Concordance Testing Comparing STR Multiplex Kits with a Standard Data Set

Becky Hill, Margaret Kline, David Duewer,

Peter Vallone, and John Butler

National Institute of Standards and Technology

American Academy of Forensic Sciences

Chicago, IL

February 25, 2011

Disclaimer

NIST Funding: Interagency Agreement 2008-DN-R-121 between the National Institute of Justice and NIST Office of Law Enforcement Standards

NIST Disclaimer: 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.

Points of view are mine and do not necessarily represent the official position or policies of the US Department of Justice or the National Institute of Standards and Technology.

Outline of Topics to Discuss

- Introduction and importance of concordance testing
  - Overlapping markers with different primer configurations
- NIST role in concordance testing
  - SRM 2391b concordance with new kits
  - Standard sample set, DNA sequencing
- Commercial STR multiplex kits examined
  - Applied Biosystems, Promega, and Qiagen
- Concordance results with various STR multiplex kits
  - Primer binding site mutations and null alleles
- Summary and conclusions

Why are concordance studies important?

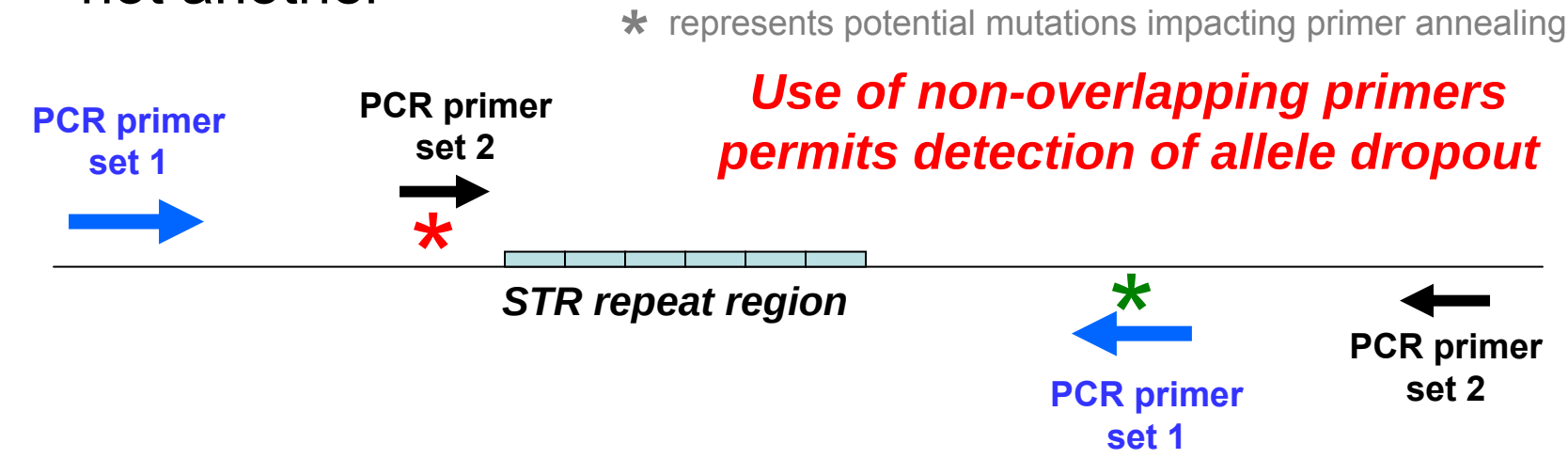
Importance of Concordance Testing

- There are a variety of commercial STR multiplex kits with different configurations of STR markers
  - Different primer sequences are used to amplify the same markers
  - Discordant results can impact DNA databases
- Detection of primer binding site mutations that cause **null alleles**, or allele drop-out
  - Can only be determined with concordance testing and DNA sequencing
- Concordance with NIST reference materials
  - Important to test with all new STR typing kits

Hill, C.R., Kline, M.C., Duewer, D.L., Butler, J.M. (2010) Strategies for concordance testing. *Profiles in DNA (Promega)*, 13(1).

Purpose of Concordance Studies

When different primer sets are utilized, there is a concern that allele dropout may occur due to primer binding site mutations that impact one set of primers but not another

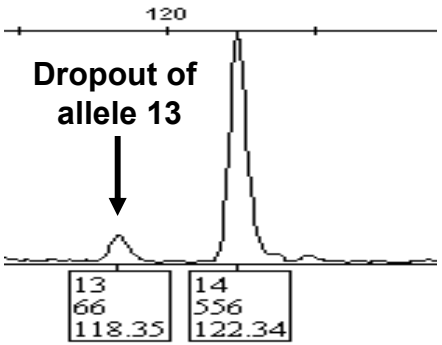


If no primer binding site mutations

If a primer binding site mutation exists

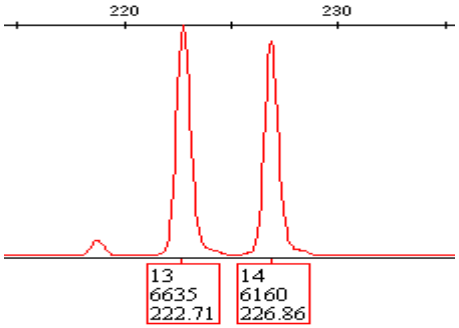
Example Primer Binding Site Mutation that Causes a Null Allele

Identifiler = 14,14



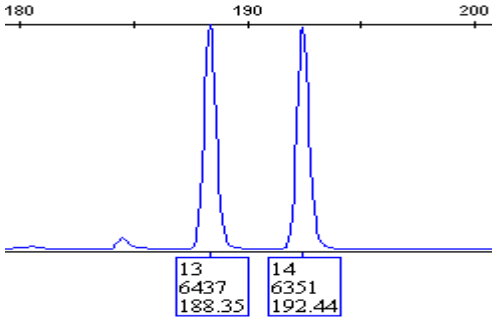
PHR = 11.9%

PP ESX 17 = 13,14



PHR = 92.8%

PP ESI 17 = 13,14



PHR = 98.7%

G → A SNP

D19S433 repeat region

ggaaagaaggtaggaaggaaggaaggaagtaagtaaggaaggaaggaaggaaggaaggaaggagagaggaagaaagagagaagatttttattcgggtaat

X

This region could potentially represent where the reverse primer is located to include the primer binding site mutation

*Applied Biosystems does not publish their primer sequences

To Avoid Overlapping PCR Product Size Ranges with STR Loci in the Same Dye Channel

- Applied Biosystems (Strategy 1)
  - **Maintains primer sequences** (except MiniFiler & NGM kits)
  - Utilizes mobility modifiers or additional dyes, no primer redesign is necessary
  - Enables comparison to legacy data with earlier kits but null alleles may go undetected with the potential for incorrect genotypes within data sets

- Promega Corporation (Strategy 2)
  - Moves primer sequences to change PCR product size ranges
  - Primer redesign can be difficult, but can be moved from primer-binding-site mutations
  - **Requires concordance studies to check for potential allele dropout**

Why is NIST involved in concordance studies?

Purpose of Concordance Studies

1. To test SRM 2391b (PCR-based DNA Profiling Standard) components with all new STR multiplex kits and verify results against certified reference values

2. To gain a better understanding of primer binding site mutations that cause null alleles

What are the NIST strategies for concordance testing?

STR Kit Concordance Testing

Profiles in DNA Article Published April 2010

Article Type: Feature

Volume 13 No. 1, April 2010

Strategies for Concordance Testing

Carolyn R. Hill, Margaret C. Kline, David L. Duewer and John M. Butler
National Institute of Standards and Technology, Biochemical Science Division, Gaithersburg, Maryland, USA

Concordance evaluations are important to conduct to determine if there are any allelic dropout or "null alleles" present in a data set. These studies are performed because there are a variety of commercial short tandem repeat (STR) multiplex kits with different configurations of STR markers available to the forensic community. The placement of the markers can vary between kits because the primer sequences were designed to amplify different polymerase chain reaction (PCR) product sizes. When multiple primer sets are used, there is concern that allele dropout may occur due to primer-binding-site mutations that affect one set of primers but not another.

