Hispanic samples, *in press FSI*
- Chromosomal duplication issues
  - Butler, J.M., Decker, A.E., Kline, M.C., Vallone, P.M. (2005) Chromosomal duplications along the Y-chromosome and their potential impact on Y-STR interpretation, *in press JFS*

Y-Chromosome Standard NIST SRM 2395

www.cstl.nist.gov/biotech/strbase/srm2395.htm

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

$318

Certified for all loci in commercial Y-STR kits:

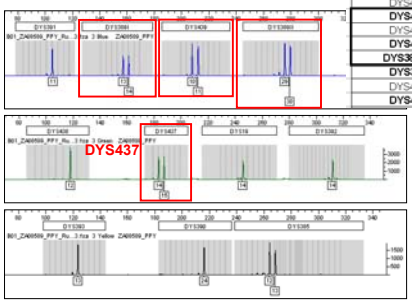
Y-PLEX 6
Y-PLEX 5
Y-PLEX 12
PowerPlex Y

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

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

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

Duplication at Multiple Loci with Single-Source Sample



Y-chromosome mapping

q-arm

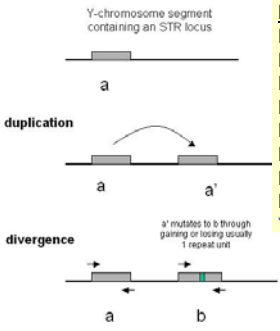
Y STR Marker	Position (Mb)
DYS391	13.413
DYS635 (C4)	13.690
DYS434	13.777
DYS437	13.778
DYS435	13.807
DYS439	13.826
DYS389 I/II	13.923
DYS388	14.057
DYS442	14.071
DYS438	14.248

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), *in press JFS*

Duplication and Divergence Model



Locus	# dup*	>1 repeat
DYS19	23	2
DYS389I	5	0
DYS389II	9	2
DYS390	1	0
DYS391	3	1
DYS392	0	0
DYS393	3	0
DYS385a/b	17	0

*from www.yhrd.org, literature, and our work

92% have single repeat difference

Since single-step mutations are most common, then single repeat spacing in duplicated alleles is expected

Butler et al. (2005), *in press JFS*

Our Recent Y-Chromosome Work

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

- Schoske, R., Vallone, P.M., Kline, M.C., Redman, J.W., Butler, J.M. (2004) High-throughput Y-STR typing of U.S. populations with 27 regions of the Y chromosome using two multiplex PCR assays, *Forensic Sci. Int.* 139: 107-121.
- Vallone, P.M. and Butler, J.M. (2004) Multiplexed assays for evaluation of Y-SNP markers in U.S. populations. *Progress in Forensic Genetics 10*, Elsevier Science: Amsterdam, The Netherlands, International Congress Series 1261, 85-87.
- Butler, J.M. and Schoske, R. (2004) Forensic value of the multi-copy Y-STR marker DYS464. *Progress in Forensic Genetics 10*, Elsevier Science: Amsterdam, The Netherlands, International Congress Series 1261, 278-280.
- Butler, J.M. and Schoske, R. (2004) Duplication of DYS19 flanking regions in other parts of the Y chromosome. *Int. J. Legal Med.*, 118: 178-183.
- Vallone, P.M. and Butler, J.M. (2004) Y-SNP typing of U.S. African American and Caucasian samples using allele-specific hybridization and primer extension. *J. Forensic Sci.* 49(4): 723-732.
- Butler, J.M., Decker, A.E., Kline, M.C., Vallone, P.M. (2005) Chromosomal duplications along the Y-chromosome and their potential impact on Y-STR interpretation, *J. Forensic Sci., in press.*
- Butler, J.M., Decker, A.E., Vallone, P.M., Kline, M.C. (2005) Allele Frequencies for 27 Y-STR Loci with U.S. Caucasian, African American, and Hispanic Samples, *Forensic Sci. Int., in press.*

