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

NIST Research Update

John M. Butler

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

January 18, 2005

SWGDAM

NIST Human Identity Project Team


John Butler (Project Leader)


Pete Vallone


Margaret Kline


Jan Redman

Former Team Member


Rich Schoske


Amy Decker


Mike Coble


Dave Duewer

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

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, 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)

More minor additions

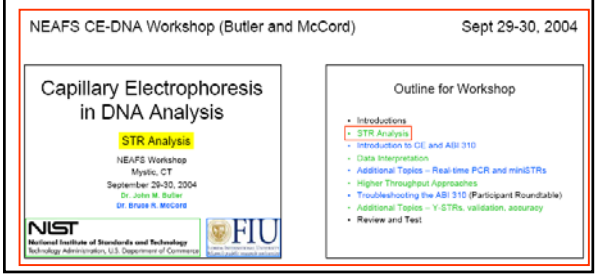
- Additional commercial STR kit schematics (Yfiler, PowerPlex Y)
- Published Promega primers (added PP16)
- Additional NIST publications/presentations (8 new talks, 9 new papers)
- Additional variant alleles & scientist addresses

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)



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

1	Introduction	1397
1.1	General aspects	1397
1.2	Early work with CE	1400
2	Sample preparation and injection	1401
3	Sample separation	1402
3.1	The polymer separation matrix	1403
3.2	The buffer	1403
3.3	The capillary	1404
4	Sample detection	1405
5	Sample interpretation	1406
5.1	Software used	1406
5.2	Assessing resolution of DNA separations	1406
6	Applications of forensic DNA testing	1407
6.1	Forensic casework	1407
6.2	DNA databasing	1408
7	Increasing sample throughput	1408
7.1	Capillary array electrophoresis systems	1408
7.2	Microchip CE systems	1409
7.3	Future methods for DNA typing with STR markers	1410
8	References	1410

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

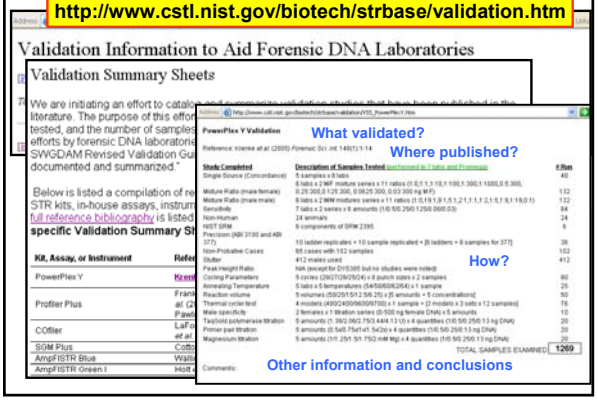
- "to some degree it can be, however, validation is specific to the platform, kits, …",
- "a start-up lab should do much more than an experienced lab…",
- "validation builds on previous work by lab or published data",
- "parts of it can be standardized; I don't think the non-probative cases could be", 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

We invite updates to this table. Please contact John Butler <john.butler@nist.gov> if you would like to add a summary of your lab-ratory's validation studies with a particular forensic DNA test, instrument, or software program. Please submit information in a standard format summarizing the studies conducted, a description of samples run, and the number of samples examined using the downloadable Excel file [click here].

Summaries of Validation Studies Conducted in Individual Laboratories (not published in the literature)

Kit, Assay or Instrument	Laboratory	Submitter
PowerPlex 16 Kit with ABI 310	Pennsylvania State Police	Christine Tomasy
Quantifiler with ABI 7000	Alabama Department of Forensic Sciences	Angelo Della Manna

Soliciting Information on Studies Performed by the Community

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

Expert System Calibration Data Set

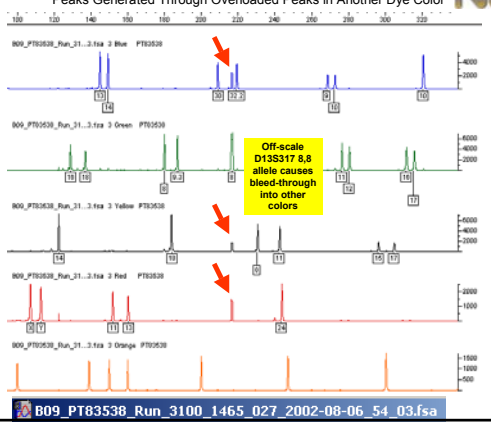
"Electronic SRM" to help meet NDIS Appendix B requirements

These data will be able to be used to check software upgrades to ensure reliable performance of the Expert System software

