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

Finding Point Mutations, Deletions and New Alleles Through STR Allele Sequencing

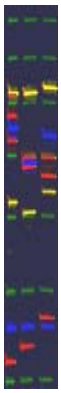
Margaret C. Kline
Michael D. Coble, Jill E. Appleby, Richard Schoske,
John M. Butler

4TH Annual Bode East Coast Advanced DNA Technology Workshop May 23-26 , 2005 Duck Key, FL

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.

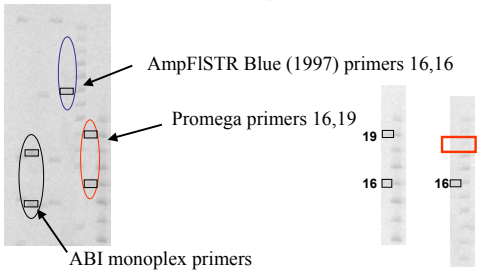


Outline for Presentation

- Explanation of Null and Variant Alleles
- Variant Allele Cataloging and Characterization on STRBase
- STR Allele Sequencing Approach
- Examples

vWA Allele Dropout Observed

What started our interest in sequencing variant alleles



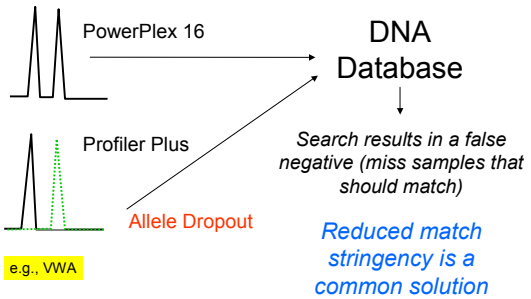
Kline, M.C., Jenkins, B. & Rodgers, S. (1998) Non-amplification of a vWA allele. *J Forensic Sci.*, 43(1), p250

Null Alleles

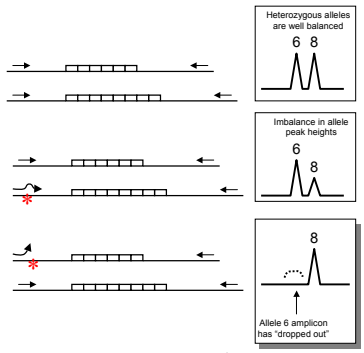
- Allele is present in the DNA sample but fails to be amplified due to a **nucleotide change in a primer binding site**
- Allele dropout is a problem because a heterozygous sample appears falsely as a homozygote
- Two PCR primer sets can yield different results on samples originating from the same source
- 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

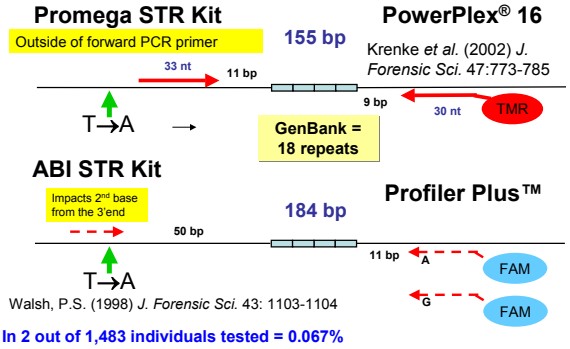


Impact of DNA Sequence Variation in the PCR Primer Binding Site



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%

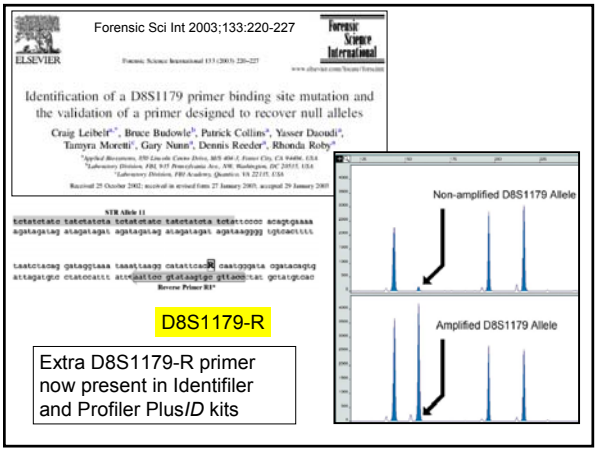
Lazaruk *et al.* (2001) *Forensic Sci Int.* 119:1-10

Apparent Null Alleles Observed During Concordance Studies

10/13 CODIS loci affected so far

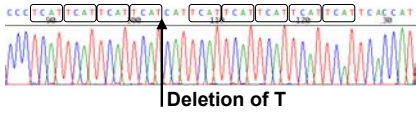
Locus	STR Kits/Assays Compared	Results	Reference
VWA	PP1.1 vs ProPlus	Loss of allele 19 with **ProPlus**; fine with PP1.1	Kline et al. (1998)
D5S818	PP16 vs ProPlus	Loss of alleles 10 and 11 with **PP16**; fine with ProPlus	Alves et al. (2003)
D13S317	Identifiler vs miniplexes	Shift of alleles 10 and 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



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 Common Variant Alleles eg. D7S820

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

GATAGAACACTTGTCATAGTTTAGAACGAACTAACGATAGATAGATAGATAGATAGATAG
CTATCTTGTGAACAGTATCAAATCTTGCTTGATTGCTATCTATCTATCTATCTATCTATC

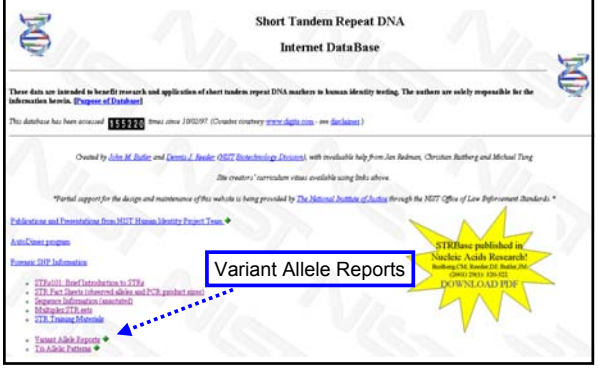
13 repeat units = $(GATA)_{13}$

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

9T's nominal "on ladder"

ATAGATAGATAGATAGATAGATAGATAGACAGATTGATAGTTTTTTTTTAATCTCACTAAA
TATCTATCTATCTATCTATCTATCTATCTGTCTAACTATCAAAAAAAAATTAGAGTGATTT

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



Variant Alleles Cataloged in STRBase

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

Updated as of May 9, 2005

Variant Allele Reports

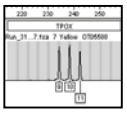
Currently **269** at 13/13 CODIS loci

- CSF1PO (12)
- D2S1338 (3)
- D3S1358 (19)
- D5S818 (5)
- D7S820 (22)
- D8S1179 (4)
- D13S317 (10)
- D16S539 (11)
- D18S51 (30)
- D19S433 (11)
- D21S11 (24)
- FESFPS (1)
- FGA (69)
- HUMTH01 (7)
- Penta D (14)
- Penta E (13)
- TPOX (8)
- VWA (6)

Tri-Allelic Patterns

Currently **63** at 13/13 CODIS loci

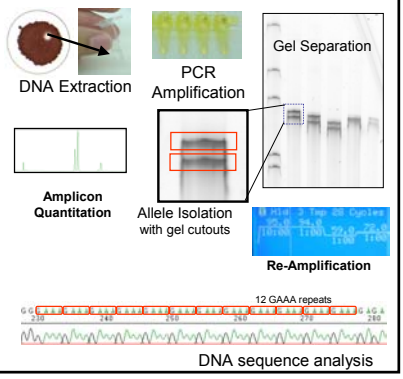
- CSF1PO (2)
- D3S1358 (4)
- D5S818 (2)
- D7S820 (2)
- D8S1179 (5)
- D13S317 (3)
- D16S539 (2)
- D18S51 (7)
- D21S11 (4)
- FGA (10)
- HUMTH01 (1)
- TPOX (13)
- VWA (8)



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

Steps in STR Allele Sequencing

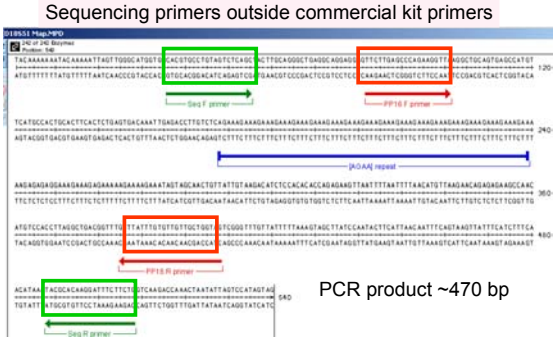
- DNA Extraction
- Amplification with primers external to kit primers
- Gel Cutouts with Heterozygotes
- Re-Amplification
- Amplicon Quantitation
- ExoSAP
- Cycle Sequencing
- Dye Terminator Removal
- F/R Sequence Alignment to Reference Sequence



Steps in STR Allele Sequencing

PCR Amplification D18S51

Sequencing primers outside commercial kit primers

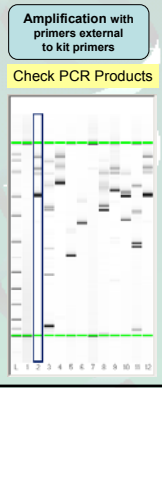


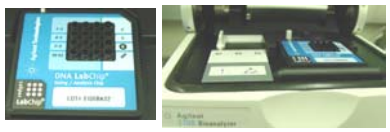
PCR product ~470 bp

Steps in STR Allele Sequencing

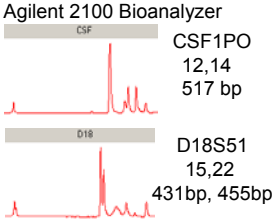
Amplification with primers external to kit primers

Check PCR Products





Agilent 2100 Bioanalyzer



CSF1PO
12,14
517 bp

D18S51
15,22
431bp, 455bp

Steps in STR Allele Sequencing

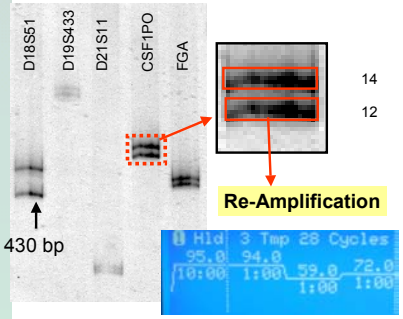
Gel Cutouts with Heterozygotes

Re-Amplification

Allele Isolation with gel cutouts

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

Gel Separation



Steps in STR Allele Sequencing

Amplicon Quantitation

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

ExoSAP

Cycle Sequencing

Dye Terminator Removal

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)

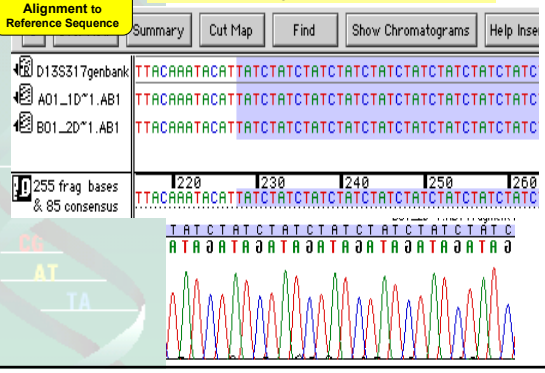
Cycle Sequencing (R primer only)

DYE Terminator removal

Steps in STR Allele Sequencing

F/R Sequence Alignment to Reference Sequence

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



EXAMPLES

(1) Deletions impacting miniSTRs with D13

(2) Deletions impacting TPOX with PP 1.1/ IDfiler vs PP 2.1/PP 16

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

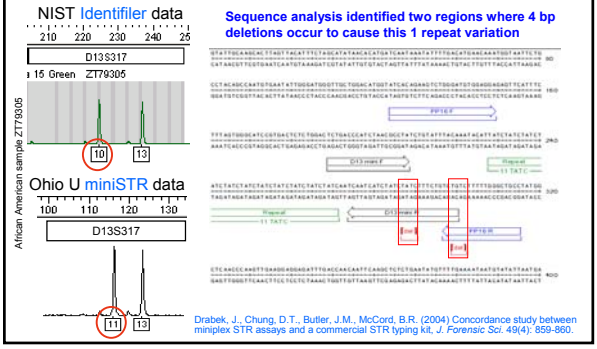
(4) Small D18S51 allele "5.3"

(5) Large D18S51 allele 40

(6) Characterization of DYS635 allele 21.3 variant

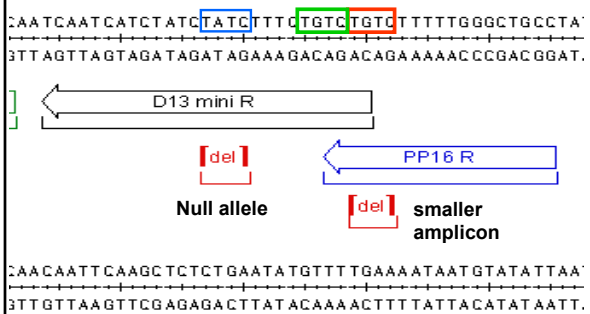
D13S317 Flanking Region Deletion

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

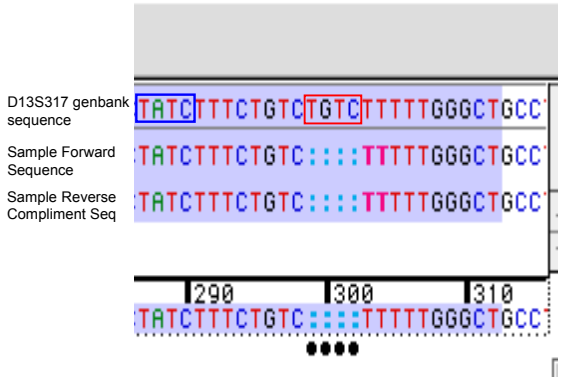


Drabek, J., Chung, D.T., Butler, J.M., McCord, B.R. (2004) Concordance study between miniplex STR assays and a commercial STR typing kit, *J. Forensic Sci.* 49(4): 859-860.

D13S317 Flanking Region Deletion



ZT79305: D13S317 TGTC deletion



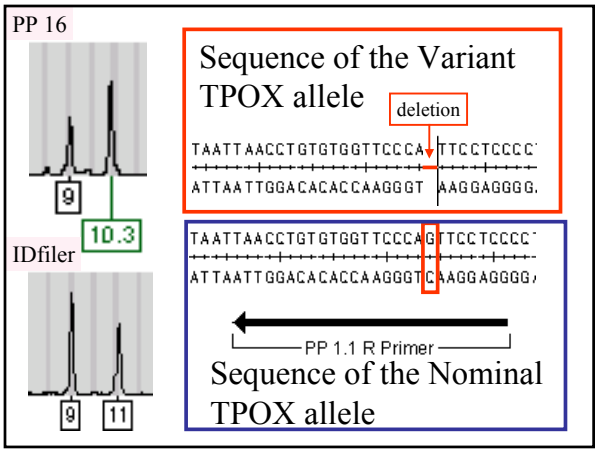
Variation in the Flanking Region Can Cause Discordant Alleles eg TPOX

- Received a sample with discordant results at the TPOX locus
- PP 1.1 and Identifiler type as 9,11
- PP 2.1 / PP 16 type as 9, 10.3
- Sequencing the sample we found:
  - A deletion157 bases from the end of the repeats
  - This deletion is within region where the PP 1.1 primer anneals (8 bases from the 5' end of the reverse primer).

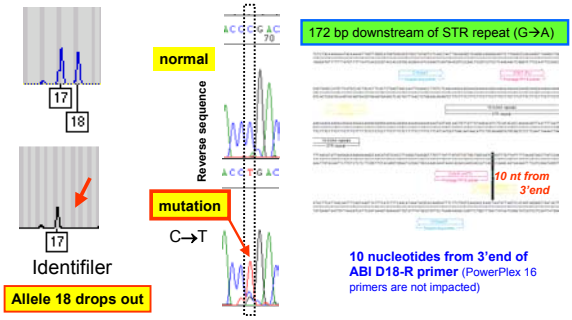
Sample courtesy of Maryland State Police Laboratory

Nominal TPOX Allele 11

Deletion results in a 10.3 allele call with PP 16 or PP 2.1. The deletion does not influence PP 1.1 or Identifiler results.

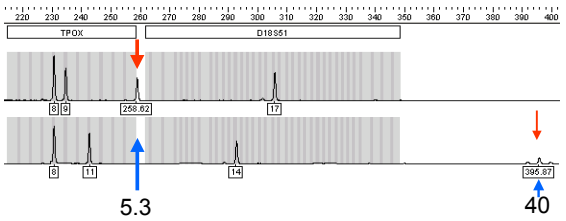


D18S51 Null Allele from Kuwait Samples with ABI Primers PowerPlex 16



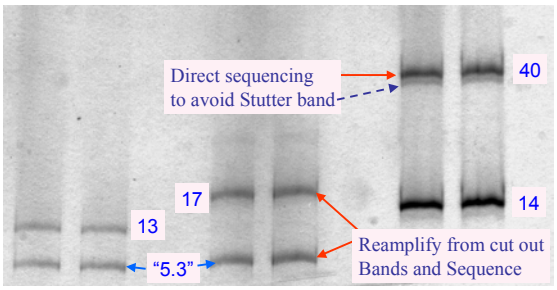
Clayton et al. (2004) **Primer binding site mutations affecting the typing of STR loci contained within the AMPFISTR SGM Plus kit.** *Forensic Sci Int.* 139(2-3): 255-259

Extreme D18S51 Variants Range 258 – 396 bp

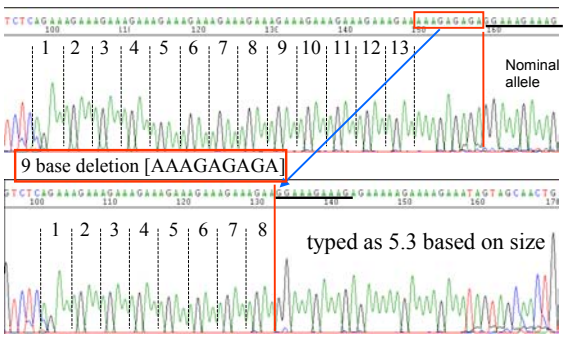


Samples amplified with Identifiler.Nominal range 8 - 27
Small D18 allele just outside of the TPOX allele range.
Large D18 is 46 bp larger than the D18 allele range.

Gel Separation Of D18 Variant Alleles

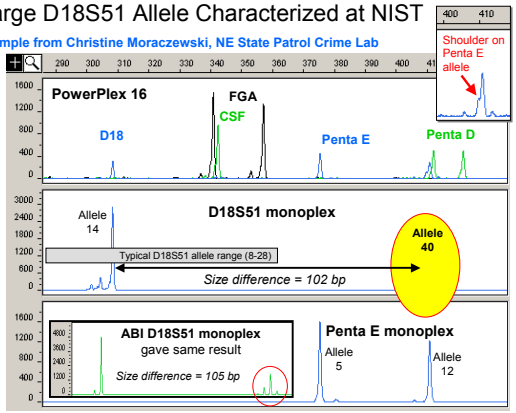


D18S51 sequencing Sample 188



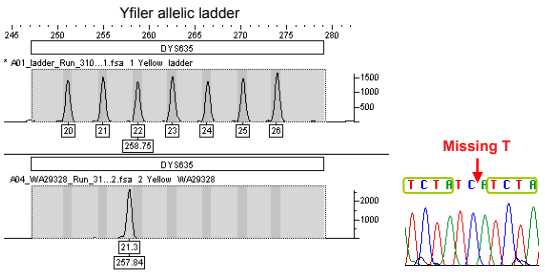
Nominal allele range 8-27 repeats. Sample courtesy of DNA Solutions

Large D18S51 Allele Characterized at NIST



DYS635 Variant Allele 21.3

(Y-GATA-C4)

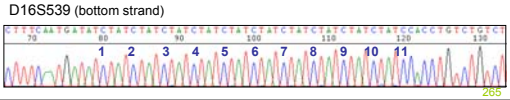


$[TCTA]_4(TGTA)_2[TCTA]_2(TGTA)_2[TCTA]_2(TGTA)_2[TCTA]_5$ TC-A $[TCTA]_2$

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)



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

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

Sample Suppliers and Collaborators

- **Those who have sent samples:**
- NE State Patrol Crime Lab
- DNA Solutions Inc., OK
- FSS
- Kuwait
- MD State Police
- **Mini STR Concordance Collaborator**
  - Bruce McCord, Denise Chung

Acknowledgements

Funding:

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

margaret.kline@nist.gov

john.butler@nist.gov

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