Mitochondrial DNA Work

- Evaluation of Roche LINEAR ARRAY screening assay

Kline et al. (2005) JFS 50: 377-385

- Comparison of LINEAR ARRAY resolution to control region sequencing performed by AFDIL
- Collaboration with AFDIL for developing coding SNP assays using SNaPshot

Coble, M.D., Just, R.S., O'Callaghan, J.E., Letmanyi, I.H., Peterson, C.T., Irwin, J.A., Parsons, T.J. (2004) Single nucleotide polymorphisms over the entire mtDNA genome that increase the power of forensic testing in Caucasians. *Int. J. Legal Med.*, 118: 137-146.

Vallone, P.M., Just, R.S., Coble, M.D., Butler, J.M., Parsons, T.J. (2004) A multiplex allele-specific primer extension assay for forensically informative SNPs distributed throughout the mitochondrial genome. *Int. J. Legal Med.*, 118: 147-157.

Mito "Strips"

- Roche Applied Science (Indianapolis, IN) recently released a mtDNA typing kit
- LINEAR ARRAY Mitochondrial DNA HVI/HVII Region-Sequence Typing Kit
- Cat. No. 03 527 867 001
- Cost $1500 for 50 reactions
- NIST was involved in beta-testing and performed a population study with these LINEAR ARRAYs

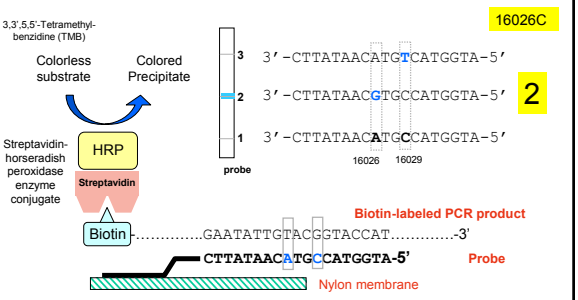
Previous Publications on mtDNA Typing Assays with SSO Probes (dot blot, reverse dot blot, linear arrays)

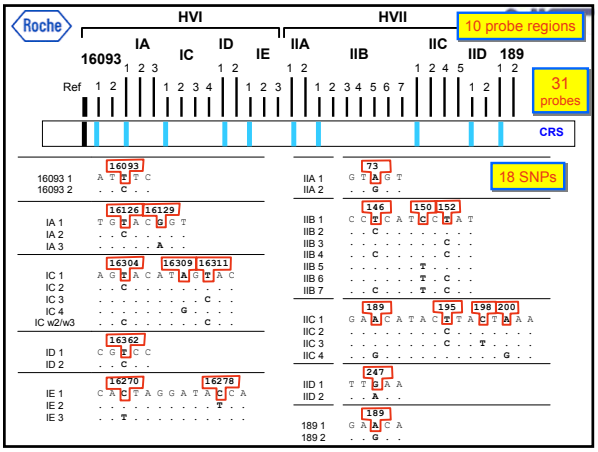
- Stoneking *et al.* (1991) Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes. *Am. J. Hum. Genet.* 48:370-382
- Skowasch, K., *et al.* (1994) Development of PCR-based reverse dot-blot typing system for the control region of mtDNA. *Proceedings of the Fifth International Symposium on Human Identification*, Madison, WI: Promega, p. 127.
- Comas, D., *et al.* (1999) Eur. J. Hum. Genet. 7:459-468
- Calloway, C.D., *et al.* (2000) *Am. J. Hum. Genet.* 66:1384-1397
- Reynolds, R., *et al.* (2000) *J. Forensic Sci.* 45(6):1210-1231
- Gabriel, M.N., *et al.* (2001) *Croatian Medical Journal* 42(3):328-335
- Gabriel, M.N., et al. (2003) *Croatian Medical Journal* 44(3):293-298
- Calloway, C., et al. (2003) Validation of the LINEAR ARRAY Mitochondrial DNA HVI/HVII Region-Sequence Typing kit. *Proceedings of the 14th International Symposium on Human Identification*.
- Calloway, C., et al. (2003) Applications of the LINEAR ARRAY Mitochondrial DNA HVI/HVII Region-Sequence Typing kit. *Proceedings of the 14th International Symposium on Human Identification*.
- Kline, M.C., et al. (2003) Semi-automation of mtDNA arrays: results from 666 population samples and comparisons. *Proceedings of the 14th International Symposium on Human Identification*.