To Help Meet NDIS Appendix B Requirements for Evaluating Expert Systems…

- **200 calibration samples needed**
- Types of challenges (at least 5 of each type)
  - Off-ladder alleles
  - Tri-allelic patterns
  - Non-template addition
  - Spikes and signal overload (bleed-through into another dye channel)
  - Mixtures
  - Degraded DNA
- We welcome suggestions as to other types of challenges to include in the data set and **what annotation format would be most useful**
- **Samples are currently being gathered** with plans to generate data using Profiler Plus/COfiler, Identifiler, PowerPlex 16, and SGM Plus (kits have already been purchased)

Peaks Generated Through Overloaded Peaks in Another Dye Color

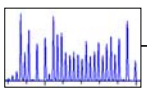


NIST QA/QC Software

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)*

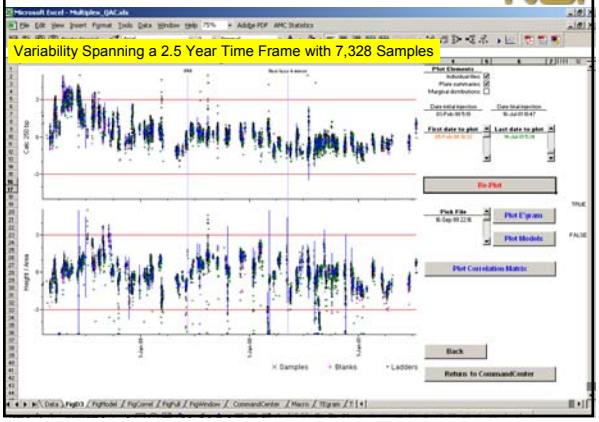


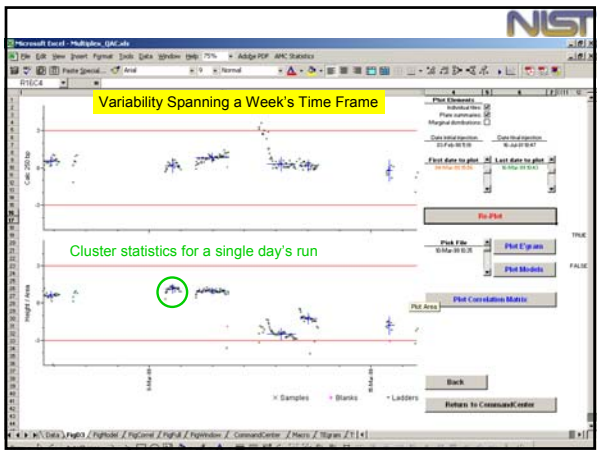
NCBI Program → Peak Height, Area, Size → NIST **Multiplex QAC** Software → Date vs Resolution (or Signal/Noise)

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.

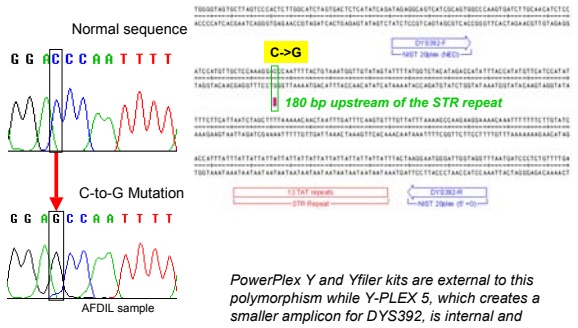




Variant Allele Sequencing

- 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

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…

Interlaboratory Studies

DNA Quantitation (2004),
Mixture Interpretation (2005)

NIST Quantitation Study 2004 (QS04)

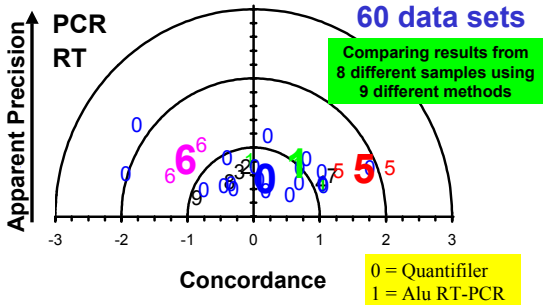
Consisted of:
•8 DNA extracts labeled A – H
•Shipped Dec 2003 –Jan 2004 to 84 laboratories for quantification; data received back by April 2004
•Labs were requested to use multiple methods / multiple analysts

