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

NIST – GMI Seminar
Innsbruck, Austria
September 5, 2011

Outline of Topics to Discuss

- Introduction and importance of concordance testing
  - Overlapping markers with different primer configurations
- NIST role in concordance testing
  - SRM 2391b/2391c 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?

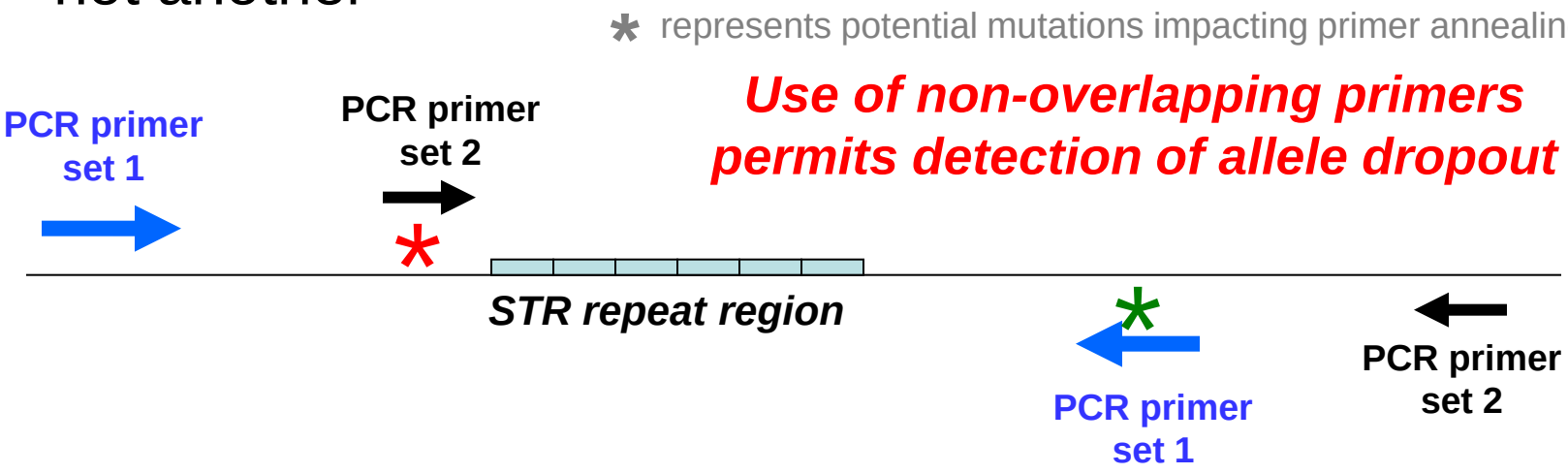
Importance of Concordance Testing

- 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

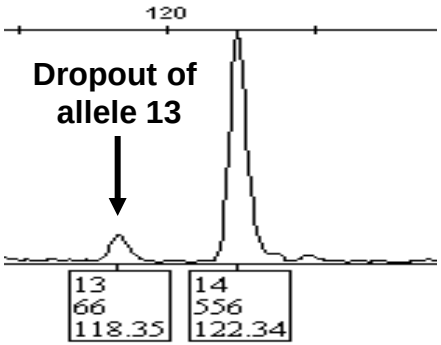


If no primer binding site mutations

If a primer binding site mutation exists

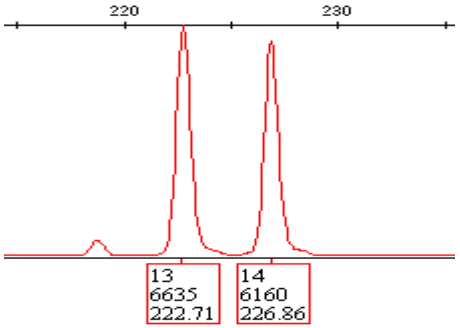
Example Primer Binding Site Mutation that Causes a Null Allele

Identifiler = 14,14



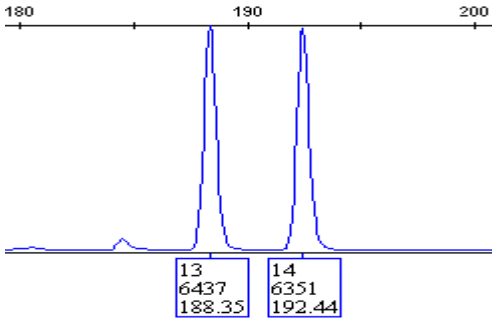
PHR = 11.9%

PP ESX 17 = 13,14



PHR = 92.8%

PP ESI 17 = 13,14



PHR = 98.7%

D19S433 repeat region

G → A SNP

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/2391c (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