Terry Melton population studies…

Detection Chemistry

Saiki, R.K., *et al.* (1989) Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes. *Proc. Natl. Acad. Sci. USA* 86:6230-6234





Lessons Learned with LINEAR ARRAYS

- pH is important to wash solutions
  - If above protocol's pH 7.4, then the blue color will not develop correctly and signal will be lost
- Quantification value obtained on Agilent is not equivalent to yield gel information provided by Roche
  - May need to make adjustments—do a sensitivity titration in your lab

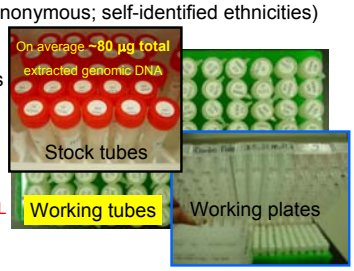
NIST U.S. Population Samples

As of 06/2003 666 males (anonymous; self-identified ethnicities)

286 Caucasians
252 African Americans
128 Hispanics

Whole blood received from Interstate Blood Bank (Memphis, TN)

Working tubes/plates 1 ng/uL



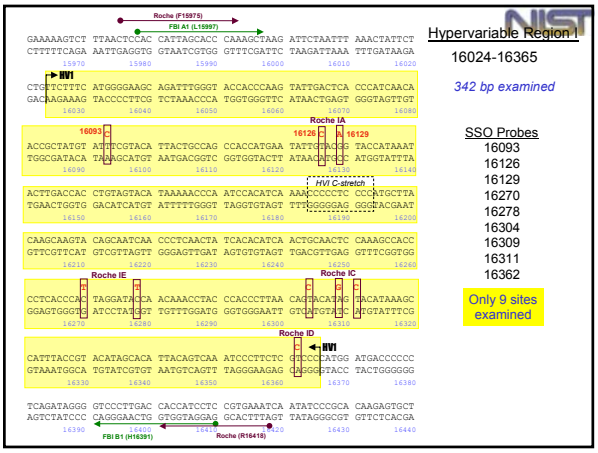
Samples supplied to **OhioU** for miniSTR typing and **AFDIL** for whole mtGenome sequencing

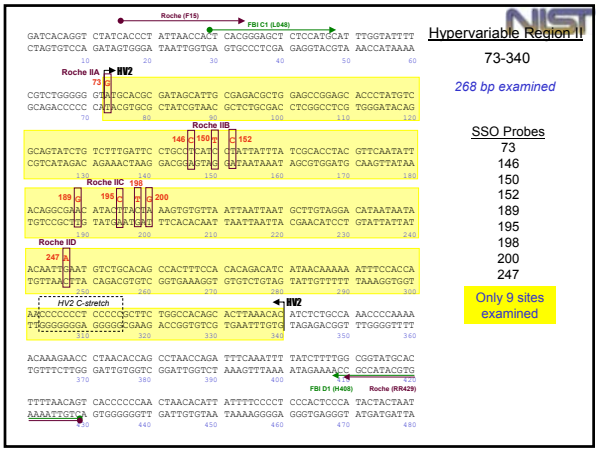
PCR Amplification

- Protocol calls for 5 pg input DNA (based on nuclear DNA measurement)
- We used 1 ng DNA instead and reduced PCR cycle number from 34 (protocol) to 28
- Thermal cycling on GeneAmp 9700:
  - 94 ºC for 14 minutes
  - **28 cycles**: 92 ºC for 15s, 59 ºC for 30s, and 72 ° C for 30s
  - 72 ºC for 10 minutes
  - hold at 10 ºC
- 50 uL PCR volume (protocol) with duplex amplification