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



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

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

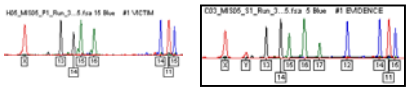
		% Quantitative Results[a]							
Target [DNA] ng/μL		1.5	0.5	0.5	0.16	0.16	0.05	0.05	0.05
Method	N_{eval}	A	B	E	C	F	D	G	H
Quantifiler	37	100	100	100	100	100	100	100	100
Other RT-PCR	23	100	100	100	100	100	100	100	100
"ACES"	14	100	100	100	100	100	100	100	100
AluQuant	13	100	100	100	100	100	100	100	100
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 Jan, ~70 labs are enrolled for participation (17 from overseas)
- Four mock cases supplied with "victim" and "evidence" electropherograms (GeneScan .fsa files – that can be converted for Mac or GeneMapper; gel files will be 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 being shipped in mid-January 2005
- Responses are due before February 28, 2005
- 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 (July 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 will speak 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

Information on New Loci

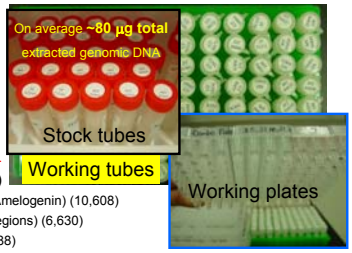
Autosomal SNPs, Y-Chromosome

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

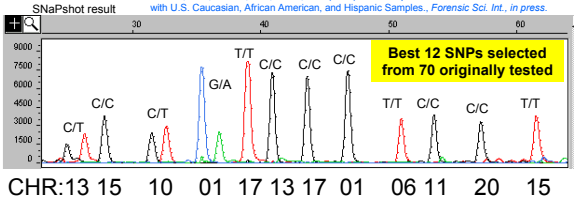
260 Caucasians, 260 African Americans, 140 Hispanics, 3 Asians = **663 males**

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., submitted**
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

SNaPshot result

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., in press.*



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

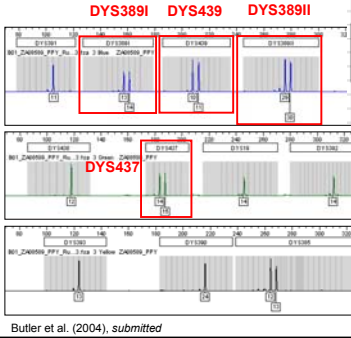
New Y-STR Loci, Issues, and Assays

- 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. (2004) Allele frequencies for 27 Y-STR loci with U.S. Caucasian, African American, and Hispanic samples, *submitted.*
- Chromosomal duplication issues
  - Butler, J.M., Decker, A.E., Kline, M.C., Vallone, P.M. (2004) Chromosomal duplications along the Y-chromosome and their potential impact on Y-STR interpretation, *submitted.*

Duplication at Multiple Loci with Single-Source Sample

Y-chromosome mapping

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

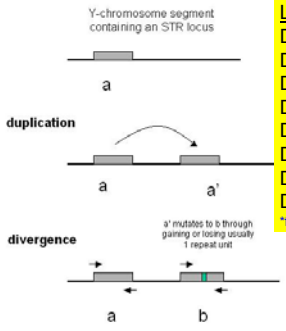


Entire region of Y-chromosome has likely been duplicated and then diverged

Most duplications have a single repeat spread in allele patterns

Butler et al. (2004), *submitted*

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

1 repeat unit difference

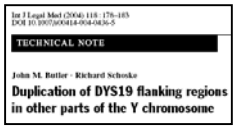
92% have single repeat difference

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

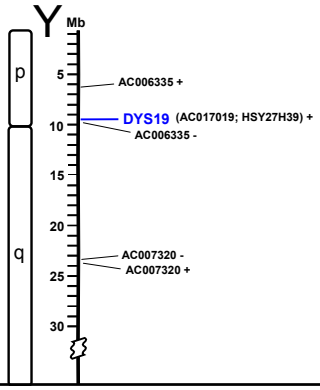
Butler et al. (2004), *submitted*

DYS19 Flanking Region Duplication

Int. J. Legal Med. (2004) 118: 178-183



Different primers around DYS19 repeat result in selection of different regions of the Y-chromosome



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. (2004) Chromosomal duplications along the Y-chromosome and their potential impact on Y-STR interpretation, *J. Forensic Sci., submitted.*
- Butler, J.M., Decker, A.E., Vallone, P.M., Kline, M.C. (2004) Allele Frequencies for 27 Y-STR Loci with U.S. Caucasian, African American, and Hispanic Samples, *Forensic Sci. Int., submitted.*

Acknowledgements

john.butler@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