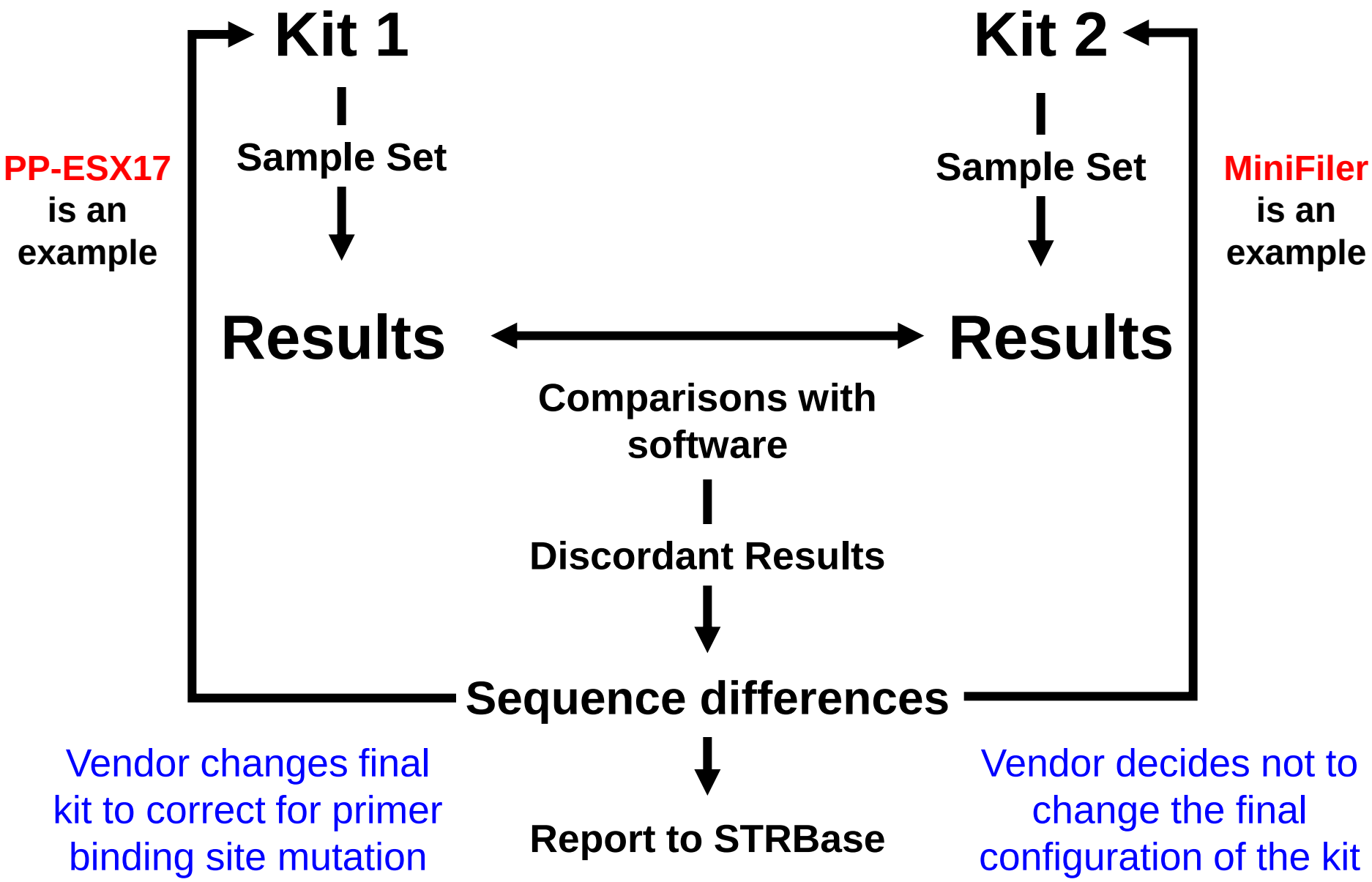
The 4 “S’s” of Concordance

- 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
- ESSplex SE

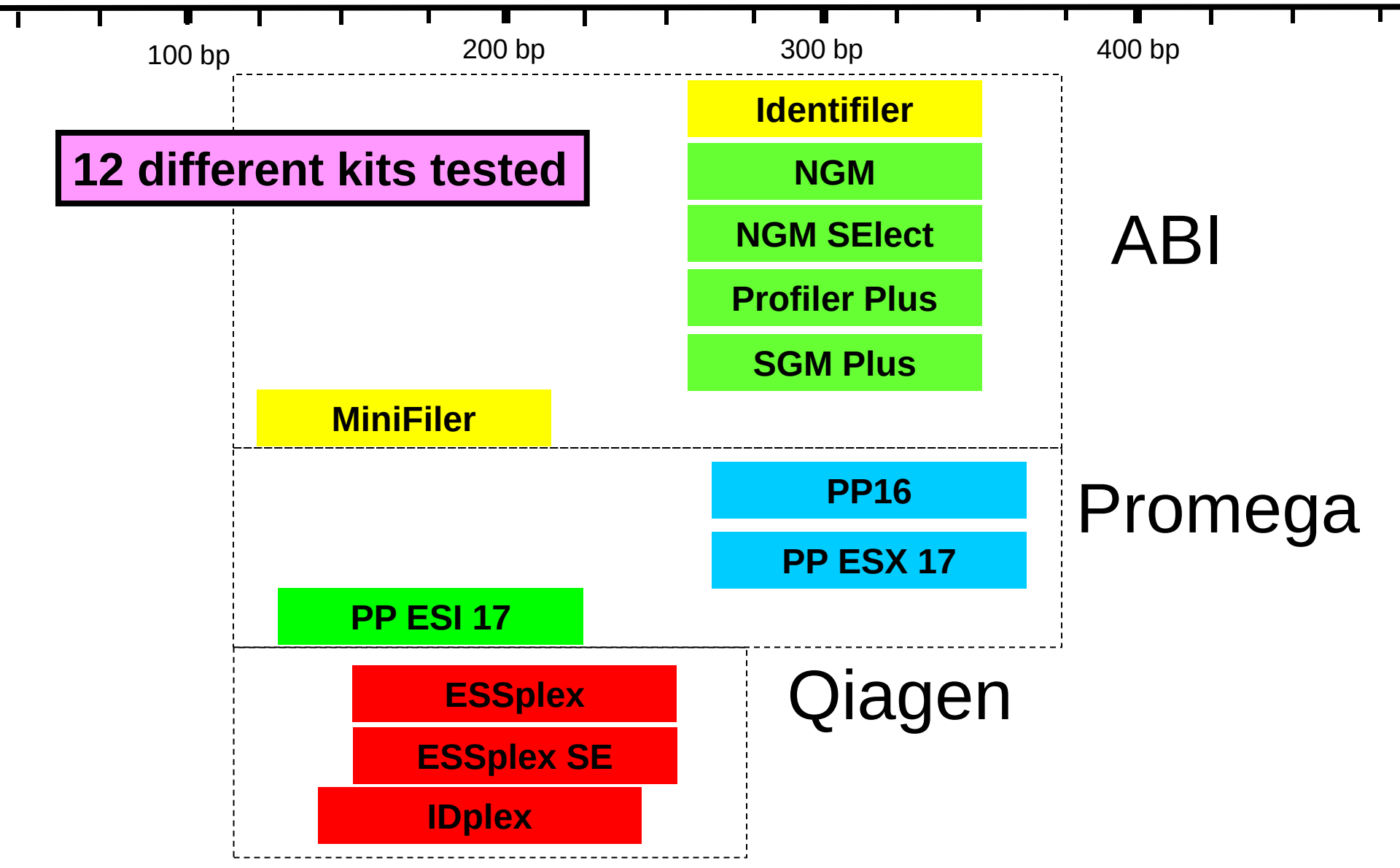
What samples are used at NIST to perform concordance testing?

NIST Sample Set (>1450 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
- **NIST SRM 2391c**, PCR-based DNA Profiling Standard
  - 4 genomic DNA (one mixture)
  - 2 cell lines (903 and FTA paper)

What are the results from the completed concordance studies?

D18S51 Concordance Checking

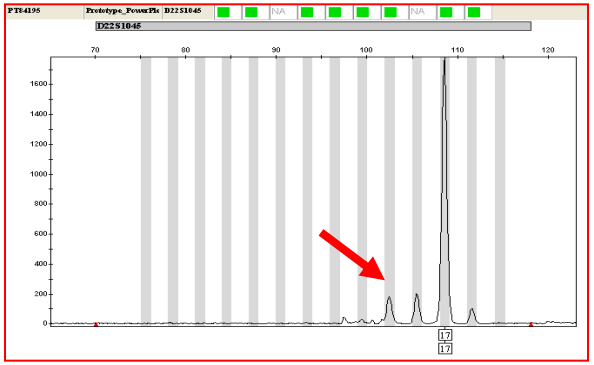


Primer Set Compared

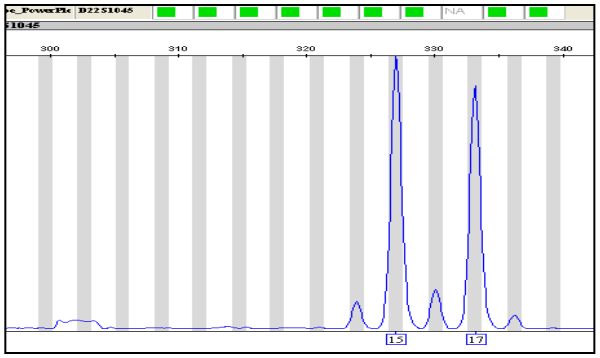
Marker	# of Sets	Marker	# of Sets
Amelogenin	**13**	D2S441	**9**
D18S51	**12**	D19S433	**9**
D21S11	**12**	D1S1656	**7**
FGA	**12**	D12S391	**7**
D3S1358	**11**	SE33	**5**
TH01	**11**	D5S818	**4**
D16S539	**11**	D7S820	**4**
vWA	**11**	D13S317	**4**
D8S1179	**11**	TPOX	**3**
D2S1338	**10**	CSF1PO	**4**
D10S1248	**9**	Penta D	**1**
D22S1045	**9**	Penta E	**1**

D22S1045 Discordance

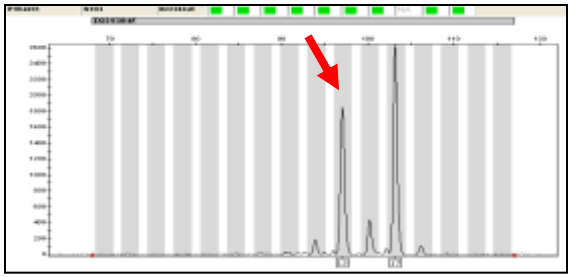
ESX 17 (prototype) = 17,17



ESI 17 (prototype) = 15,17

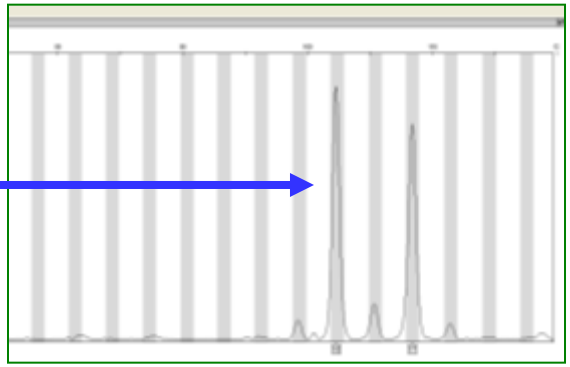


NIST NC01 = 15,17

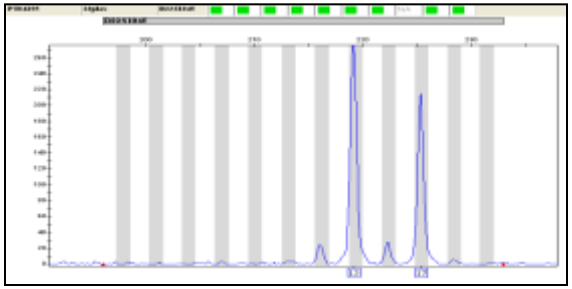


Destabilized some (but uses lower annealing temperature with fewer amplicons in multiplex)