HVI primers amplify a **444 bp PCR product**:
Forward (L15975-15993) 5'-biotin-CTCCACCATTAGCACCCAA-3'
Reverse (H16418-16401) 5'-biotin-ATTTCACGGAGGATGGTG-3'

HVII primers amplify a **416 bp PCR product**:
Forward (L15-34) 5'-biotin-CACCCTATTAACCACTCACG-3'
Reverse (H429-410) 5'-biotin-CTGTTAAAAGTGCATACCGC-3'





Hypervariable Region II

73-340

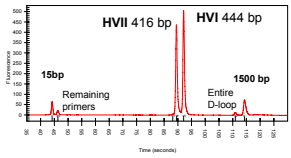
268 bp examined

SSO Probes

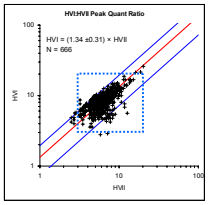
73
146
150
152
189
195
198
200
247

Only 9 sites examined

Analysis of HVI/HVII PCR Products



Agilent 2100 Bioanalyzer sized and quantified HVI/HVII products



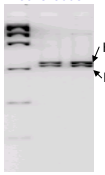
Usually 2-20 ng/uL obtained for each amplicon

Aimed for 50 ng on each mito strip

12 samples can be run in ~45 minutes

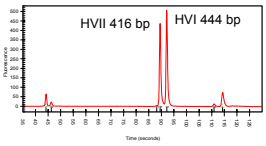
Gel Results

Roche ladder

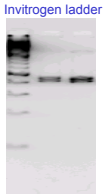


HVI 444 bp

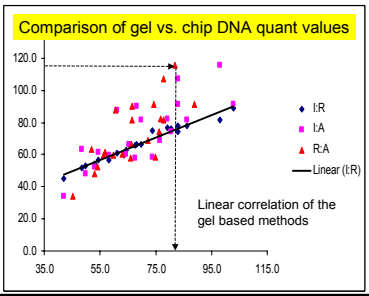
HVII 416 bp



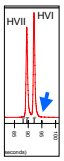
Invitrogen ladder



Comparison of gel vs. chip DNA quant values

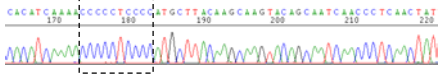


Analysis of HVI/HVII PCR Products

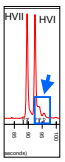


16189T

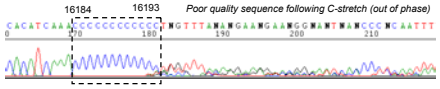
Good quality sequence



The Agilent electropherogram also gives an indication of the HVI "C-stretch" by the presence of extra peaks (see Butler *et. al.* (1998) *Electrophoresis* 19, 119-124).



16184 16193 *Poor quality sequence following C-stretch (out of phase)*



HV1 C-stretch

Automated Washing & Color Development

Tecan Profiblot – processes sample through wash steps

24 strips per run
~2 hours per run
2 or 3 runs easily performed per day

Step	File	Time	Temp	Solution
1	Temp		55 ºC	
2	Disp		55 ºC	Wash
3	Pause		55 ºC	
4	Inc	15 min	55 ºC	
5	Asp		55 ºC	
6	Disp		55 ºC	Wash
7	Asp		55 ºC	
8	Disp		55 ºC	SA-HRP Conjugate
9	Inc	5 min	55 ºC	
10	Asp		55 ºC	
11	Disp		55 ºC	Wash
12	Asp		55 ºC	
13	Disp		55 ºC	Wash
14	Inc	12 min	55 ºC	
15	Asp		55 ºC	