http://www.promega.com/profiles/1301/1301_08.html

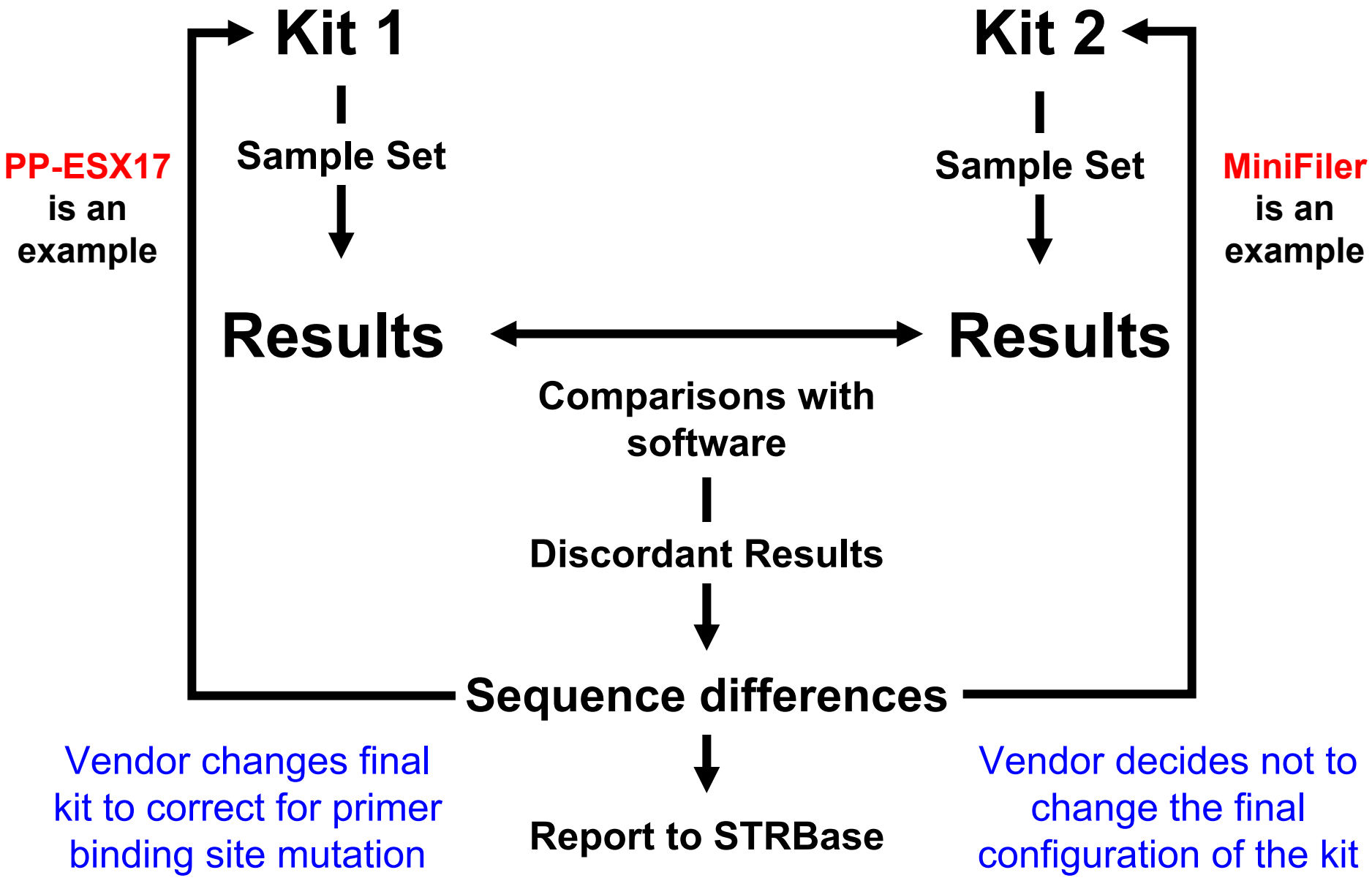
The 4 “S’s” of Concordance

- NIST Standard Samples
  - Run same samples with multiple kits to compare results
- Concordance Software
  - Allows comparison of data sets using NIST developed software

    http://www.cstl.nist.gov/biotech/strbase/software.htm
- DNA Sequencing
  - To validate and determine the exact cause for the null allele
- STRBase website
  - To report verified null alleles and discordant results to the forensic community

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

NIST Concordance Testing Steps



What concordance studies have been completed thus far?

Applied Biosystems AmpFℓSTR Kits

- Identifiler
- **MiniFiler**
- Profiler Plus
- SGM Plus
- NGM
- NGM SElect (studies are ongoing)

Hill, C.R., Kline, M.C., Mulero, J.J., Lagace, R.E., Chang, C.-W., Hennessy, L.K., Butler, J.M. (2007) Concordance study between the AmpFlSTR MiniFiler PCR Amplification Kit and conventional STR typing kits. *J. Forensic Sci*. 52(4): 870-873.

Promega PowerPlex Systems

- PowerPlex 16
- **PowerPlex ESX 17**
- **PowerPlex ESI 17**
- PowerPlex 18D (rapid and direct kit)

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Concordance and population studies along with stutter and peak height ratio analysis for the PowerPlex® ESX 17 and ESI 17 Systems

Carolyn R. Hill[a,*], David L. Duewer[a], Margaret C. Kline[a], Cynthia J. Sprecher[b], Robert S. McLaren[b], Dawn R. Rabbach[b], Benjamin E. Krenke[b], Martin G. Ensenberger[b], Patricia M. Fulmer[b], Douglas R. Storts[b], John M. Butler[a]

[a] National Institute of Standards and Technology, Chemical Science and Technology Laboratory, Gaithersburg, MD 20899-8312, USA
[b] Promega Corporation, Madison, WI 53711-5399, USA

Qiagen Investigator HID Kits

- ESSplex
- IDplex
- Hexaplex ESS (studies are ongoing)

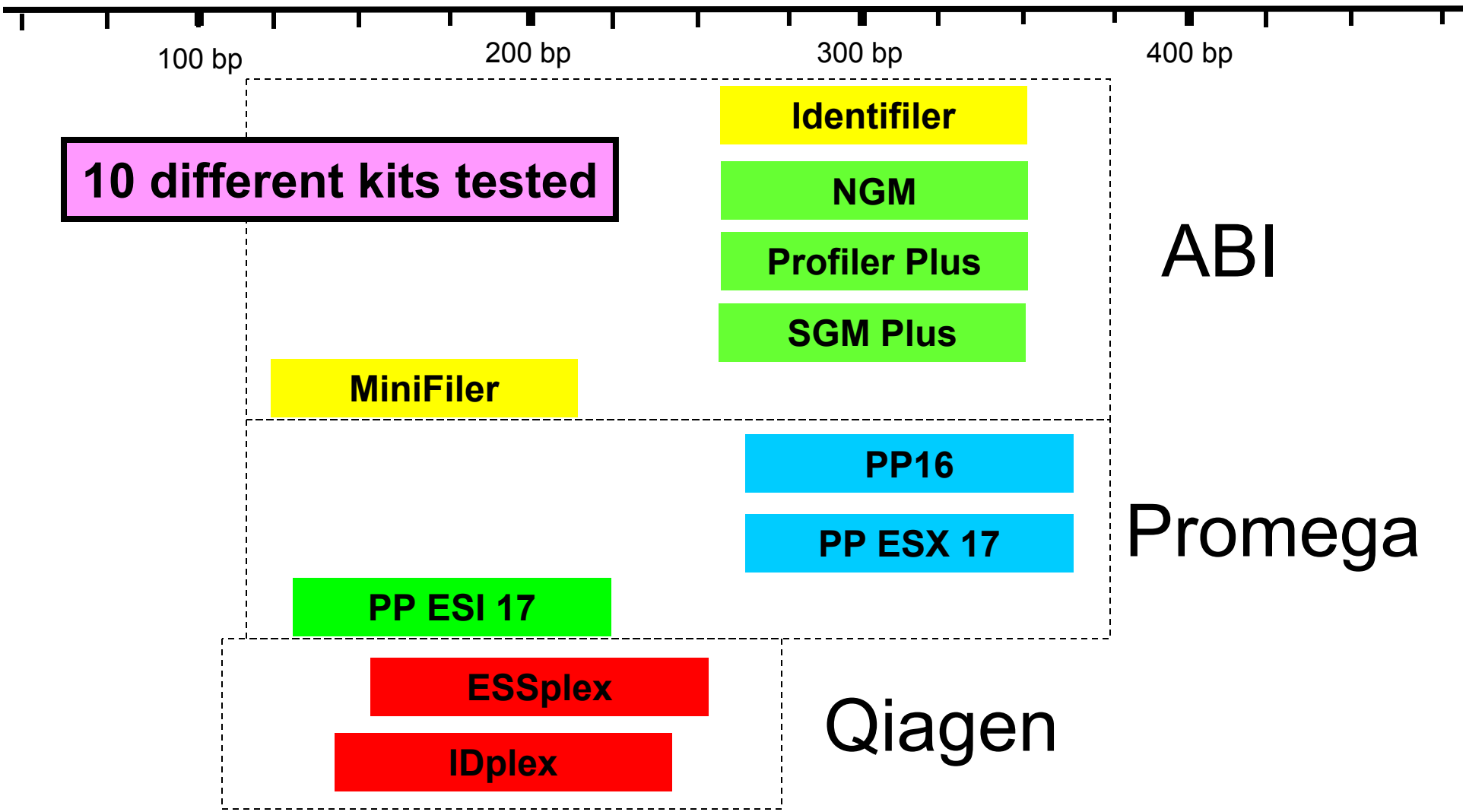
What samples are used at NIST to perform concordance testing?

NIST Sample Set (>1400 Samples)

- **NIST U.S. population samples**
  - 254 African American, 261 Caucasian, 139 Hispanic, 3 Asian
- **U.S. father/son paired samples**
  - 178 African American, 198 Caucasian, 190 Hispanic, 198 Asian
- **NIST SRM 2391b**, PCR-based DNA Profiling Standard (highly characterized)
  - 10 genomic DNA samples, 2 cell line samples
  - Includes 9947A and 9948

What are the results from the completed concordance studies?

D18S51 Concordance Checking



D18S51 marker positions for all kits tested

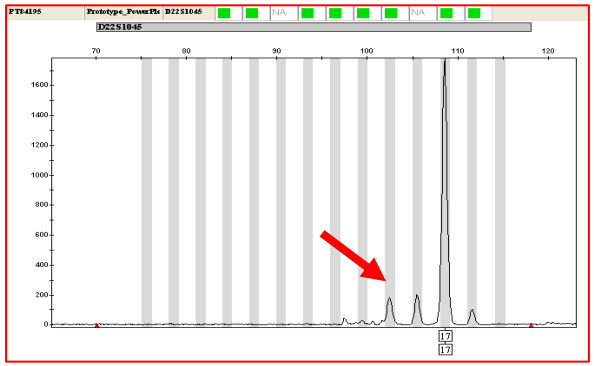
Primer Set Compared

Marker	# of Sets
D18S51	**10**
Amelogenin	**10**
D21S11	**10**
FGA	**10**
D3S1358	**9**
TH01	**9**
D16S539	**9**
vWA	**9**
D8S1179	**9**
D2S1338	**8**
D10S1248	**7**
D22S1045	**7**

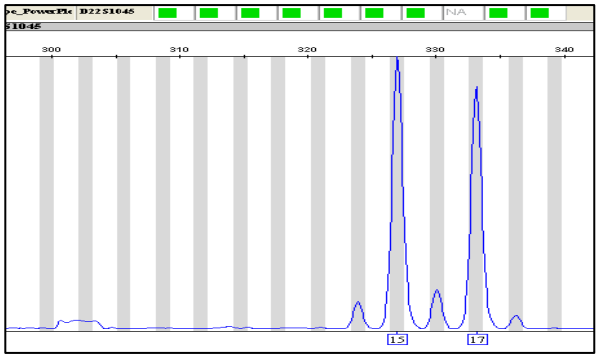
Marker	# of Sets
D2S441	**7**
D19S433	**7**
D1S1656	**5**
D12S391	**5**
D5S818	**4**
D7S820	**4**
D13S317	**4**
SE33	**4**
TPOX	**3**
CSF1PO	**3**
Penta D	**1**
Penta E	**1**

D22S1045 Discordance

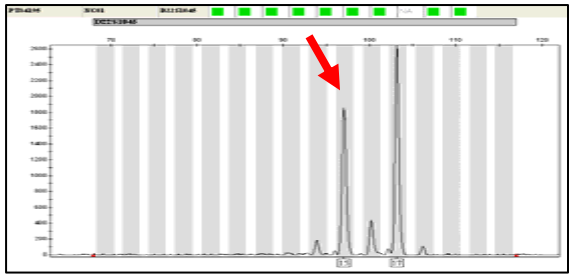
ESX 17 (prototype) = 17,17



ESI 17 (prototype) = 15,17

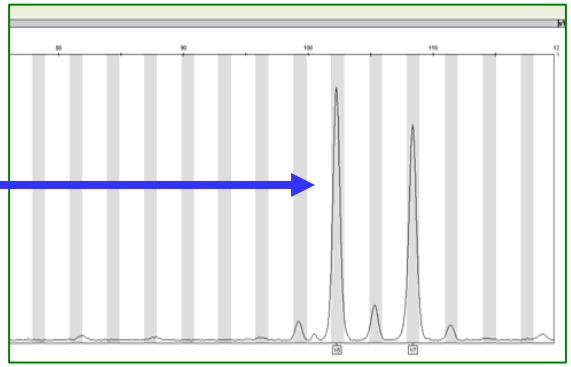


NIST NC01 = 15,17

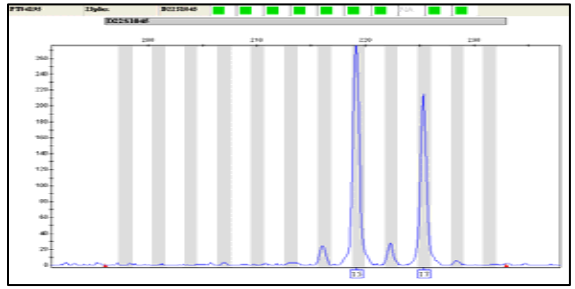


Destabilized some (but uses lower annealing temperature with fewer amplicons in multiplex)

ESX 17 (final) = 15,17



NIST 23plex = 15,17

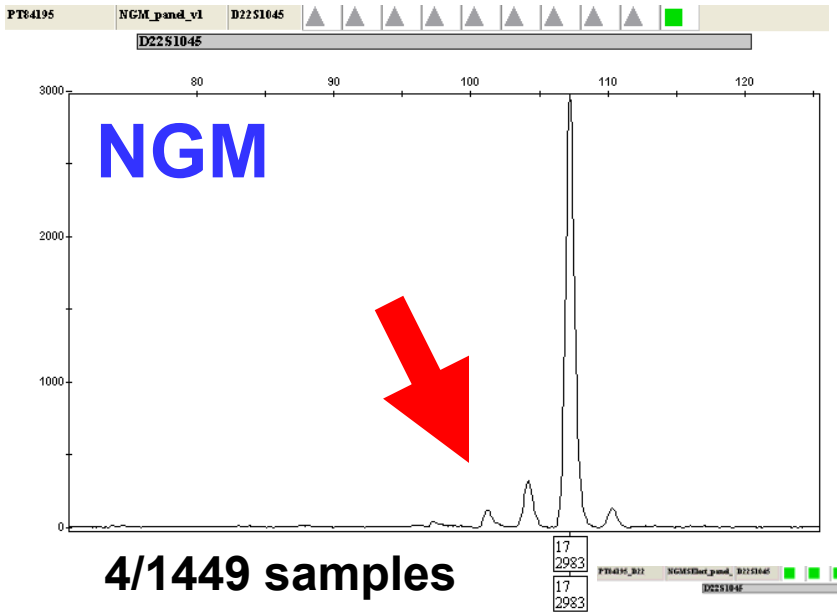


NIST PT84195

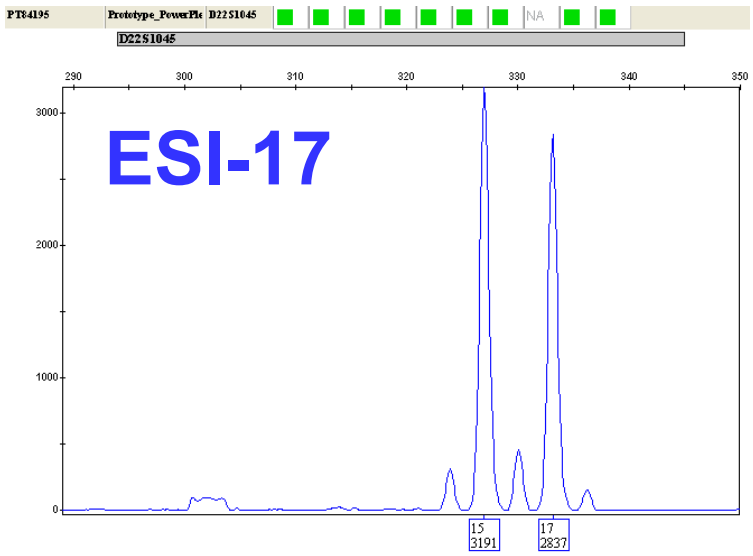
G→T 15 bp upstream impacting forward primer binding with ESX17

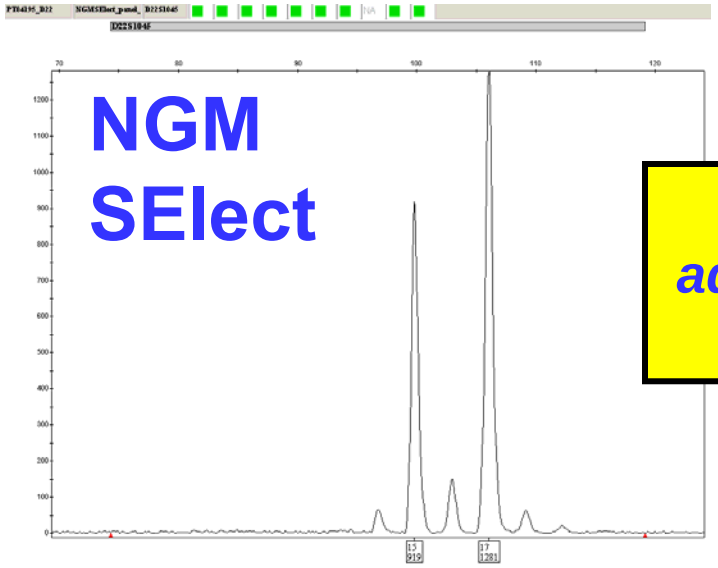
Promega added additional primer to correct issue

D22S1045 Null Allele



4/1449 samples



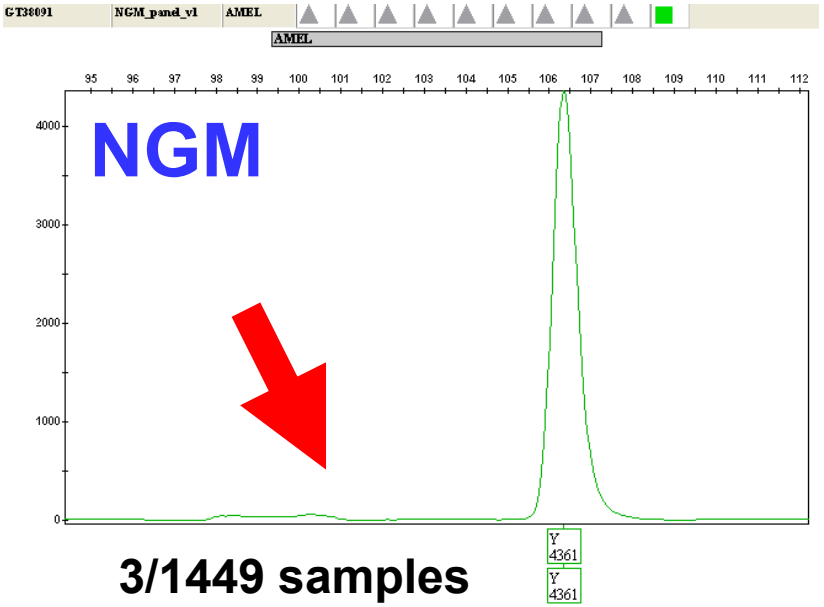


Correct type (15,17)

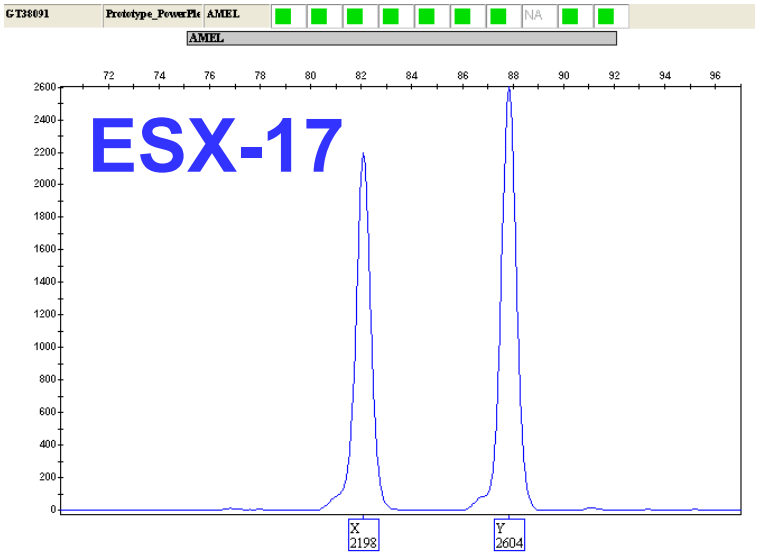
ABI added an additional primer to correct issue

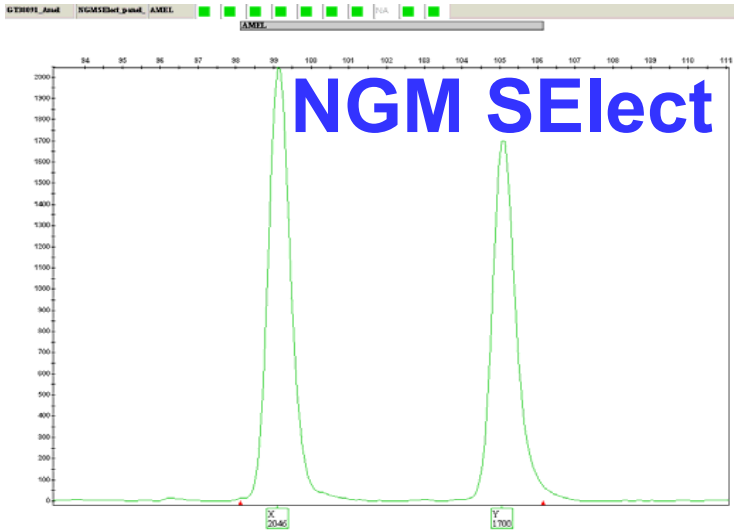
G→T 15 bp upstream impacting forward primer binding with NGM

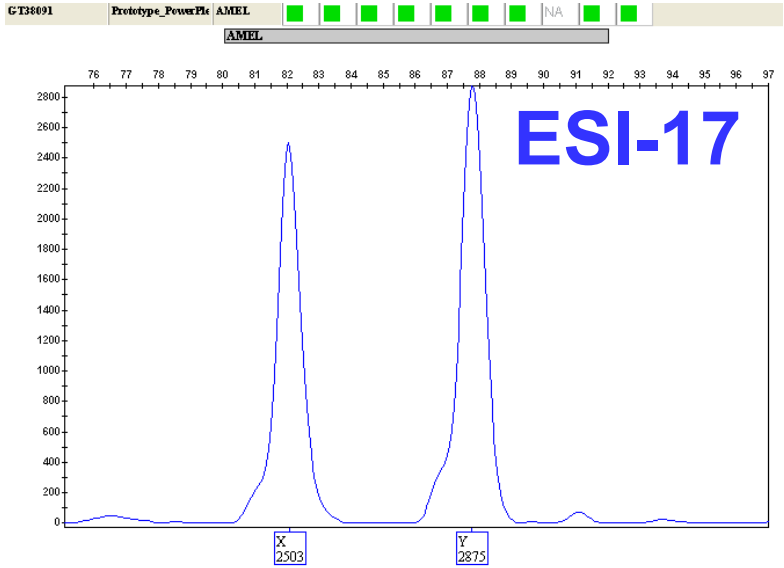
Amelogenin X Null Allele



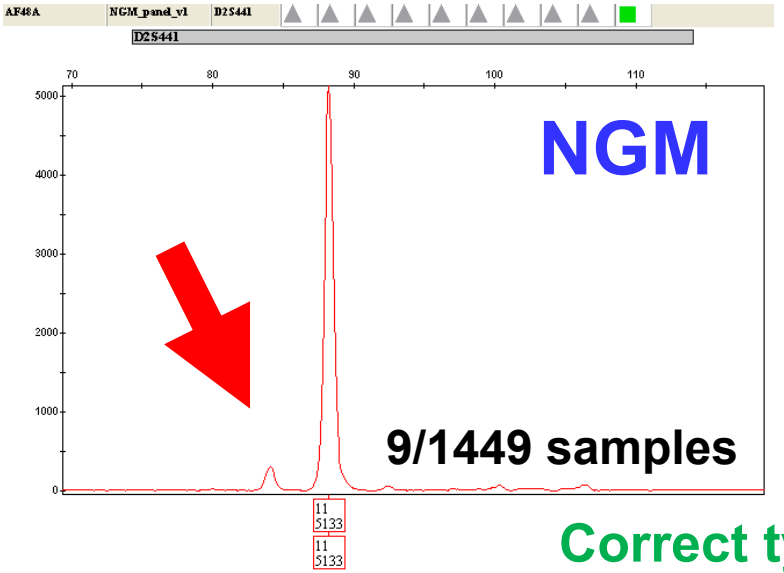
3/1449 samples

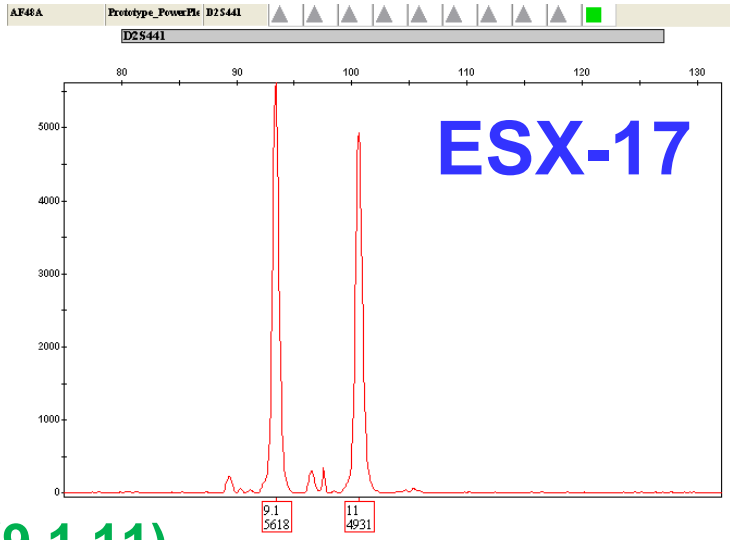




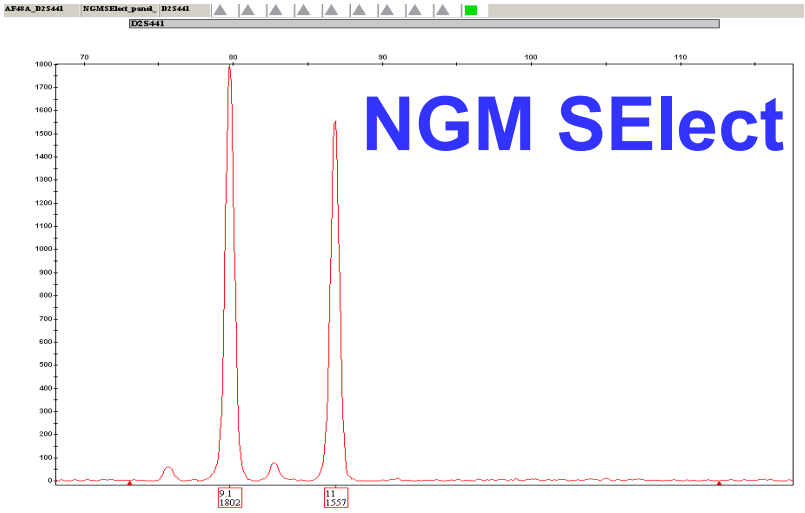


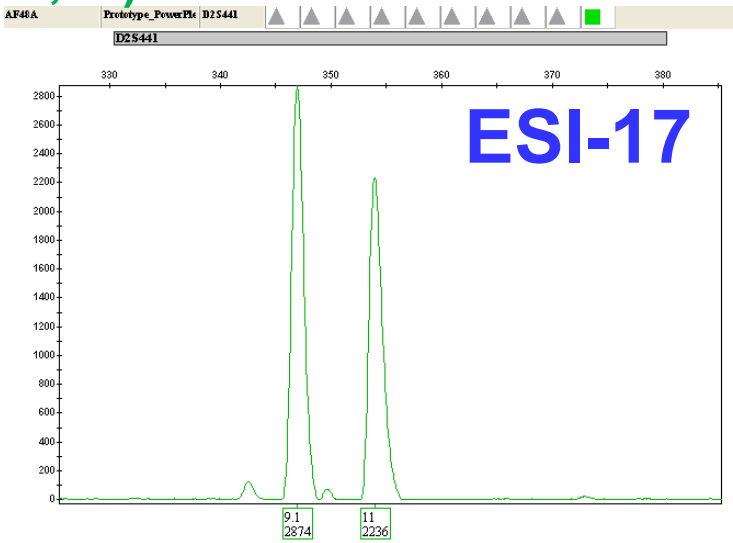
D2S441 Null Allele





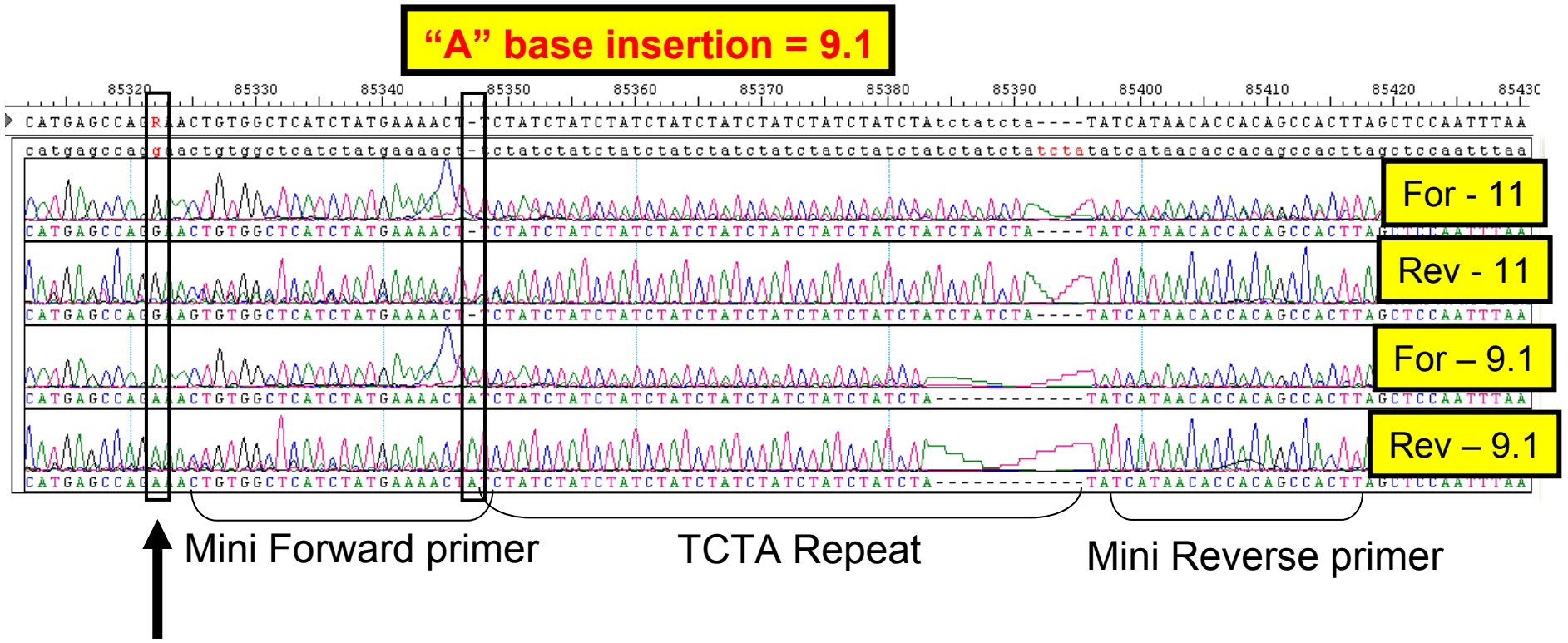
Correct type (9.1,11)





8/9 null alleles were from Asian samples

D2S441 Sequencing

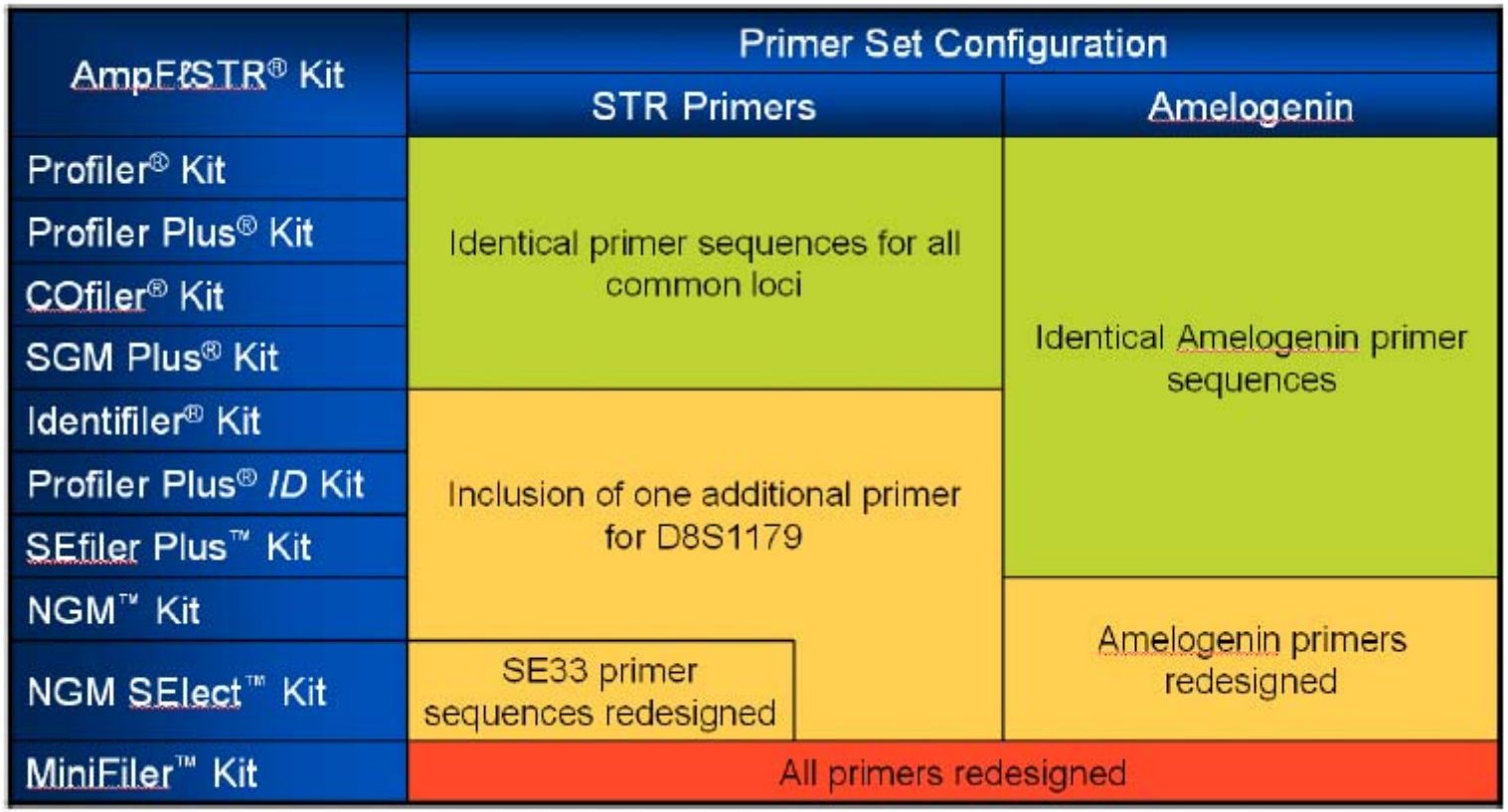


G → A SNP 25/26 bp upstream of the repeat

True Genotype = 9.1,11

NGM Genotype = 11,11

Primer Changes with ABI Kits

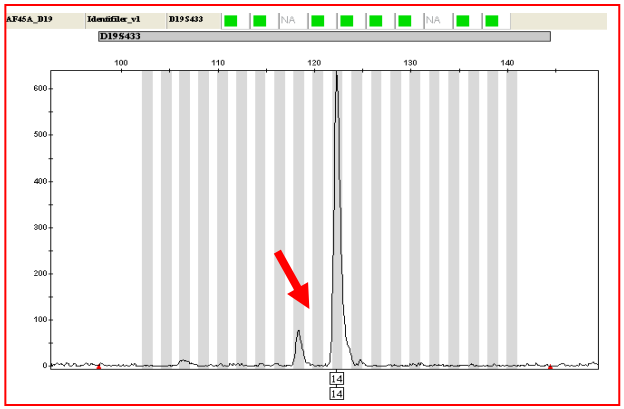


D2S441 and D22S1045 have an additional primer in NGM and NGM SElect

Table 4 from "Development of the AmpFℓSTR NGM SElect Kit: New Sequence Discoveries and Implications for Genotype Concordance", Forensic News (January 2011)

D19S433 Discordance

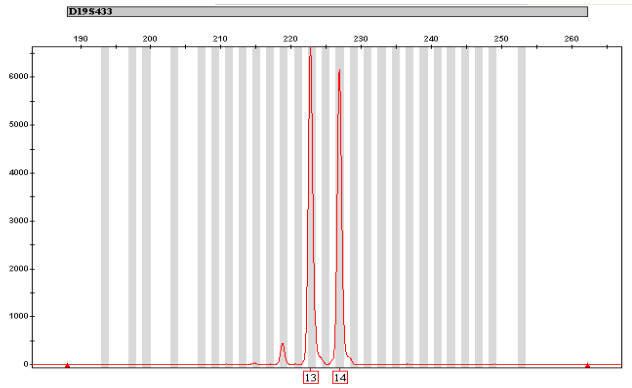
Identifiler & NGM = 14,14



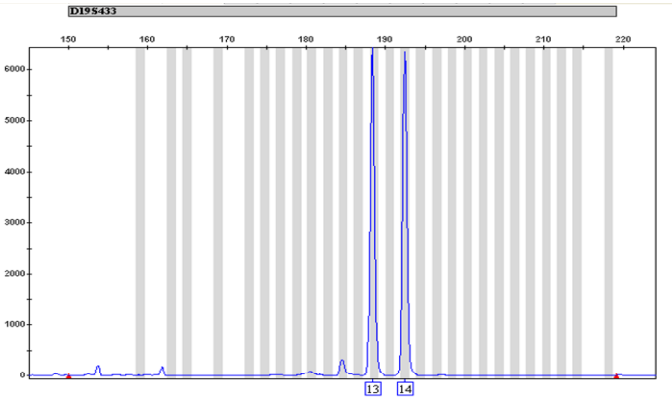
AF45A (Asian)

Allele 13 was missing in **two different Asian samples** with ABI primers = 2/2886 = **0.07% discordance**

ESX 17 = 13,14



ESI 17 = 13,14



Frequencies [for] the silent allele were determined to be 0.0114 in 176 people from Shizuoka (Honshu) and 0.0128 in 156 people from Okinawa

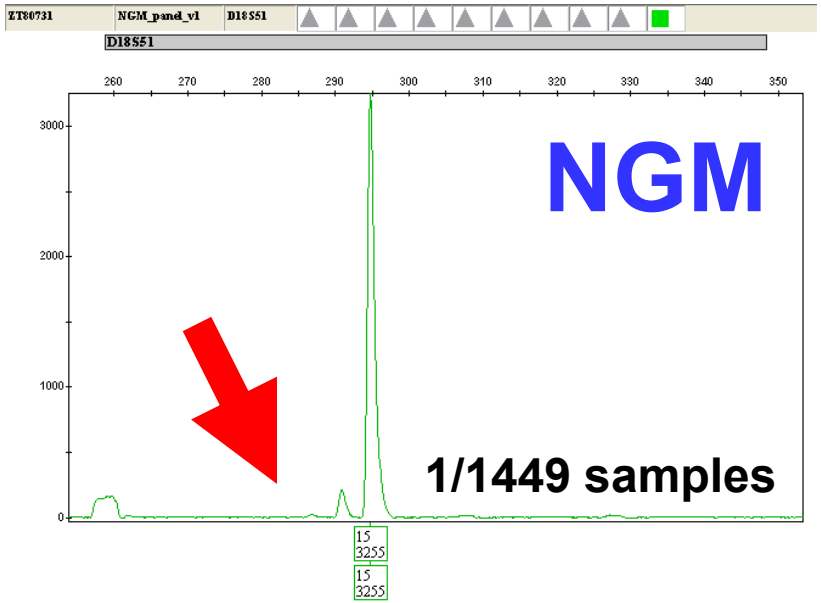
J Forensic Sci, September 2008, Vol. 53, No. 5
doi: 10.1111/j.1556-4029.2008.00806.x
Available online at: www.blackwell-synergy.com

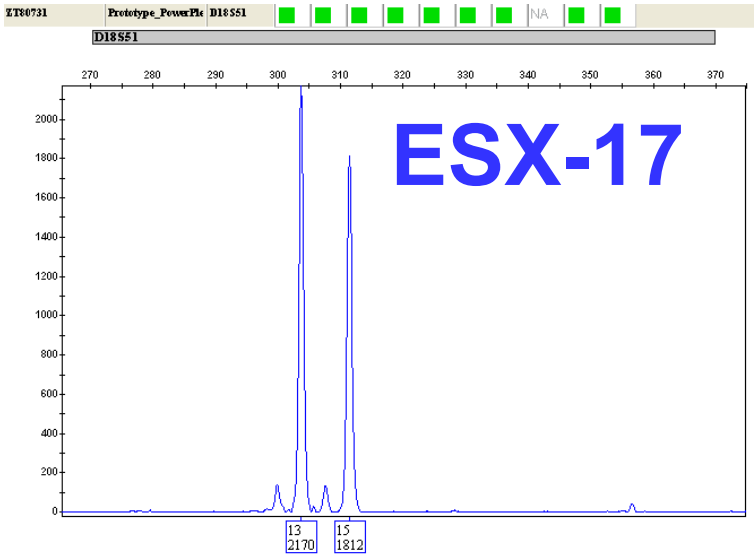
Natsuko Mizuno,[1] D.V.M.; Tetsushi Kitayama,[1] M.Sc.; Koji Fujii,[1] Ph.D.; Hiroaki Nakahara,[1] D.V.M.; Kanako Yoshida,[1] Ph.D.; Kazumasa Sekiguchi,[1] Ph.D.; Naoto Yonezawa,[2] Ph.D.; Minoru Nakano,[2] Ph.D.; and Kentaro Kasai,[1] Ph.D.

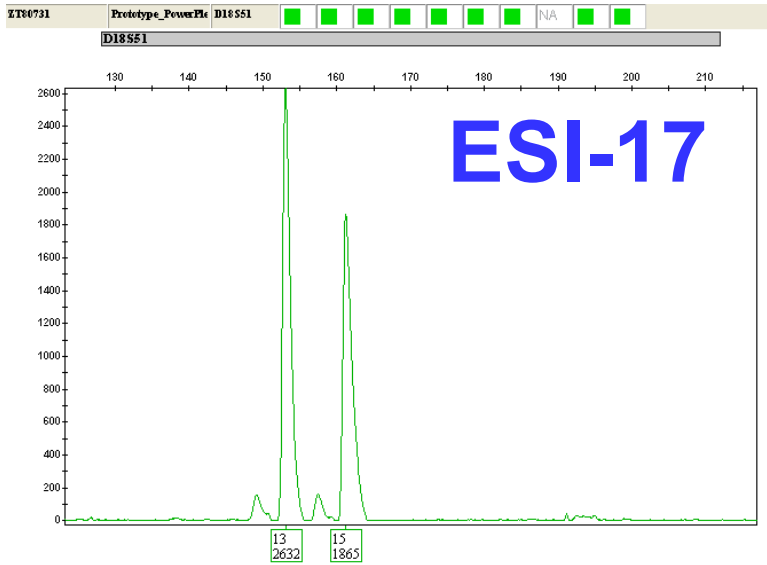
A D19S433 Primer Binding Site Mutation and the Frequency in Japanese of the Silent Allele It Causes

T→A 8 bp downstream impacting reverse primer binding with Identifiler (and thus SGM Plus)

D18S51 Null Allele



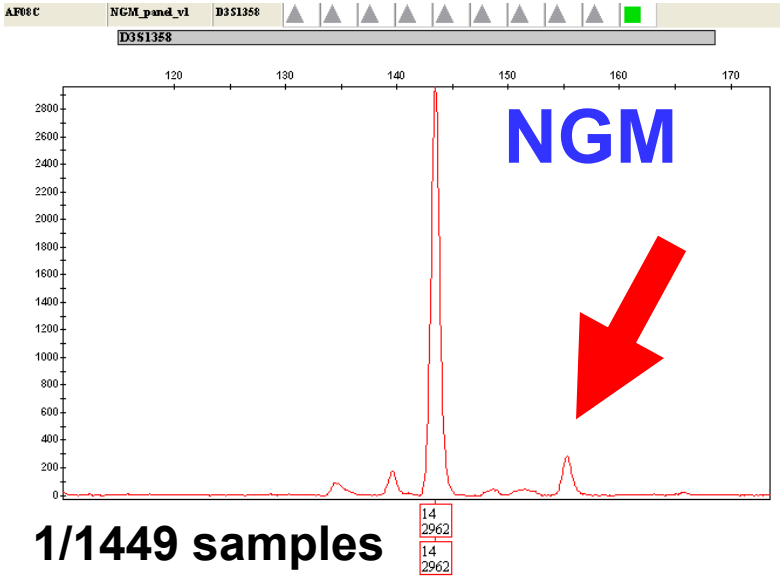




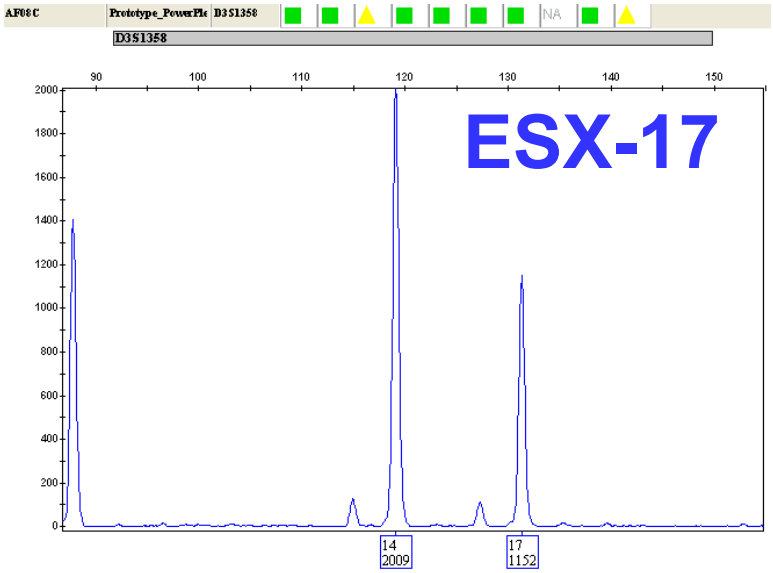
Correct type (13,15)

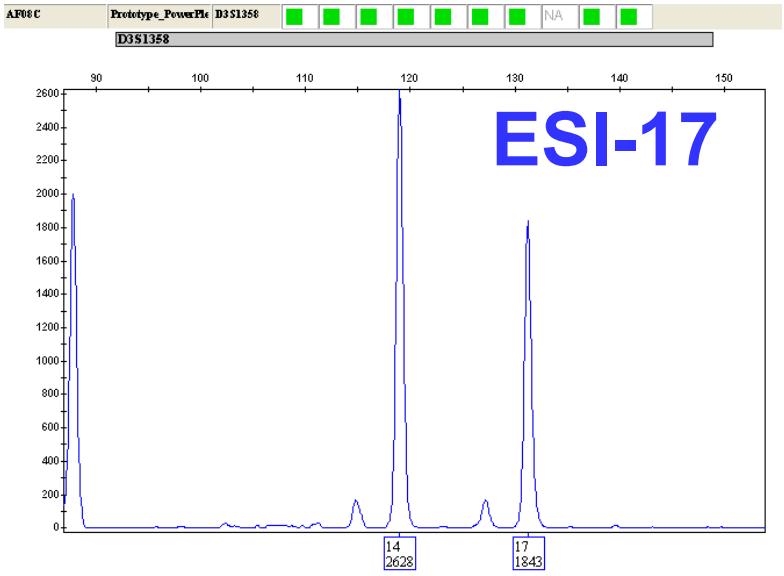
C→T SNP 172 bp downstream from repeat

D3S1358 Null Allele



1/1449 samples





Correct type (14,17)

G → C SNP 11 bp downstream from repeat

Completed Concordance Studies

Kits compared	Samples	Loci compared	Comparisons	# Differences	Concordance (%)
SGM-ID	1436	11	15,796	1	99.994%
ID-ProPlus	1427	10	14,270	1	99.993%
SGM-NGM	1436	11	15,796	4	99.975%
ID-NGM	1449	11	15,939	3	99.981%
ProPlus-NGM	1427	10	14,270	4	99.972%
SGM-ESI	1436	11	15,796	5	99.968%
ProPlus-ESX	1427	7	9,989	3	99.970%
ESI-NGM	1449	16	23,184	15	99.935%
ESX-NGM	1449	16	23,184	17	99.927%
ESI-ESX	1455	17	24,735	3	99.988%
		TOTAL	**172,959**	**56**	**99.970%**

172,959 comparisons
56 differences
99.97% concordance

Kits (except Identifiler) were kindly provided by ***Promega and Applied Biosystems*** *for concordance testing performed at NIST*

Qiagen Investigator Kit Concordance Results

- Results have been obtained for the ESSplex and IDplex kits and shared with Qiagen

- Confirmation DNA sequencing is currently being performed to verify the results

Was there complete concordance with SRM 2391b?

SRM 2391b
PCR-Based Profiling Standard

- The first set of samples run with new STR multiplex kits is SRM 2391b

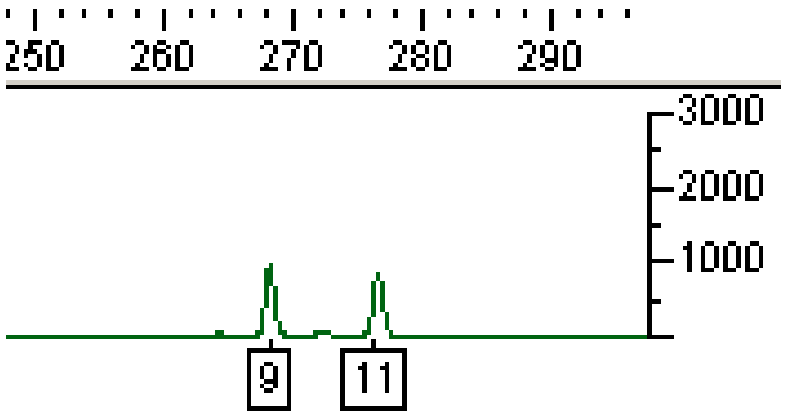
- All new kits tested have been completely concordant with the certified values of all markers for each component

- One exception: MiniFiler
  - Genomic 8 with D16S539

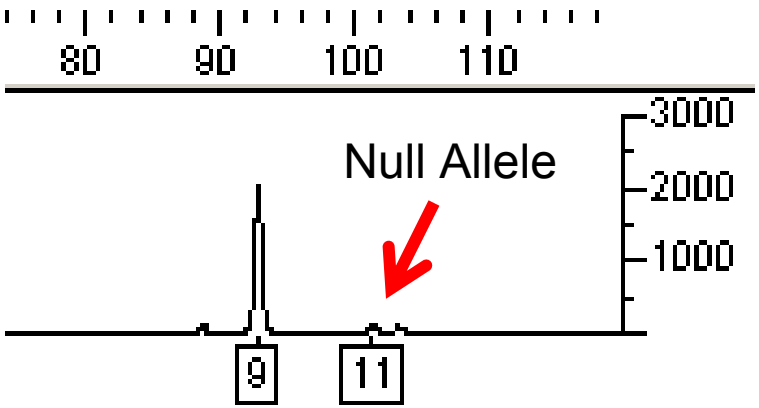
SRM 2391b Genomic 8 with D16S539

All allele calls with MiniFiler for CSF1PO, D7S820, D13S317, D18S51, D21S11, FGA, and D16S539 (with the exception noted below) ***match previously certified values.***

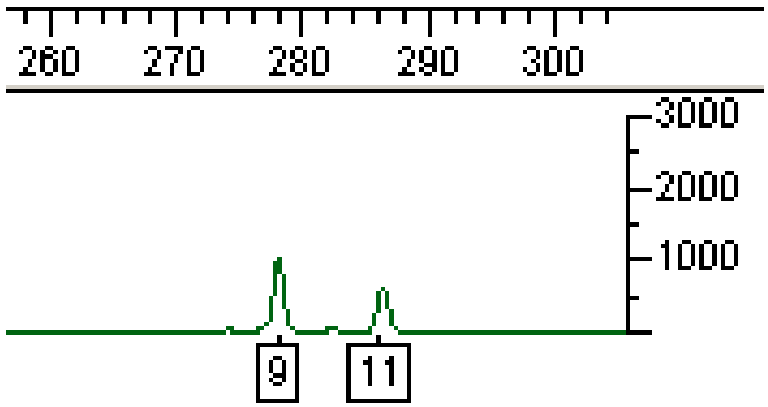
Identifiler



MiniFiler



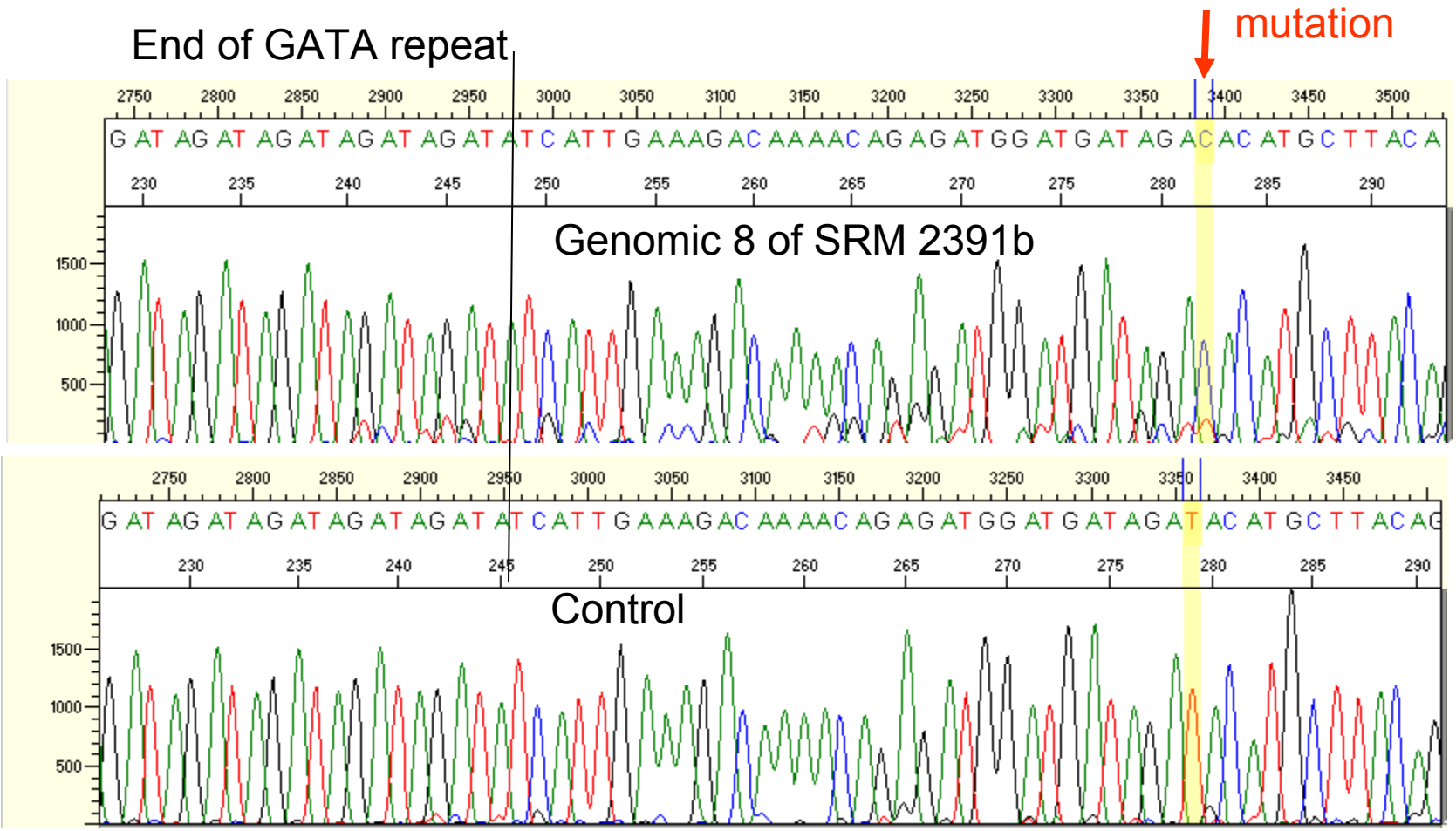
PowerPlex 16



***Due to primer binding site mutation**

D16S539 SRM 2391b Genomic 8

T→C mutation 34 bp downstream of the repeat



Position of the T→C probably affects the reverse primer of Minifiler and is the 3rd base found the 5'end of the Reverse PP16 primer. This could explain the imbalance of the allele seen when using PP16.

Current and Future Concordance Studies

Current Concordance Studies

- SRM 2391c (PCR-based DNA Profiling Standard) is currently being certified for all loci in the new kits (including European loci)
  - Genotype and concordance results have been obtained using ALL available STR multiplex kits, including newly released kits
  - All types are concordant thus far

Future Concordance Studies

- Rapid PCR
  - Vallone, P.M., et al. (2008). Demonstration of rapid multiplex PCR amplification involving 16 genetic loci. *FSI: Genetics, 3, 42-45.*
  - First available rapid PCR kit is **PowerPlex 18D**

- Will there be greater susceptibility to primer binding site mutations with rapid PCR?
  - Concordance testing is STILL important

PP18D Study
New Promega Kit for **Direct** & **Rapid** PCR

- 15 STRs present in Promega PP16 + 2 Identifiler loci
  - **D2S1338 & D19S433**
- Initial testing with 50 anonymous blood samples
  - Blood stored on FTA and 903 paper
- Concordance typing performed with PP16HS and Identifiler
- All comparisons concordant except for
  - One sample at D2S1338 (null allele?)
    - PP18D = 17,23
    - Identifiler = 23,23

DNA sequencing is underway to confirm this result

Thermal Cycling Protocol[1]
96°C for 2 minutes, then:
94°C for 10 seconds 60°C for 1 minute for 27 cycles, then: ~1.5 h
60°C for 20 minutes
4°C soak

PP18D work performed by Peter Vallone and Erica Butts

Summary &
Final Thoughts

Conclusions

- Concordance testing is valuable when different sets of primers are used to amplify the same markers
- Null alleles and discordant results are reported on STRBase:

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

- NIST plays an important role in concordance testing to aid the community
  - SRM 2391b concordance
  - Several null alleles have been fixed before the final release of new STR multiplex kits

Acknowledgments

NIST Funding: Interagency Agreement 2008-DN-R-121 between the National Institute of Justice and NIST Office of Law Enforcement Standards

NIST Disclaimer: 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.

Points of view are mine and do not necessarily represent the official position or policies of the US Department of Justice or the National Institute of Standards and Technology.

NIST Team for This Work

John Butler

Dave Duewer

Margaret Kline

Pete Vallone

A special thanks to Applied Biosystems, Promega, and Qiagen for providing the kits used in this study

Contact Info: becky.hill@nist.gov, 301-975-4275