ESX 17 (final) = 15,17



NIST 23plex = 15,17

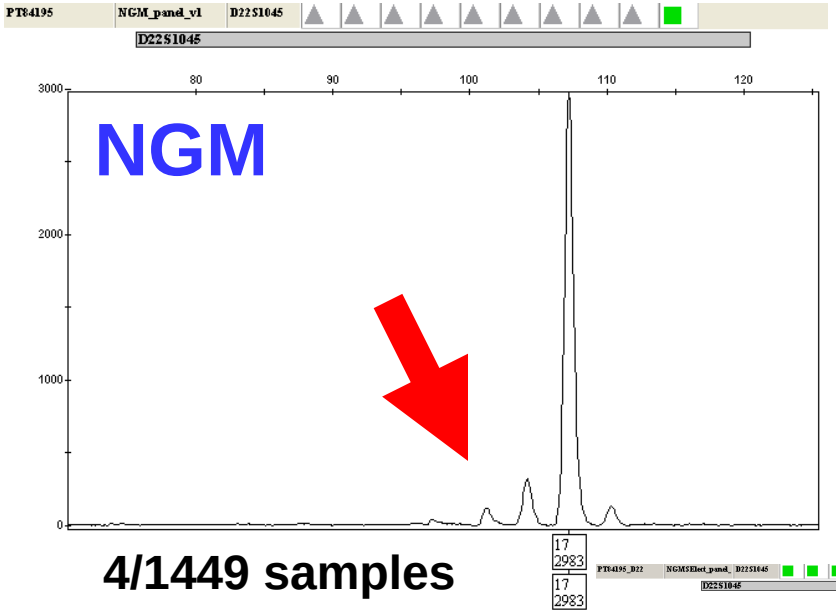


NIST PT84195

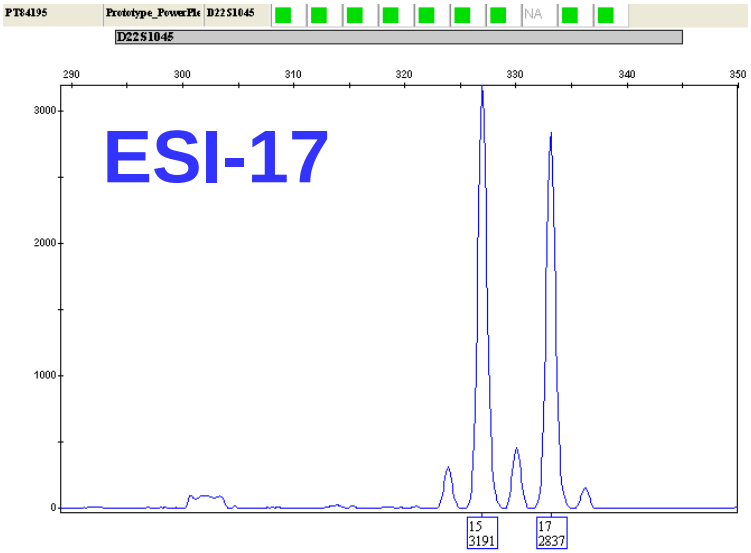
G→T 15 bp upstream impacting forward primer binding with ESX17

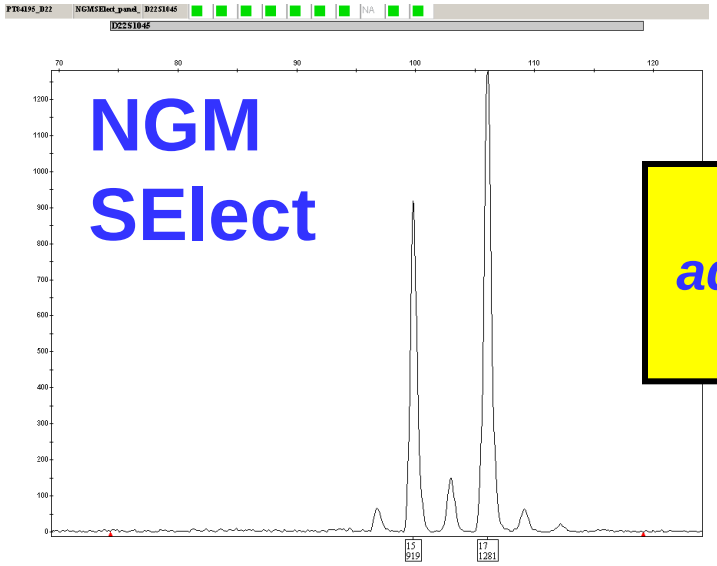
Promega added additional primer to correct issue

D22S1045 Null Allele



4/1449 samples



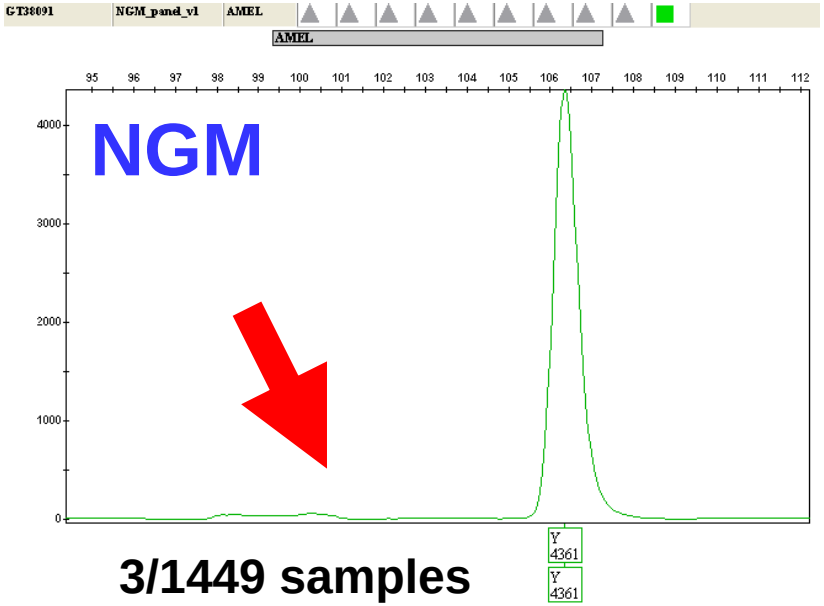


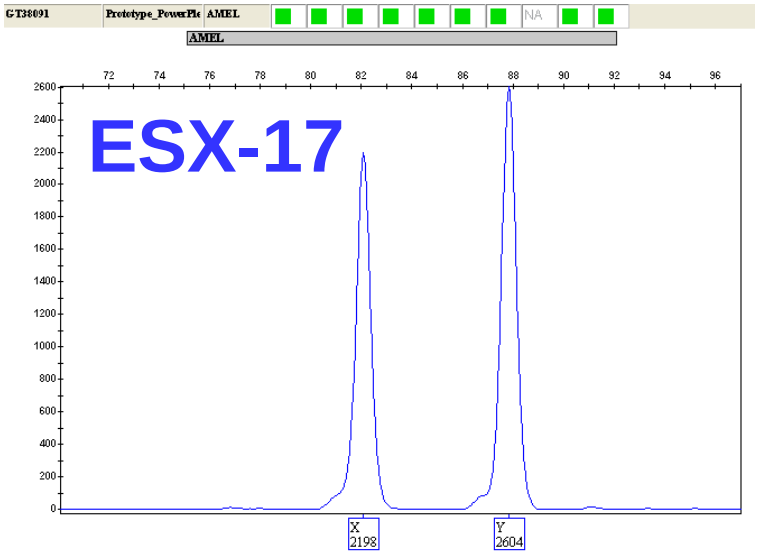
Correct type (15,17)

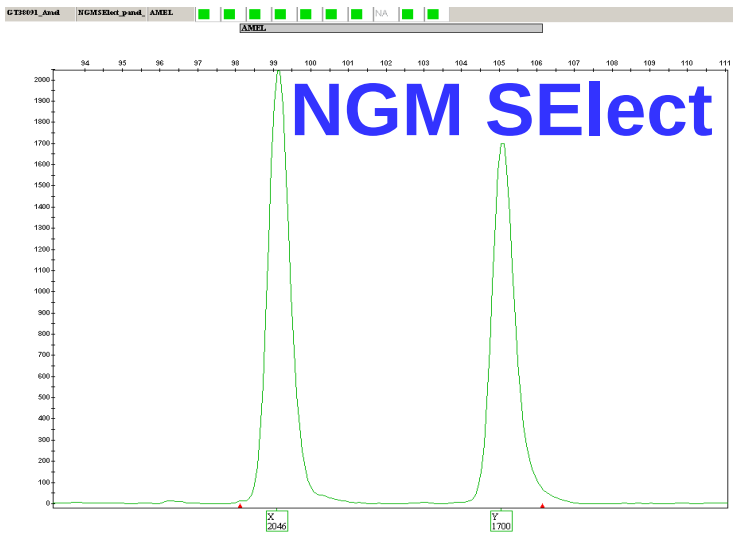
ABI added an additional primer to correct issue

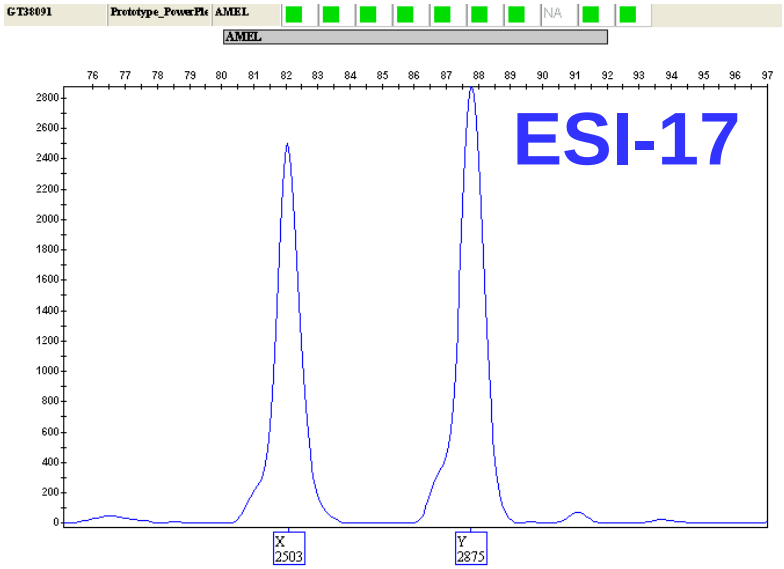
G→T 15 bp upstream impacting forward primer binding with NGM

Amelogenin X Null Allele

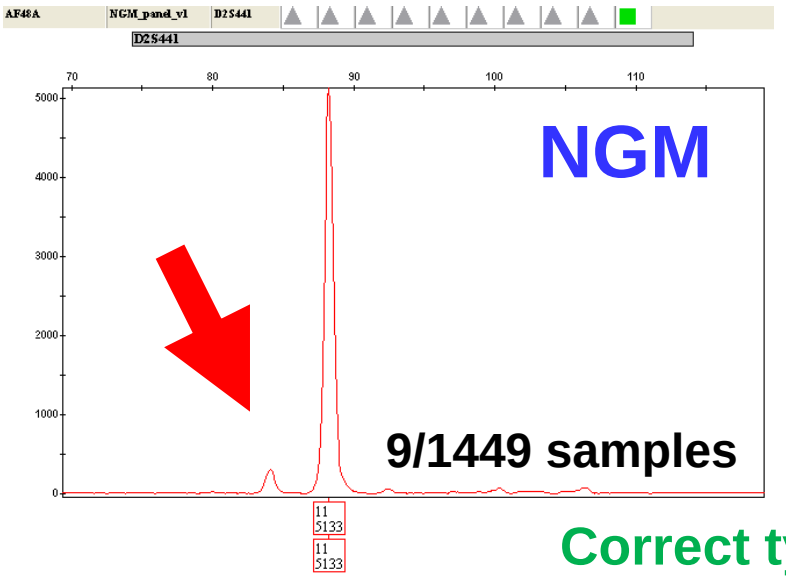


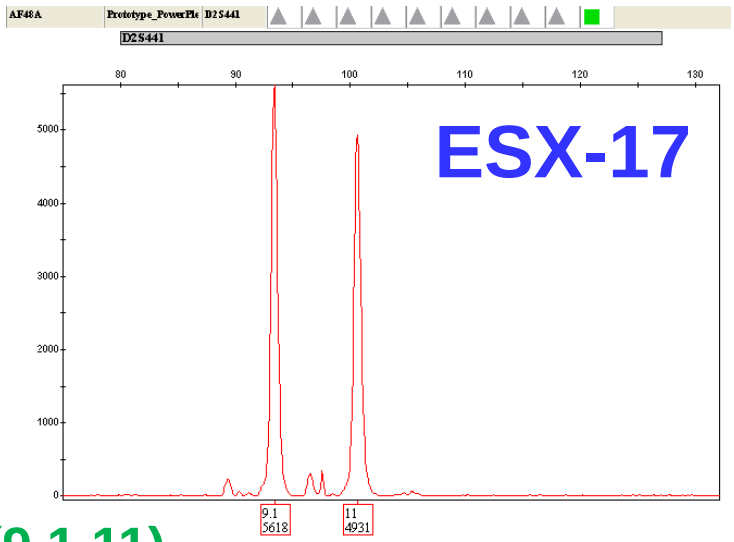




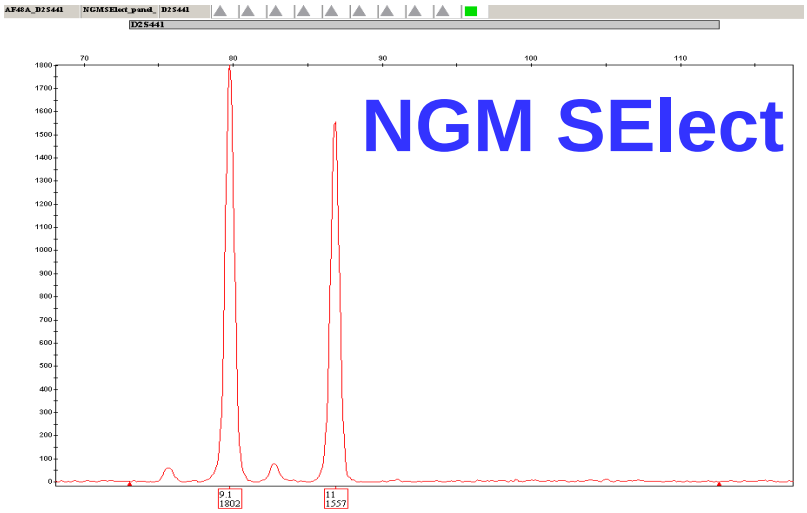


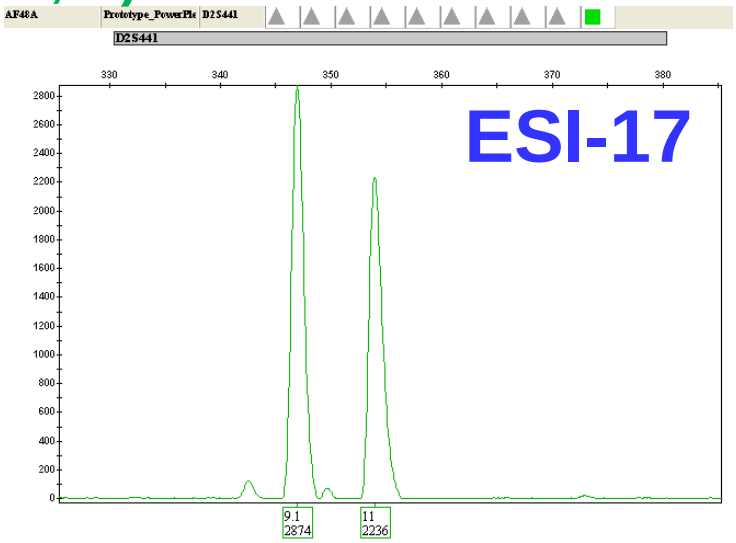
D2S441 Null Allele





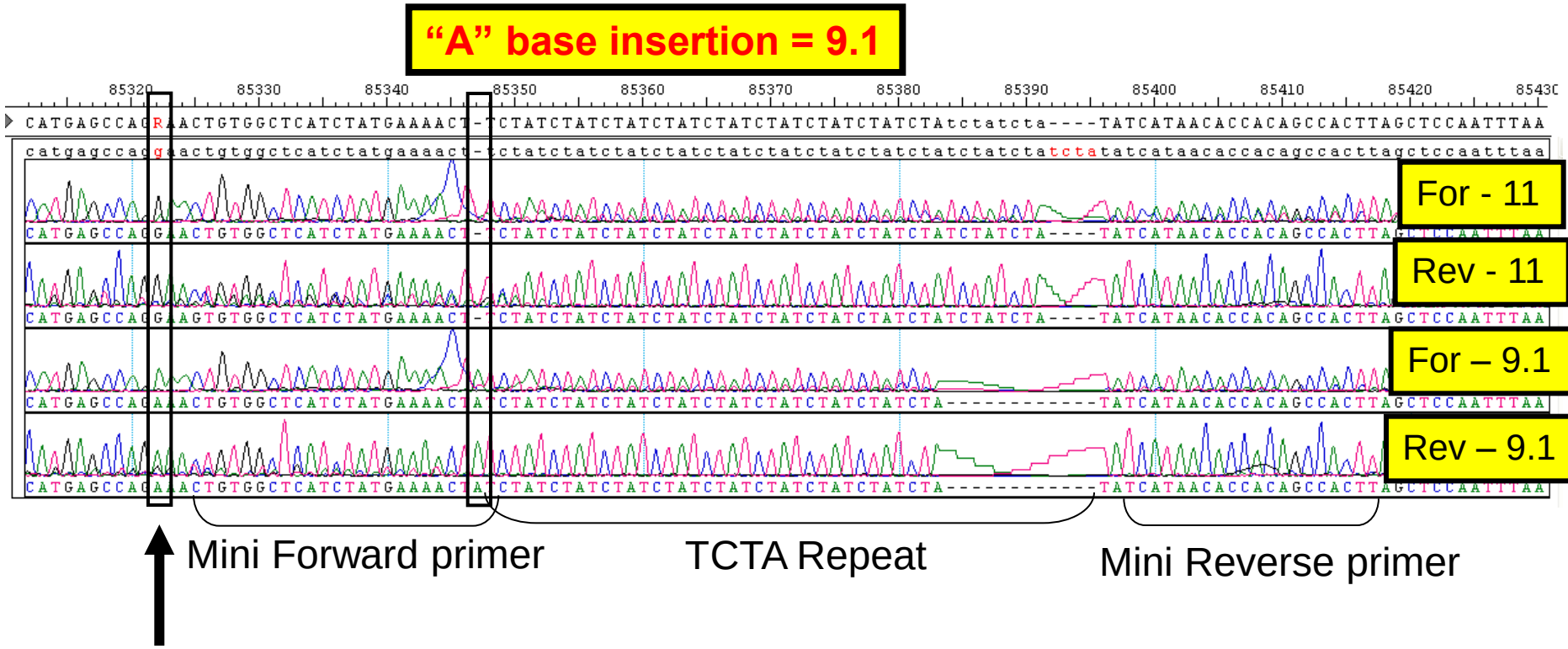
Correct type (9.1,11)





8/9 null alleles were from Asian samples

D2S441 Sequencing

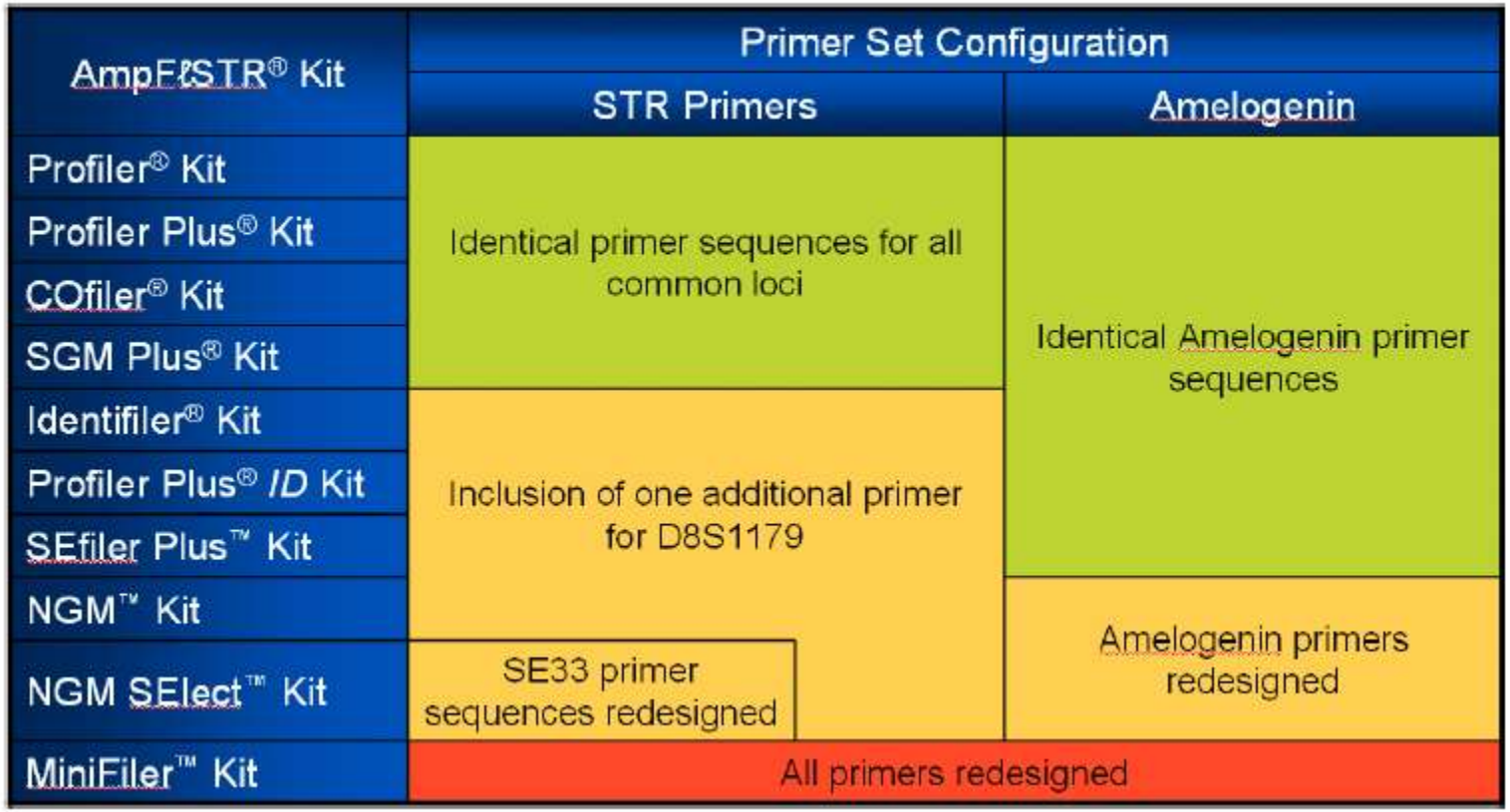


G → A SNP 25/26 bp upstream of the repeat

True Genotype = 9.1,11

NGM Genotype = 11,11

Primer Changes with ABI Kits

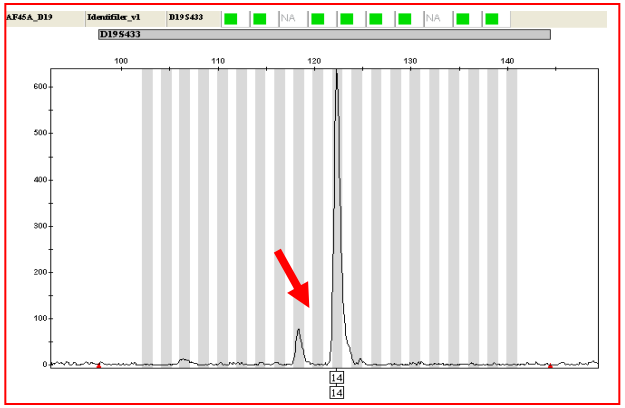


D2S441 and D22S1045 have an additional primer in NGM and NGM SElect

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

D19S433 Discordance

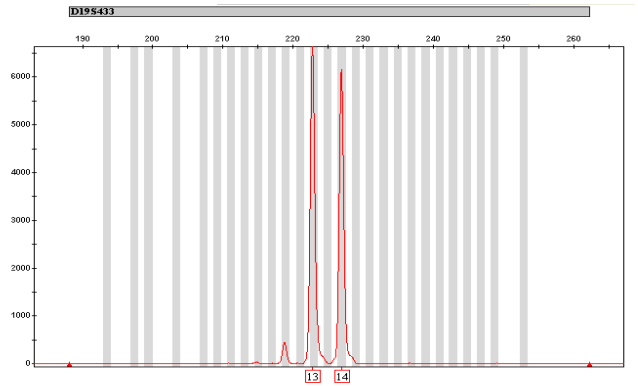
Identifiler & NGM = 14,14



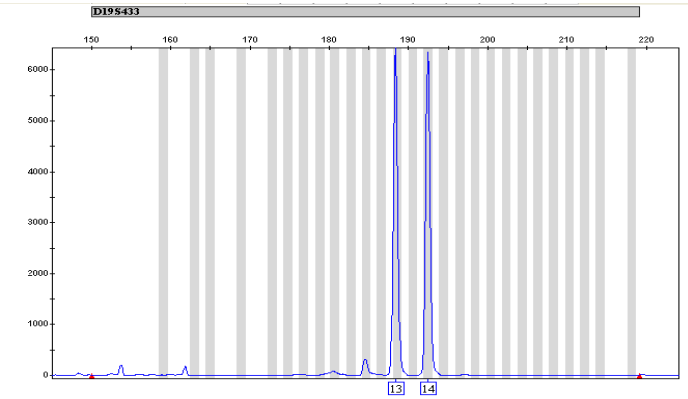
AF45A (Asian)

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

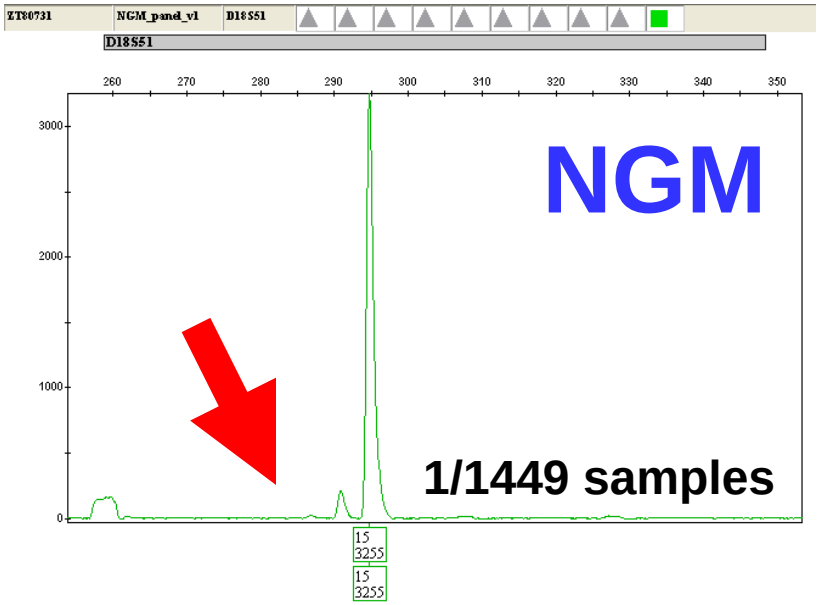
J Forensic Sci, September 2008, Vol. 53, No. 5
doi: 10.1111/j.1556-4029.2008.00806.x
Available online at: www.blackwell-synergy.com

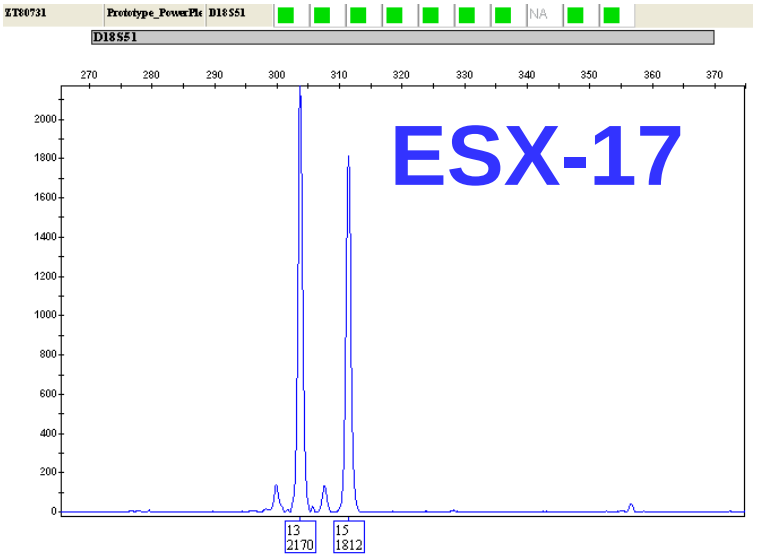
Natsuko Mizuno,[1] D.V.M.; Tetsushi Kitayama,[1] M.Sc.; Koji Fujii,[1] Ph.D.; Hiroaki Nakahara,[1] D.V.M.; Kanako Yoshida,[1] Ph.D.; Kazumasa Sekiguchi,[1] Ph.D.; Naoto Yonezawa,[2] Ph.D.; Minoru Nakano,[2] Ph.D.; and Kentaro Kasai,[1] Ph.D.

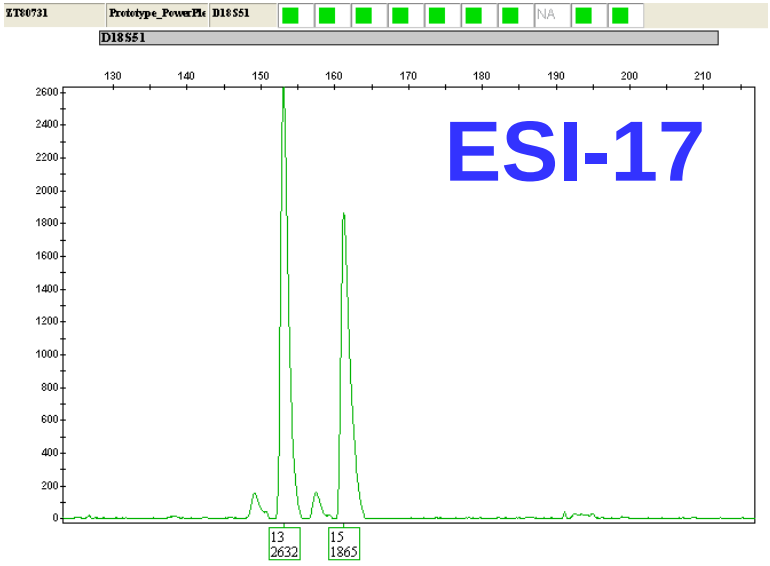
A D19S433 Primer Binding Site Mutation and the Frequency in Japanese of the Silent Allele It Causes

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

D18S51 Null Allele



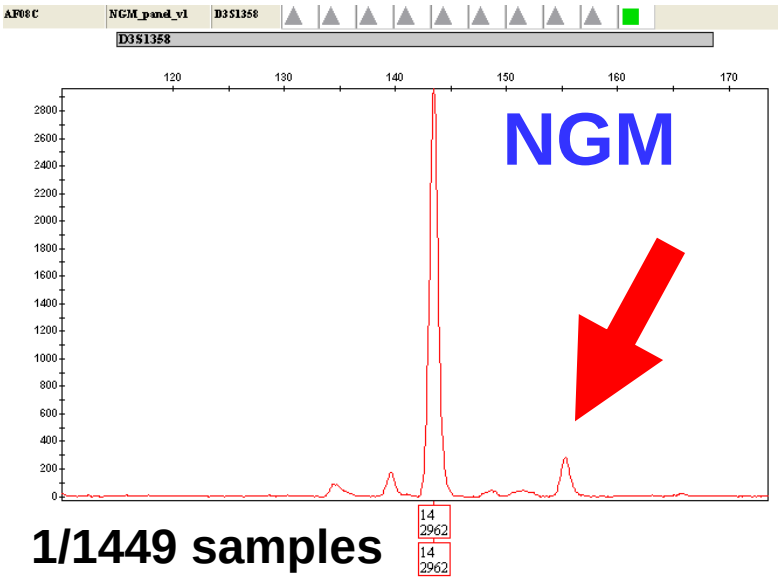




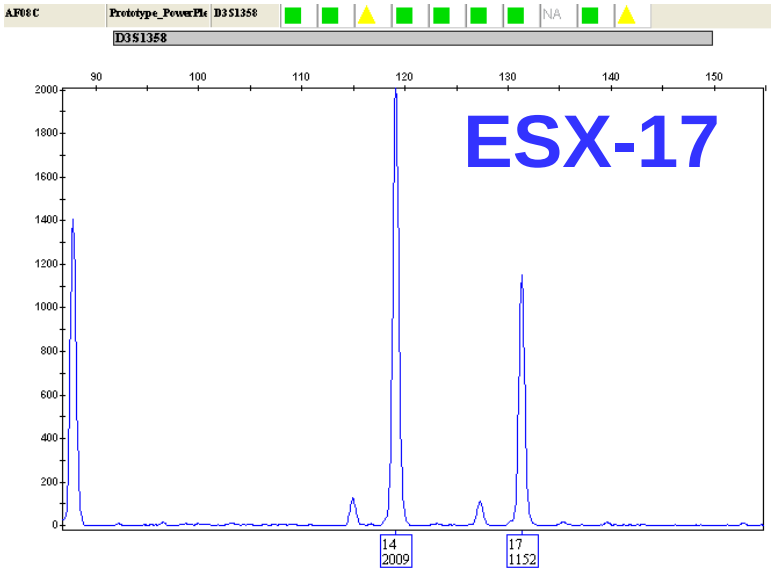
Correct type (13,15)

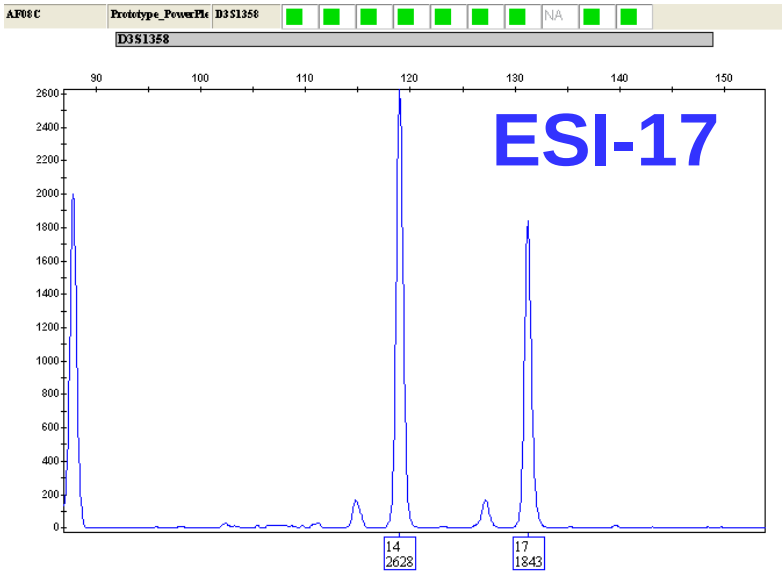
C→T SNP 172 bp downstream from repeat

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 (%)
111	102,345	1,021	948,301	1,109	99.883

948,301 allele comparisons

1,109 total differences

99.88% concordance

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

Kits compared	Samples	Loci Compared	Comparisons	# Differences	Concordance (%)
ID-SGM+	1424	11	15,664	1	99.994
ID-Pro+	1415	10	14,150	1	99.993
ID-IDplex	1426	16	22,816	29	99.873
ID-PP16	662	14	9,268	4	99.957
ID-MiniFiler	1137	9	10,233	26	99.746
ID-NGM	1437	11	15,807	3	99.981
ID-NGMs	663	11	7,293	0	100.000
ID-ESX17	1443	11	15,873	5	99.968
ID-ESI17	1443	11	15,873	4	99.975
ID-ESSplex	1433	11	15,763	28	99.822
ID-ESSplexSE	662	11	7,282	17	99.767
ID-Hexaplex	653	2	1,306	1	99.923
PP16-SGM+	651	9	5,859	1	99.983
PP16-Pro+	647	10	6,470	2	99.969
PP16-IDplex	657	14	9,198	3	99.967
PP16-MiniFiler	656	8	5,248	14	99.733
PP16-NGM	657	9	5,913	3	99.949
PP16-NGMs	662	9	5,958	1	99.983
PP16-ESX17	662	9	5,958	1	99.983
PP16-ESI17	662	9	5,958	0	100.000
PP16-ESSplex	653	9	5,877	16	99.728
PP16-ESSplexSE	662	9	5,958	16	99.731
PP16-Hexaplex	653	2	1,306	1	99.923
SGM+-Pro+	1415	7	9,905	0	100.000
SGM+-IDplex	1424	11	15,664	5	99.968
SGM+-MiniFiler	1137	6	6,822	10	99.853
SGM+-NGM	1424	11	15,664	4	99.974
SGM+-NGMs	651	11	7,161	0	100.000
SGM+-ESX17	1424	11	15,664	6	99.962
SGM+-ESI17	1424	11	15,664	5	99.968
SGM+-ESS	1424	11	15,664	5	99.968
SGM+-ESSplexSE	651	11	7,161	5	99.930
SGM+-Hexaplex	651	2	1,302	1	99.923
Pro+-IDplex	1415	10	14,150	5	99.965
Pro+-MiniFiler	1137	6	6,822	16	99.765
Pro+-NGM	1415	7	9,905	4	99.960
Pro+-NGMs	647	7	4,529	0	100.000
Pro+-ESX17	1415	7	9,905	4	99.960
Pro+-ESI17	1415	7	9,905	3	99.970
Pro+-ESS	1415	7	9,905	4	99.960
Pro+-ESSplexSE	647	7	4,529	4	99.912
Pro+-Hexaplex	647	1	647	1	99.845
IDplex-MiniFiler	1137	9	10,233	48	99.531
IDplex-NGM	1426	11	15,686	30	99.809
IDplex-NGMs	657	11	7,227	17	99.765
IDplex-ESX17	1426	11	15,686	28	99.821
IDplex-ESI17	1426	11	15,686	27	99.828
IDplex-ESS	1426	11	15,686	1	99.994
IDplex-ESSplexSE	657	11	7,227	1	99.986
IDplex-Hexaplex	653	2	1,306	1	99.923
MiniFiler-NGM	1137	6	6,822	13	99.809
MiniFiler-NGMs	656	6	3,936	10	99.746
MiniFiler-ESX17	1137	6	6,822	10	99.853
MiniFiler-ESI17	1137	6	6,822	9	99.868
MiniFiler-ESS	1137	6	6,822	35	99.487
MiniFiler-ESSplexSE	656	6	3,936	35	99.111
MiniFiler-Hexaplex	653	1	653	1	99.847
NGM-NGMs	657	16	10,512	14	99.867
NGM-ESX17	1437	16	22,992	16	99.930
NGM-ESI17	1437	16	22,992	18	99.922
NGM-ESS	1433	16	22,928	42	99.817
NGM-ESSplexSE	657	16	10,512	22	99.791
NGM-Hexaplex	653	7	4,571	9	99.803
NGMs-ESX17	662	17	11,254	4	99.964
NGMs-ESI17	662	17	11,254	14	99.876
NGMs-ESS	653	16	10,448	17	99.837
NGMs-ESSplexSE	662	17	11,254	34	99.698
NGMs_Hexaplex	653	7	4,571	3	99.934
ESX17-ESI17	1443	17	24,531	19	99.923
ESX17-ESS	653	16	10,448	34	99.675
ESX17-ESSplexSE	662	17	11,254	25	99.778
ESX17-Hexaplex	657	7	4,599	6	99.870
ESI17-ESS	653	16	10,448	28	99.732
ESI17-ESSplexSE	662	17	11,254	30	99.733
ESI17-Hexaplex	657	7	4,599	3	99.935
ESS-ESSplexSE	653	16	10,448	0	100.000
ESS-Hexaplex	653	7	4,571	3	99.934
ESSplexSE-Hexaplex	653	7	4,571	3	99.934
SE33-ESX17	1443	1	1,443	6	99.584
SE33-ESI17	1443	1	1,443	17	98.822
SE33-NGMs	663	1	663	4	99.397
SE33-ESSplexSE	662	1	662	21	96.828
ESI17p-ESX17	477	17	8,109	7	99.914
ESI17p-NGMs	477	17	8,109	2	99.975
ESI17p-ESSplexSE	477	17	8,109	42	99.482
ESI17p-SE33	477	1	477	4	99.161
PP18D-ID	50	16	800	2	99.750
PP18D-PP16	703	16	11,248	1	99.991
ESX17*/ESX17	1443	17	24531	4	99.984
ESX17*/ESI17p	477	17	8109	3	99.963
ESX17*/NGM	1437	16	22992	22	99.904
ESX17*/NGMs	663	17	11271	4	99.965
ESX17*/ESS	1433	16	22928	30	99.869
ESX17*/ESSplexSE	662	17	11254	44	99.609
ESX17*/Hexaplex	653	7	4571	2	99.956
26plex/ESX17	1443	3	4329	4	99.908
26plex/ESI17	1443	3	4329	0	100.000
26plex/NGM	1437	3	4311	11	99.745
26plex/NGMs	663	3	1989	0	100.000
26plex/ESS	1433	3	4299	0	100.000
26plex/ESSplexSE	662	3	1986	0	100.000
26plex/Hexaplex	653	3	1959	2	99.898
26plex/ESX17*	663	3	1989	0	100.000
miniSTRs/ESX17	663	3	1989	3	99.849
miniSTRs/ESI17	663	3	1989	0	100.000
miniSTRs/NGM	657	3	1971	3	99.848
miniSTRs/NGMs	663	3	1989	0	100.000
miniSTRs/ESS	653	3	1959	0	100.000
miniSTRs/ESSplexSE	662	3	1986	0	100.000
miniSTRs/Hexaplex	653	3	1959	2	99.898
miniSTRs/ESX17*	663	3	1989	0	100.000
Totals	102345	1021	948301	1109	99.883

Was there complete concordance with SRM 2391b and SRM 2391c?

SRM 2391b/2391c PCR-Based Profiling Standard

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

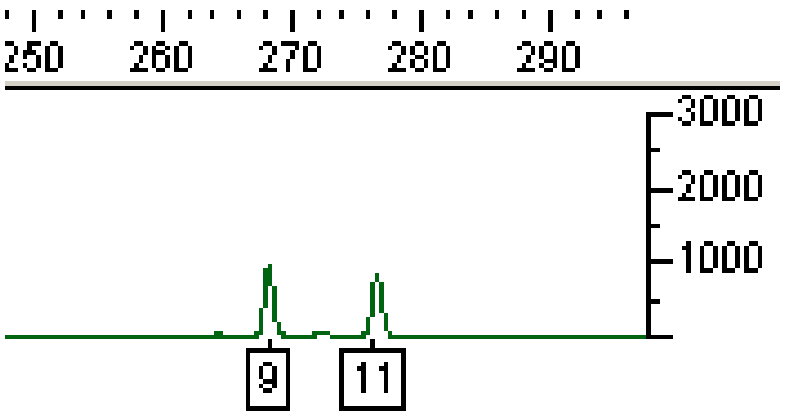
- All new kits tested have been completely concordant with the certified values of all markers for each component for SRM 2391b and 2391c

- One exception for SRM 2391b: MiniFiler
  - Genomic 8 with D16S539

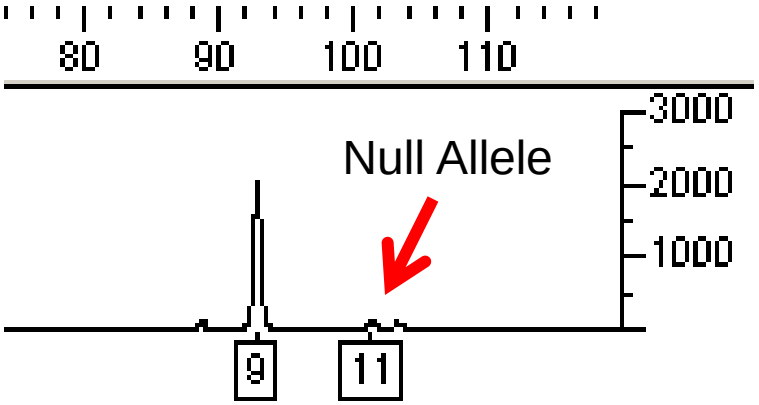
SRM 2391b Genomic 8 with D16S539

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

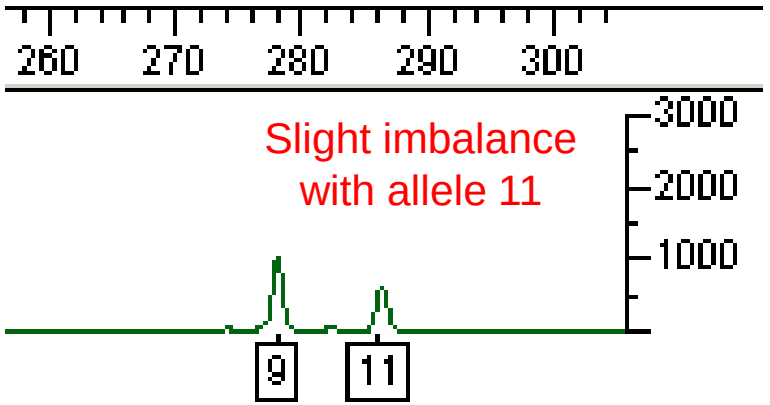
Identifiler



MiniFiler



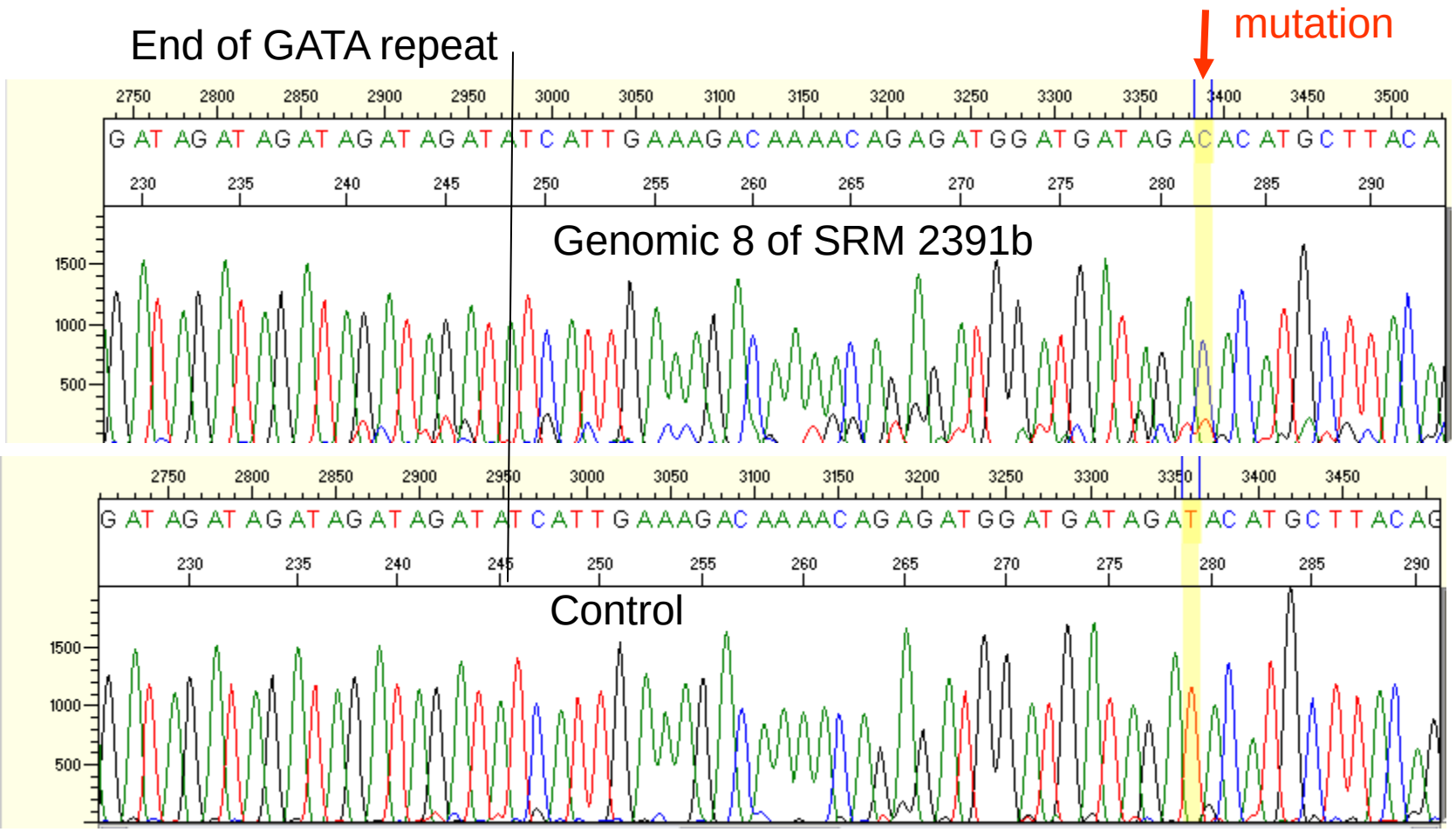
PowerPlex 16



***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.

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/2391c 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