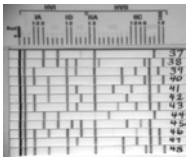
16	Cool		25 ºC	
17	Disp		25 ºC	Wash
18	Asp		25 ºC	
19	Disp		25 ºC	Citrate
20	Inc	5 min	25 ºC	
21	Asp		25 ºC	
22	Disp		25 ºC	Color Dev
23	Inc	15 min	25 ºC	
24	Asp		25 ºC	
25	Disp		25 ºC	DI Water
26	Asp		25 ºC	
27	Disp		25 ºC	DI Water
28	Inc	5 min	25 ºC	
29	Asp		25 ºC	
30	Disp		25 ºC	DI Water
31	End		25 ºC	

Digital Recording of LINEAR ARRAYS

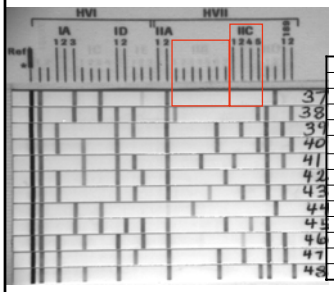
UV epi-fluorescence

GeneGnome (Syngene)

Yellow light used with filters so blue color lines could be archived as black and white images



Data Interpretation for LINEAR ARRAYs



Analysis of probe results is still manual!

	HVI					HVII			
16093	A	C	D	E	A	B	C	D	189
1	3	3	0	w2	2	0	0	2	0
1	3	3	1	1	2	1	5	1	2
1	1	0	1	1	2	6	2	1	1
1	1	1	1	0	2	6	5	1	2
1	0	1	1	1	2	4	1	1	1
1	1	1	1	2	2	7	0	1	0
1	1	0	1	1	2	5	2	1	1
1	3	3	0	0	w2	3	0	2	0
1	3	3	2	2	2	7	4	1	0
1	1	1	1	2	2	0	w2	1	1
1	1	0	1	1	2	6	2	1	1
1	1	1	1	0	2	6	5	1	2

Typing results from 50 ng of each PCR product

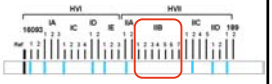
"Blank" Calls on LINEAR ARRAYs

We observed 640 "blanks" (**9.6% of calls**) on 346 different individuals (**52% of samples typed**).

*Different individuals typing as a blank for the same probe region could have different substitutions but for the purposes of data analysis the blanks are considered to represent the same variant (see Melton *et al.* (2001) *J. Forensic Sci.* 46(1):46-52)

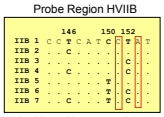
SSO Probe Region	Number Observed	Frequency	Budowle et al. (1999) Cau, AA, His
16093	23	3.5%	
HVIA	33	5.0%	3, 9, 3%
HVIC	76	11.4%	3, 20, 10%
HVID	33	5.0%	7, 17, 4%
HVIE	60	9.0%	
HVIIA	3	0.5%	0, 0, 0%
HVIIB	96	14.4%	16, 70, 55%
HVIIC	122	18.3%	11, 47, 13%
HVIID	42	6.3%	5, 5, 18%
189	152	22.8%	

Blanks expected based on full sequence analysis of 1393 individuals



PCR product fails to hybridize to any probe in region due to additional polymorphisms in the probe region that prevent hybridization

Probe Region HVIIB



Nucleotide positions 151 and 153 are common variants in African Americans

Typing frequencies for 666 NIST population samples

#*	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

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

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

HV1/HV2 Anderson (CRS) except 263G, 315.1C
Haplogroup H

51

All Roche mtDNA LINEAR ARRAYS: 1111111111

site	rCRS	1	1	1	1	2	2	3	4	4	5	12	15
3010	G	**A**	G	G	G	**A**	G	G	G	G	G	G	**A**
4793	A	A	A	A	A	A	A	A	A	**G**	A	A	A
10211	C	C	C	C	C	C	C	C	C	C	C	C	C
5004	T	T	**C**	T	T	T	T	T	T	T	T	T	T
7028	C	C	C	C	**T**	C	**T**	C	T	C	C	C	C
7202	A	A	A	A	A	A	A	A	A	A	A	A	A
16519	T	T	T	**C**	T	**C**	**C**	T	T	**C**	**C**	**C**	**C**
12858	C	C	C	C	C	C	C	C	C	C	C	C	C
4580	G	G	G	G	G	G	**A**	G	**A**	G	G	G	G
477	T	T	T	T	T	**C**	T	T	T	T	T	T	T
14470	T	T	T	**A**	T	T	T	T	T	T	T	T	T

Improved resolution of mtDNA samples with coding region SNPs

12 types were observed
4 types were unique

Comparison of Other U.S. Population Data with SSO Probes

	Population	N	#types	diversity	Most Common Type	MCT frequency
11	Caucasian	922	226	0.964	11111111	15.4%
10	African Am	805	251	0.983	12112021	6.8%
7	Hispanic	555	170	0.963	12122011	11.7%
	Total	2282	502	0.998	11111111	**7.2%**

8 regions, 21 probes, 13 SNPs

Melton *et al.* (2001) *J. Forensic Sci.* 46(1): 46-52

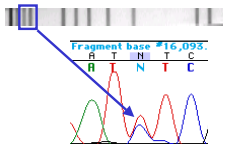
Population	N	#types	diversity	Most Common Type	MCT frequency
Caucasian	286	116	0.960	1111111111	16.4%
African Am	252	129	0.977	1141224211	10.7%
Hispanic	128	74	0.954	1102120111	16.4%
Total	666	282	0.985	1111111111	**7.7%**

10 regions, 31 probes, 18 SNPs

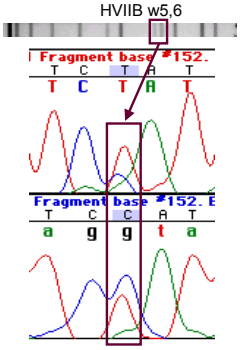
Kline *et al.* (2003) NIST population study

Heteroplasmy Detection

Accurate Detection of Heteroplasmy at position **16093**



Position 152
HVIIB w5,6

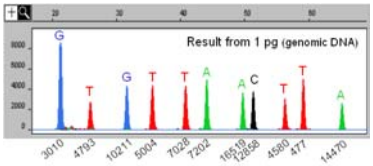


Observed 7 Times in 666 samples

AA	Cauc
HVIIC 2/4 (n=1) 152	HVID 1/2 (n=1) 16362
16093 1/2 (n=1)	HVIIB 3/4 (n=2) 146
	HVIIB 5/6 (n=1) 152
	189 1/2 (n=1)

Sites Observed: 16093, 16362, 146, 152, 189

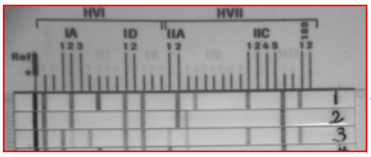
NIST mtDNA Work



Coding Region
mtSNP 11plex
(minisequencing assay)

Developed with AFDIL to resolve mtDNA most common types

Vallone et al. (2004) Int. J. Legal Med., 118: 147- 157



Roche Linear Arrays
(probes for HVI/HVII)

Kline et al.(2005) JFS 50:377-385

Beta-test/Population Study

Acknowledgements

margaret.kline@nist.gov

Funding:

Interagency Agreement between National Institute of Justice and NIST Office of Law Enforcement Standards

NIST Project Team:

John Butler Pete Vallone
Margaret Kline Jan Redman
Amy Decker Mike Coble
Dave Duewer

This presentation available as pdf file from
http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm