ISHI Workshop on New Loci and Kits
October 10, 2013 (Atlanta, GA)
New Autosomal and Y-STR Loci and Kits: Making Data Driven Decisions

NIST Studies:

Kit Concordance and U.S. Population Data

Carolyn R. (Becky) Hill
NIST Applied Genetics Group

NIST Applied Genetics

NIST

Product Disclaimer

- **I will mention commercial STR kit names and information, but I am in no way attempting to endorse any specific products.**

- **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 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 of the National Institute of Standards and Technology or the U.S. Department of Justice. **Our group receives or has received funding from the FBI Laboratory and the National Institute of Justice.**

Presentation Outline

- STR kits (including Fusion and GlobalFiler)
- NIST U.S. population samples
- Concordance study results
- SRM 2391c sequencing results

Autosomal STR Loci

Autosomal STR Kits

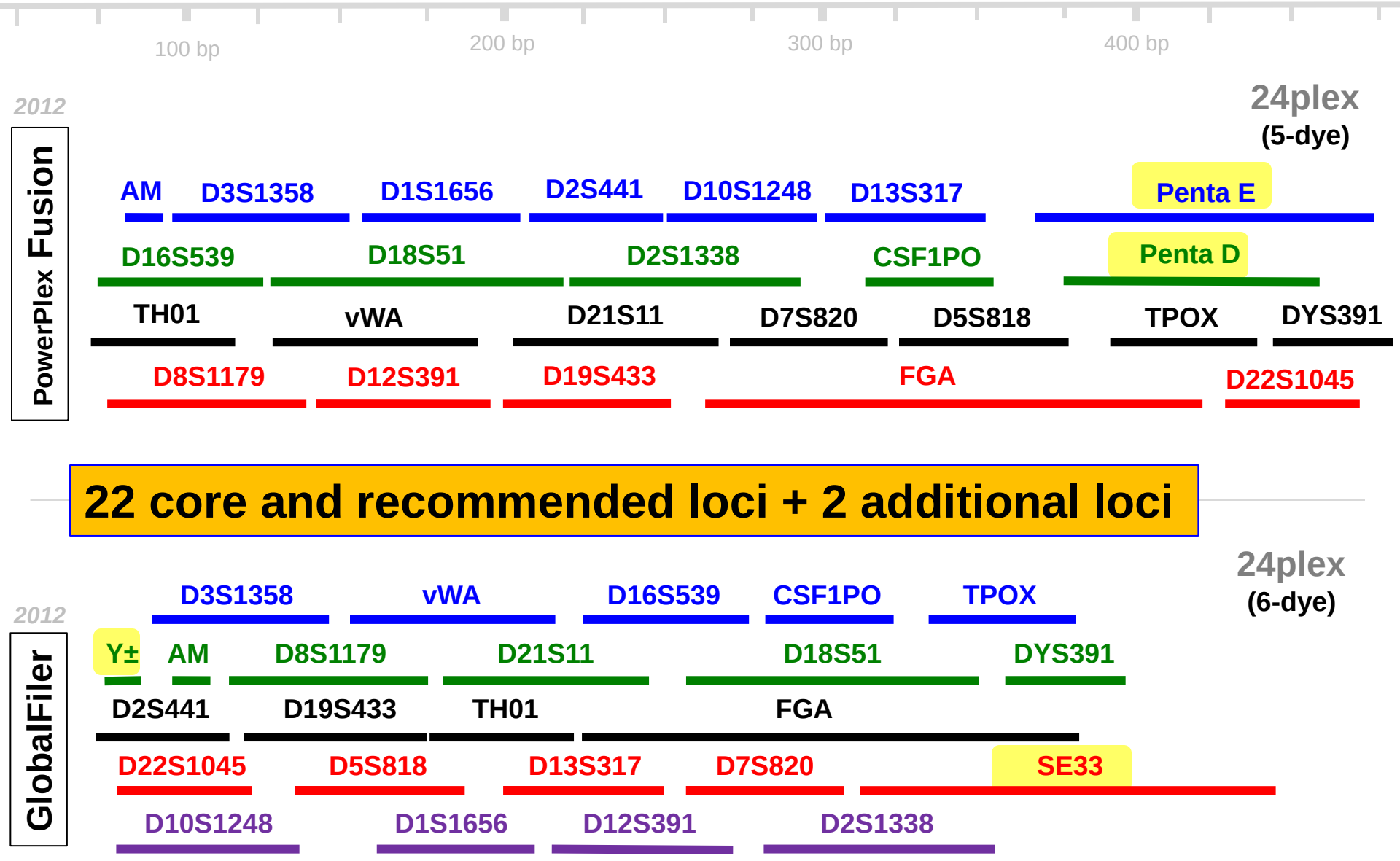
Chr	STR Locus	Repeat	Allele Range (Butler et al. 2012)	CODIS 13 (US 1997-present)	CODIS 20 (US future)	ESS 12 (EU 2009-present)
				required		
1q31	F13B	AAAT	6 to 11			
1q42	D1S1656	TAGA	10 to 19.3		■	■
2p25.3	**TPOX**	AATG	5 to 13	■		
2p14	D2S441	TCWA	8 to 17		■	■
2q35	D2S1338	TKCC	15 to 27		■	
3p21.31	**D3S1358**	TCTR	11 to 20	■	■	■
4q31.3	**FGA**	YTYY	16.2 to 43.2	■	■	■
5q23.2	**D5S818**	AGAT	7 to 15	■	■	
5q33.1	**CSF1PO**	AGAT	7 to 15	■	■	
6p24	F13A01	AAAG	3.2 to 17			
6q14	SE33	AAAG	6.3 to 36			
6q15	D6S1043	AGAY	8 to 26			
7q21.11	**D7S820**	GATA	6 to 14	■	■	
8p22	LPL	AAAT	7 to 15			
8q24.13	**D8S1179**	TCTR	8 to 18	■	■	■
9p13	Penta C	AAAAC	5 to 16			
10q26.3	D10S1248	GGAA	8 to 19		■	■
11p15.5	**TH01**	TCAT	5 to 11	■	■	■
12p13.31	**vWA**	TCTR	11 to 21	■	■	■
12p13.2	D12S391	AGAY	14 to 27		■	■
13q31.1	**D13S317**	TATC	8 to 15	■	■	
15q25	FESFPS	ATTT	5 to 14			
15q26.2	Penta E	AAAGA	5 to 25			
16q24.1	**D16S539**	GATA	5 to 15	■	■	
18q21.33	**D18S51**	AGAA	9 to 28	■	■	■
19q12	D19S433	WAGG	9 to 18.2		■	
21q21.1	**D21S11**	TCTR	24.2 to 39	■	■	■
21q22.3	Penta D	AAAGA	2.2 to 17			
22q12.3	D22S1045	ATT	8 to 19			■
Xp, Yp	Amelogenin	--	--		■	
Yq11.21	DYS391	TCTA	7 to 13		■	

Promega STR kits						
PowerPlex 16 (HS)	PowerPlex 18D	PowerPlex ESI/ESX 16	PowerPlex ESI/ESX 17	PowerPlex 21	PowerPlex CS7	**PowerPlex Fusion**
					■	
		■	■	■		■
■	■			■		■
		■	■			■
	■	■	■	■		■
■	■	■	■	■		■
■	■	■	■	■		■
■	■			■		■
■	■			■		■
					■	
			■			
				■		
■	■			■		■
					■	
■	■	■	■	■		■
					■	
		■	■	■		■
■	■	■	■	■		■
■	■	■	■	■		■
		■	■	■		■
■	■			■	■	
					■	
					■	
■	■	■	■	■		■
■	■	■	■	■		■
		■	■	■		■
■	■	■	■	■		■
					■	
		■	■	■		■
■	■	■	■	■		■
					■	
■	■	■	■	■		■

Life Technologies (ABI) STR kits										
Profiler Plus	COfiler	SGM Plus	SEfiler Plus	SinoFiler	MiniFiler	**Identifiler (Plus)**	VeriFiler	NGM	NGM SElect	**GlobalFiler**
							■	■	■	■
	■					■				■
							■	■	■	■
		■	■	■	■	■	■	■	■	■
■		■	■	■		■		■	■	■
■		■	■	■	■	■		■	■	■
	■			■	■	■				■
			■						■	■
				■			■			
■	■			■	■	■				■
■		■	■	■		■		■	■	■
							■	■	■	■
	■	■	■			■	■	■	■	■
■		■	■	■		■		■	■	■
				■			■	■	■	■
					■	■				■
	■	■	■	■	■	■		■	■	■
■		■	■	■	■	■		■	■	■
		■	■	■	■	■	■	■	■	■
							■	■	■	■
										■

Qiagen STR kits					
ESSplex	ESSplex SE	Hexaplex ESS	Nonaplex ESS	Decaplex SE	IDplex
■	■	■	■		
					■
■	■	■	■		
■	■			■	■
■	■		■	■	■
■	■		■	■	■
					■
					■
	■		■	■	
					■
■	■		■	■	■
■	■			■	■
■	■		■	■	■
■	■			■	■
■	■	■	■		
■	■		■	■	■
■	■			■	■
■	■	■	■		
■	■			■	■
■	■	■	■		
					■

STR Marker Layouts for New U.S. Kits



GlobalFiler STR Kit

Launched Friday, September 14, 2012

Human Identification

- GlobalFiler™ Kit
- Go Faster
- Go Further
- Go Global
- Powered by 6-Dye™
- Human Identification Home

Introducing the world's most powerful STR kit

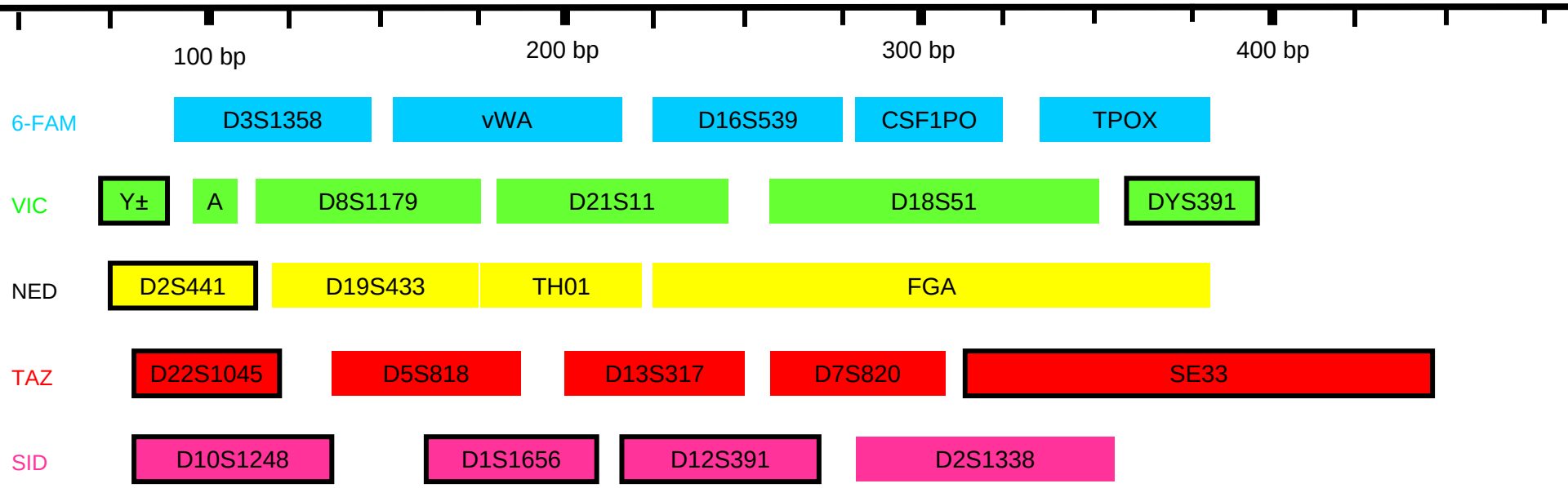
Around the world, forensic labs are being asked to do more with less. That's why the new GlobalFiler™ STR Kit combines reduced amplification time with maximum data recovery power. As part of the only fully integrated and validated forensic workflow, this breakthrough 6-dye, 24-loci technology is designed to deliver unprecedented lab performance. And, it's backed by Life Technologies best-in-class training, service, and support.

Go Faster ▸ | Go Further ▸ | Go Global ▸ | Powered by 6-Dye™ ▸

http://www.invitrogen.com/site/us/en/home/Products-and-Services/Applications/Human-Identification/globalfiler_str_kit.html

Life Technologies **GlobalFiler**

24plex



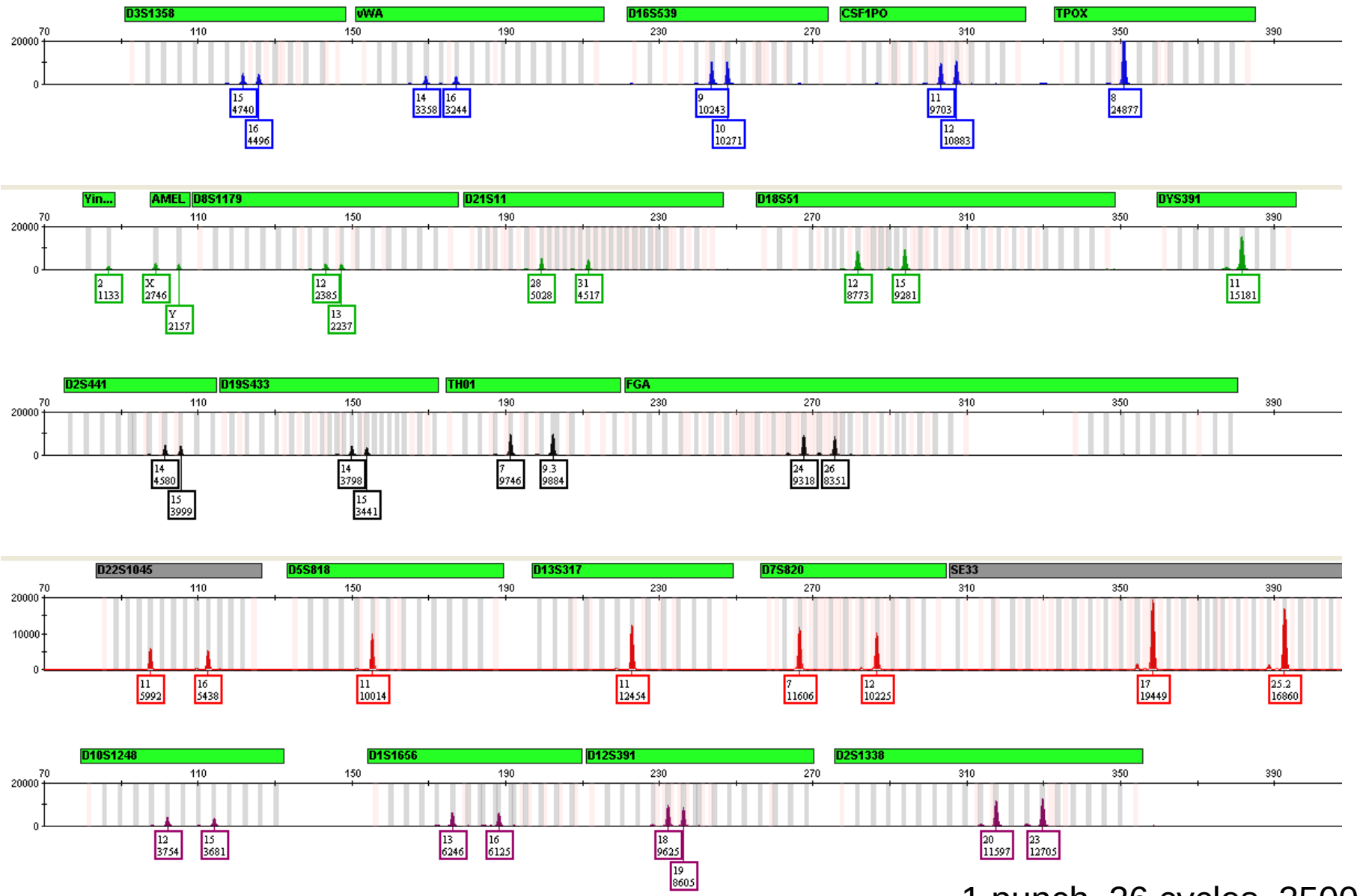
- 24 STR loci in 6 dyes (3500 use or 3130 upgrade required)
  - Includes SE33 and a Y-indel
- GlobalFiler Express: direct amplification capabilities
  - Single source samples: 40 min amplification
- GlobalFiler Casework (recently available)
  - Casework samples: 80 min amplification

Two separate kits

- GlobalFiler gives ~12 orders of magnitude improvement using the NIST 1036 data set

http://www.invitrogen.com/site/us/en/home/Products-and-Services/Applications/Human-Identification/globalfiler_str_kit.html

Positive Control 007



1 punch, 26 cycles, 3500

PowerPlex Fusion

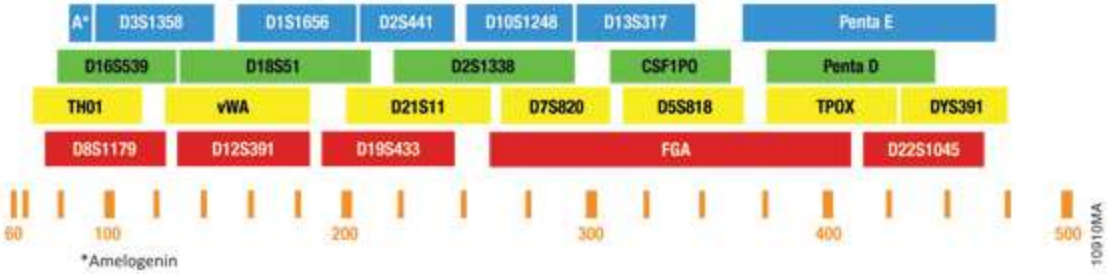
PowerPlex® Fusion System

Launched Friday, September 14, 2012

Designed to meet CODIS and European standards, the PowerPlex® Fusion System enables laboratories to:

- Achieve the most inter database compatibility and highest discrimination of any autosomal STR kit.
- Improve laboratory efficiencies with rapid cycling and direct amplification protocols.
- Obtain a higher success rate with difficult casework samples due to robustness and sensitivity.
- Simplify validation and QC efforts by using one kit for both casework and databasing sections.

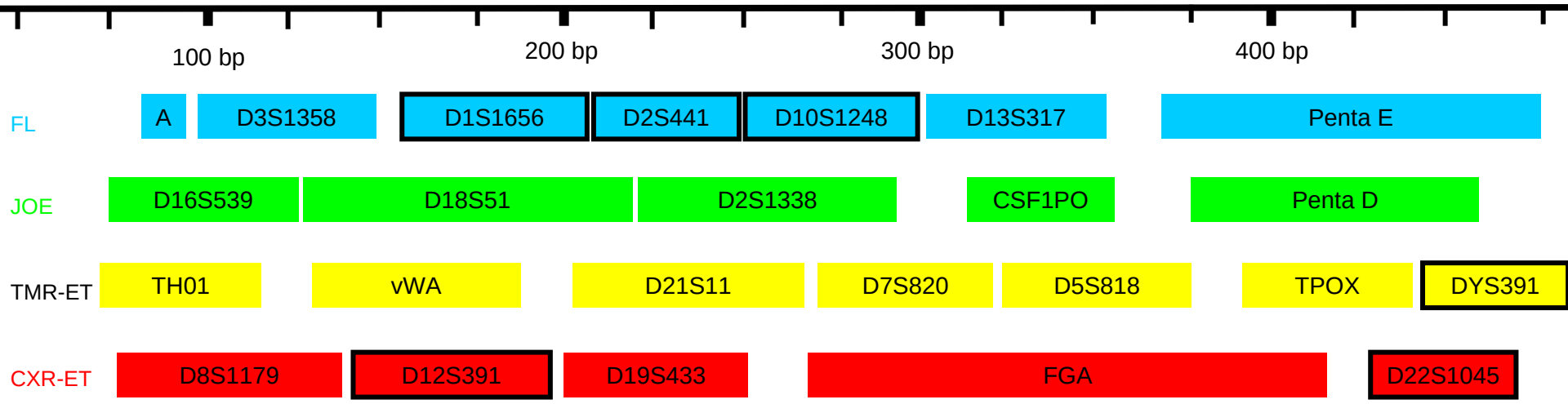
The PowerPlex® Fusion System provides all of the materials needed for co-amplification and five-color fluorescent detection of 24 loci (23 STR loci and Amelogenin), including the CODIS core loci and the European Standard Set (ESS) loci. With 24 loci, the system offers the most STR loci and highest discrimination from a single reaction and delivers more information in demanding forensic, paternity and relationship testing cases. Utilizing proven STR chemistries on existing instrument platforms and software, the PowerPlex® Fusion System requires no software or instrument upgrades.



http://www.promega.com/products/pm/genetic-identity/powerplex-fusion/

Promega PowerPlex **FUSION**

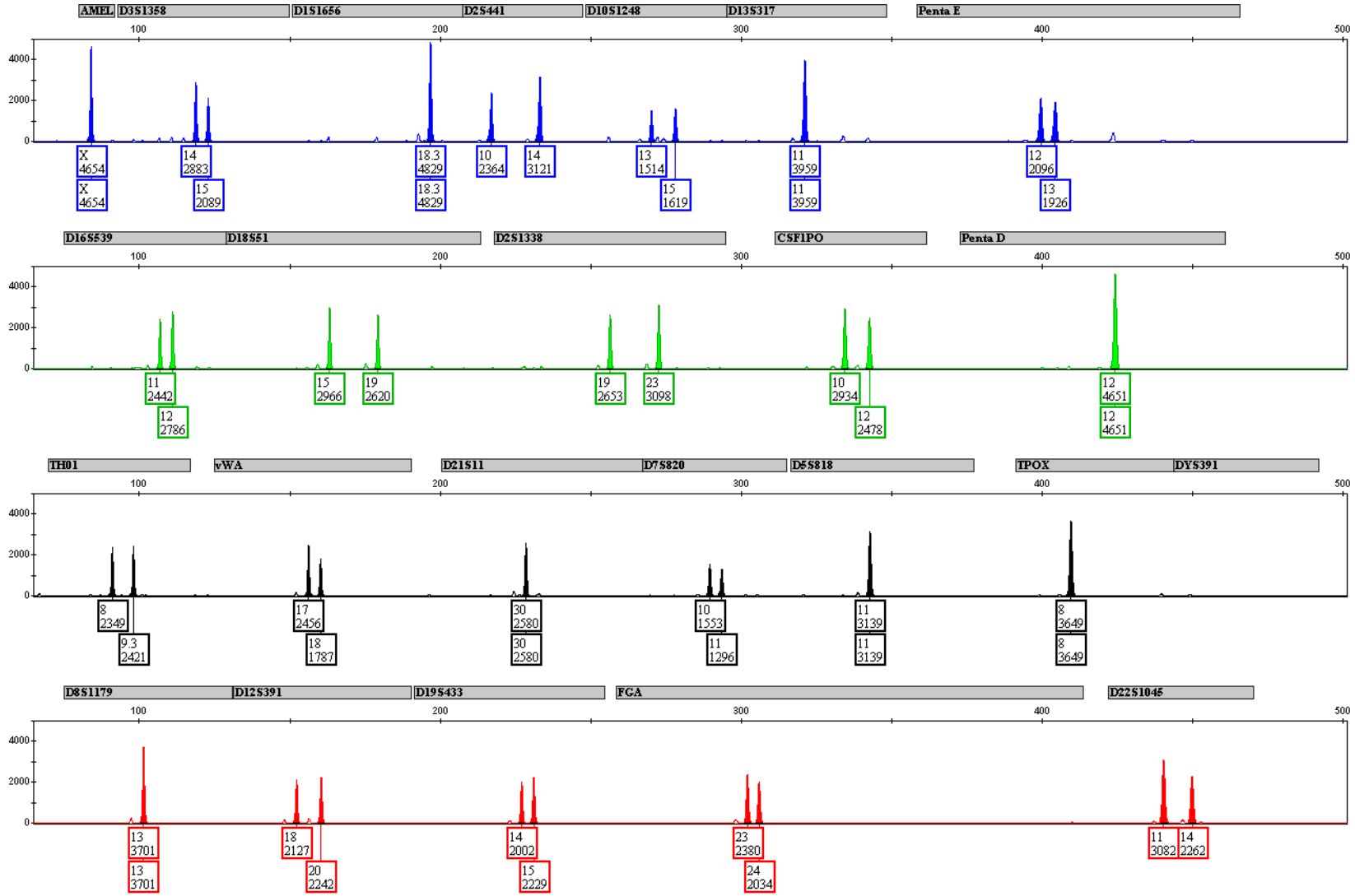
24plex



- 24 STR loci in 5 dyes (3130 and 3500 instrument use)
  - Includes Penta D and E
- Direct amplification and casework capabilities: 85 min amp for both (one kit)
- PowerPlex Fusion gives ~13 orders of magnitude improvement using the NIST 1036 data set

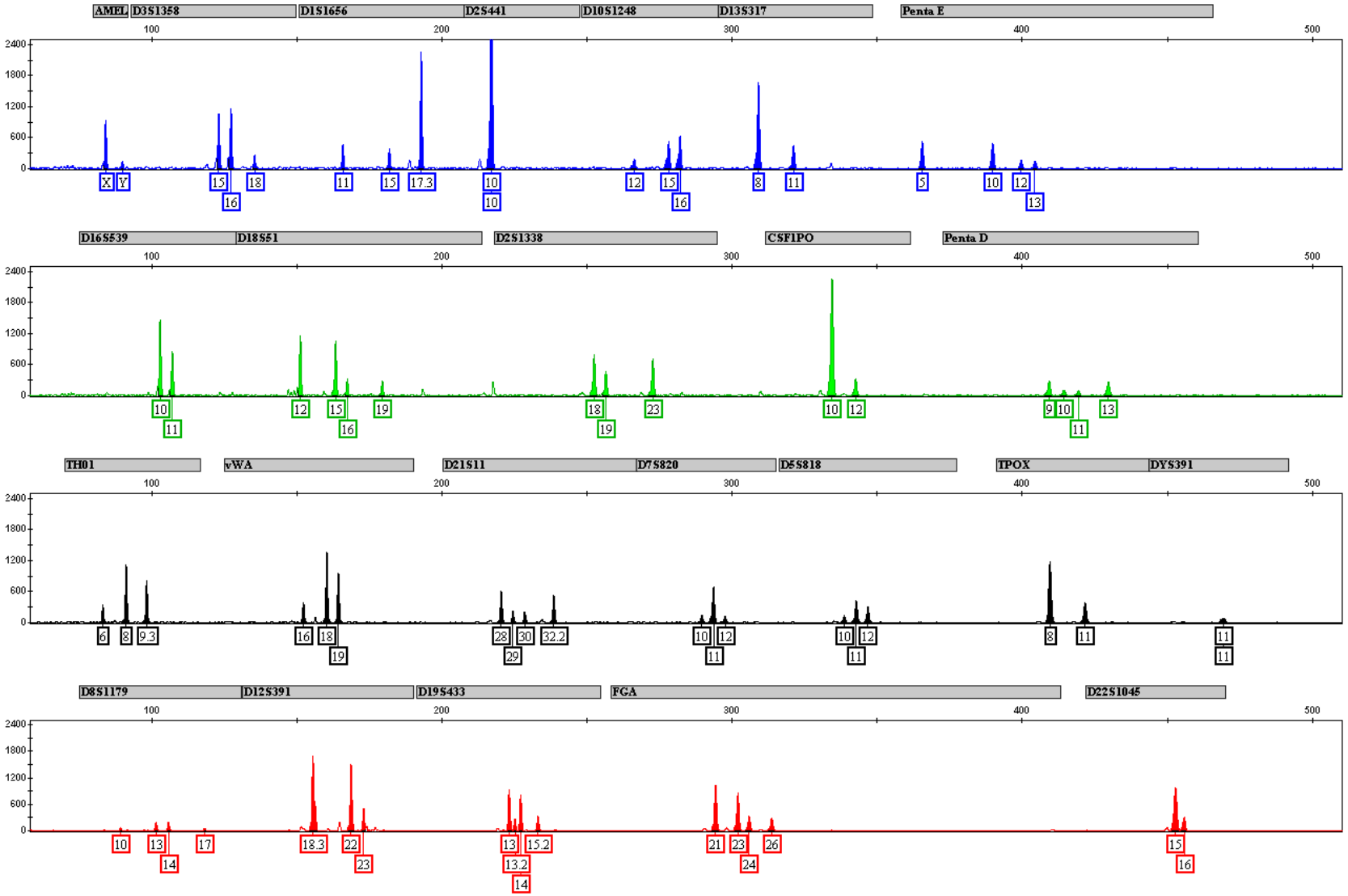
http://www.promega.com/products/pm/genetic-identity/powerplex-fusion/

SRM 2391b&c were **fully concordant** at all loci for PP Fusion kit – **9947A Profile**



1 ng DNA, 30 cycles, 3130xl

SRM 2391c Mixture Component D



1 ng DNA, 30 cycles, 3130xl

NIST U.S. Population Samples

NIST U.S. Samples (>1450)

- **NIST U.S. population samples**
  - 260 African American, 260 Caucasian, 140 Hispanic, 3 Asian
- **U.S. father/son paired samples**
  - ~**100 fathers/100 sons for each group**: 200 African American, 200 Caucasian, 200 Hispanic, 200 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)

Hill, C.R., et al. (2011) Concordance and population studies along with stutter and peak height ratio analysis for the PowerPlex® ESX 17 and ESI 17 Systems. *Forensic Sci. Int. Genet.* 5(4): 269-275.

Publications using NIST Population Samples

Data available at
http://www.cstl.nist.gov/strbase/NISTpop.htm

1. Butler et al. (2003) *J. Forensic Sci.* – Identifiler allele frequencies
2. Butler et al. (2003) *J. Forensic Sci.* – miniSTR assay development
3. Drabek et al. (2004) *J. Forensic Sci.* – miniSTR concordance
4. Schoske et al. (2004) *Forensic Sci. Int.* – Y-STR 20plex & 11plex
5. Vallone et al. (2004) *J. Forensic Sci.* – 50 Y-SNPs
6. Coble & Butler (2005) *J. Forensic Sci.* – NC01 & NC02 assay development
7. Butler et al. (2005) *J. Forensic Sci.* – PowerPlex Y with Y-STR duplications & triplications
8. Vallone et al. (2005) *Forensic Sci. Int.* – 70 autosomal SNPs
9. Butler et al. (2006) *Forensic Sci. Int.* – 27 Y-STR additional loci
10. Hill et al. (2007) *J. Forensic Sci.* – MiniFiler concordance
11. Decker et al. (2008) *FSI Genetics* - Yfiler mutation rates
12. Saunier et al. (2008) *FSI Genetics* – mtDNA control region sequencing (AFDIL)
13. Just et al. (2008) *FSI Genetics* – mtGenome analysis (AFDIL)
14. Hill et al. (2008) *J. Forensic Sci.* – NC01-NC09 miniSTR loci
15. Diegoli et al. (2009) *FSI Genetics* – mtDNA control region sequencing (AFDIL)
16. Hill et al. (2009) *J. Forensic Sci.* – NIST 26plex
17. Lao et al. (2010) *Human Mutation* – 24 ancestry SNPs, Y-SNPs, mtDNA
18. Hill et al. (2011) *FSI Genetics* – ESI 17 & ESX 17 concordance
19. Diegoli et al. (2011) *FSI Genetics Suppl. Ser.* – Argus X-12 X-STR loci
20. Fondevila et al. (2012) *Int. J. Legal Med*. – 68 InDel loci
21. Fondevila et al. (2012) *FSI Genetics* – 34 ancestry SNPs
22. Butler et al. (2012) *Profiles in DNA* – introduces NIST 1036 data set
23. Hill et al. (2013) *FSI Genetics* – 29 autosomal STRs in PowerPlex CS7 and other kits
24. Coble et al. (2013) *FSI Genetics* – 23 Y-STRs in PowerPlex Y23

Testing also completed with 16 X-STR loci and 14 rapidly mutating (RM) Y-STRs

NIST 1036 U.S. Population Samples

- 1032 males + 4 females
  - 361 Caucasians (2 female)
  - 342 African Americans (1 female)
  - 236 Hispanics
  - 97 Asians (1 female)

***Unrelated* samples**
All known or potential related individuals (based on autosomal & lineage marker testing) have been removed from the 1036 data set (e.g., only sons were used from father-son samples)

- Anonymous donors with self-identified ancestry
  - Interstate Blood Bank (Memphis, TN) – obtained in 2002
  - Millennium Biotech, Inc. (Ft. Lauderdale, FL) – obtained in 2001
  - DNA Diagnostics Center (Fairfield, OH) – obtained in 2007

- **Complete profiles with 29 autosomal STRs + PowerPlex Y23**
  - **Examined with multiple kits** and in-house primer sets enabling concordance

- Additional DNA results available on subsets of these samples
  - mtDNA control region/whole genome (AFDIL)
  - >100 SNPs (AIMs), 68 InDel markers, X-STRs (AFDIL)
  - NIST assays: miniSTRs, 26plex, >100 Y-STRs, 50 Y-SNPs

Data available on STRBase: **http://www.cstl.nist.gov/strbase/NISTpop.htm**

Benefits of NIST 1036 Data Set

- **Elimination of potential null alleles due to primer binding site mutations** through extensive concordance testing performed with different PCR primer sets from all available commercial STR kits

- **Ancestry testing performed** on DNA samples with autosomal SNPs, Y-SNPs, and mtDNA sequencing to verify self-declared ancestry categorization

- **Related individuals removed** based on Y-STR and mtDNA results

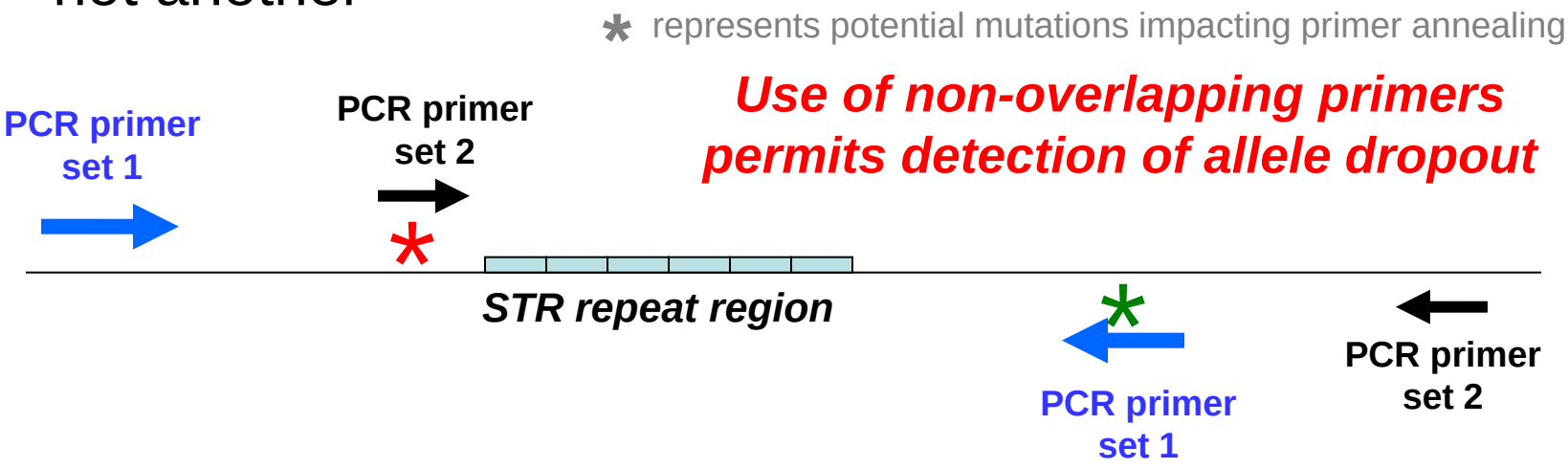
Concordance Testing at NIST

STR Kit Concordance Testing

- Many of these STR kits have different primer sequences for amplifying the same STR locus
- Need to analyze the same DNA samples with different STR typing kits looking for differences
- In some rare cases, allele dropout may occur due to mutations in primer binding regions

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

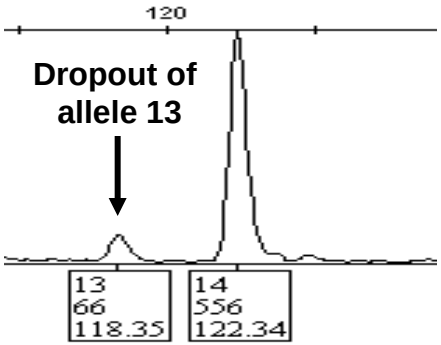


If no primer binding site mutations

If a primer binding site mutation exists

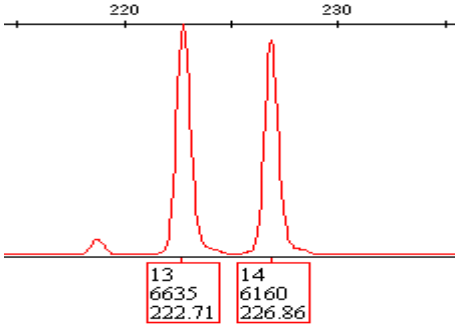
Example Primer Binding Site Mutation that Causes a Null Allele

Identifiler = 14,14



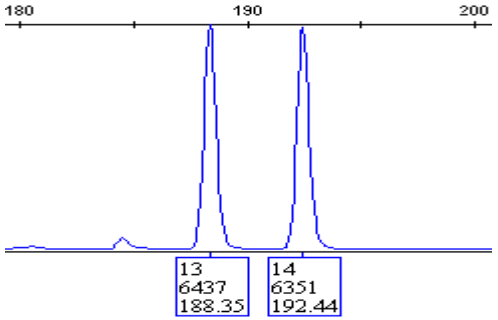
PHR = 11.9%

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

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

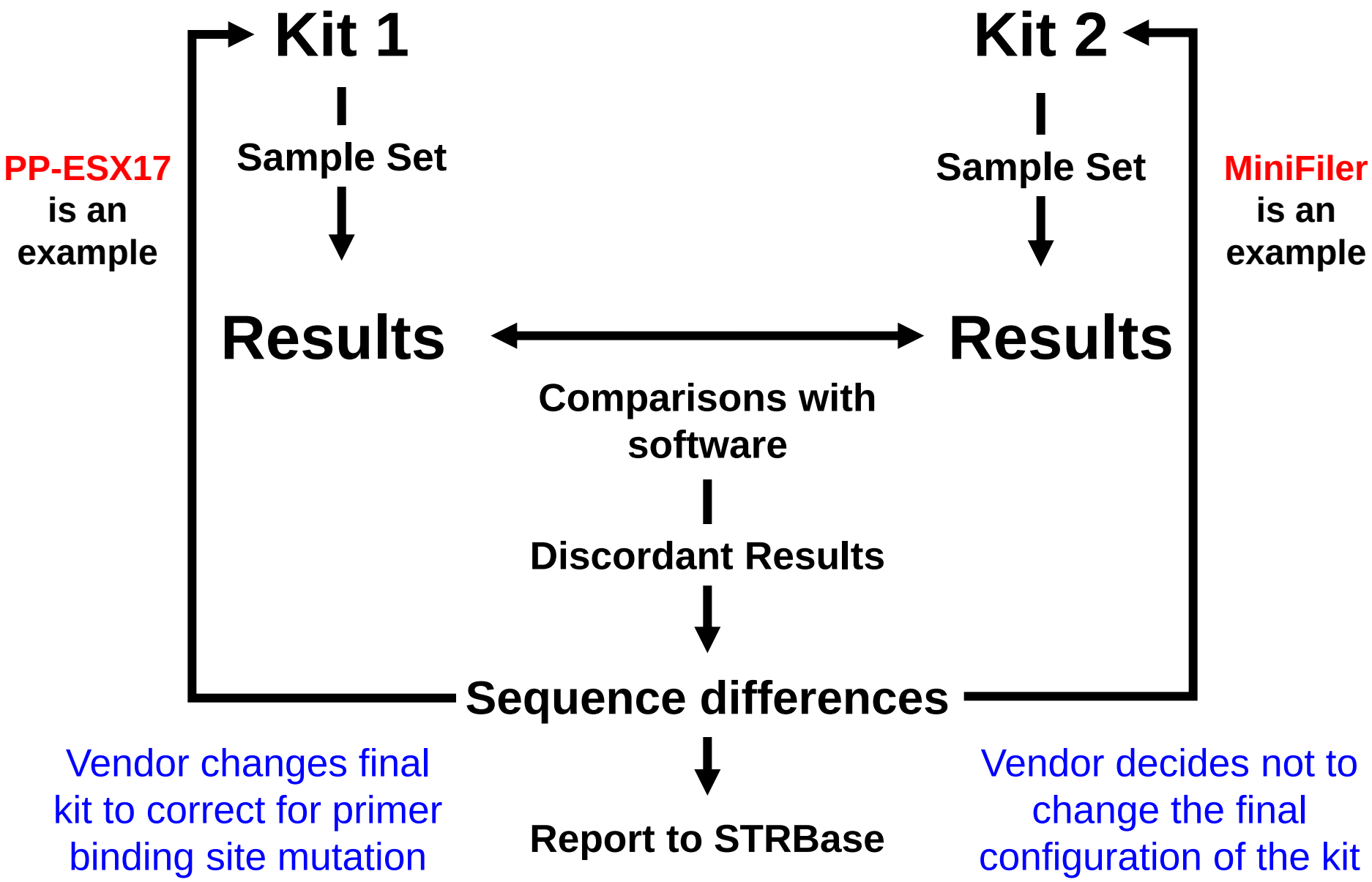
The 4 “S’s” of Concordance

- 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



Completed Concordance Studies

Applied Biosystems AmpFℓSTR Kits

- Identifiler
- **MiniFiler**
- Profiler Plus
- SGM Plus
- NGM
- NGM SElect

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 16/16HS
- **PowerPlex ESX 17** (& Fast)
- **PowerPlex ESI 17** (& Fast)
- PowerPlex ESI 17 Pro
- PowerPlex 18D (rapid and direct kit)
- PowerPlex 21
- PowerPlex Fusion


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
- ESSplex Plus
- ESSplex SE
- ESSplex SE Plus
- Hexaplex ESS
- IDplex
- IDplex Plus

Completed Concordance Studies

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
PP21*/Identifiler	761	16	12176	6	99.951
PP21*/PP16	761	16	12176	3	99.975
PP21/SGM+	761	11	8371	4	99.952
PP21/Pro+	761	10	7610	2	99.974
PP21/IDplex	761	16	12176	20	99.836
PP21/MiniFiler	761	9	6849	14	99.796
PP21/ESX17	761	13	9893	1	99.990
PP21/ESI17	761	13	9893	0	100.000
PP21/NGM	761	13	9893	5	99.949
PP21/NGMs	568	13	7384	1	99.986
PP21/ESS	761	13	9893	18	99.818
PP21/ESSplexSE	568	13	7384	16	99.783
PP21/Hexaplex	568	4	2272	1	99.956
PP21/Sinofiler	568	16	9088	1	99.989
PP21nk/Identifiler	639	16	10224	4	99.961
PP21nk/PP16	639	16	10224	1	99.990
Totals	**114144**	**1245**	**1,104,031**	**1224**	**99.889**

1,104,031 allele comparisons

1,224 total differences

99.89% concordance

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

Final Concordance Results

Marker	Kits with Correct Genotype	# Kits Compared	Correct Type	Kits with Null Allele/Discordant Genotype	Incorrect Type	Total Samples	Sequence Issue
Amel	ID/ESX17/ESI17/PP16/MiniFiler/Pro+/SGM+/ESS/ESSplexSE/IDplex/NGMs/Hexaplex	13	X,Y	NGM*	Y,Y	657	yet to be determined
Amel	ID/ESX17/ESI17/PP16/MiniFiler/Pro+/SGM+/ESS/ESSplexSE/IDplex/NGMs/Hexaplex	13	X,Y	NGM*	Y,Y	657	yet to be determined
Amel	ID/ESX17/ESI17/Pro+/SGM+/ESS/IDplex	8	X,Y	NGM*	Y,Y	780	yet to be determined
Amel	ESS/ESSplexSE/IDplex/Hexaplex call	13	X,Y	ID/MiniFiler/ESI17/ESX17/PP16/NGM/NGMs/SGM+/PP18D	Y,Y	653	yet to be determined
CSF1PO	PP16/MiniFiler/IDplex/PP18D	5	11,11	ID call	11,11.1	656	1 bp ins in ID amplicon outside of PP16 and MiniFiler primers [2]
CSF1PO	ID/MiniFiler/IDplex/PP18D	5	9,12	PP16 call	12,12	662	C/T SNP 16 bp ds from repeat
D10S1248	ESX17/ESI17/NGM/NGMs/ESS/ESSplexSE	7	14,16	Hexaplex	14,14	653	G/T SNP 24 bp ds from repeat
D10S1248	ESX17/ESI17/NGM/NGMs/ESS/ESSplexSE	7	14,16	Hexaplex	14,14	653	G/T SNP 24 bp ds from repeat
D13S317	Pro+/ID/PP16/IDplex	5	9, 11	MiniFiler	11, 11	656	4 bp del in the rev MiniFiler primer binding site [2]
D13S317	Pro+/ID/PP16/IDplex	5	9, 11	MiniFiler	11, 11	656	4 bp del in the rev MiniFiler primer binding site [2]
D13S317	Pro+/ID/PP16/IDplex	5	9, 13	MiniFiler	13, 13	656	4 bp del in the rev MiniFiler primer binding site [2]
D13S317	Pro+/ID/PP16/IDplex	5	9, 12	MiniFiler	12, 12	656	4 bp del in the rev MiniFiler primer binding site [2]
D13S317	Pro+/ID/PP16/IDplex	5	8, 12	MiniFiler	12, 12	656	4 bp del in the rev MiniFiler primer binding site [2]
D13S317	Pro+/ID/PP16/IDplex	5	8, 11	MiniFiler	11, 11	656	4 bp del in the rev MiniFiler primer binding site [2]
D13S317	Pro+/ID/PP16/IDplex	5	10, 13	MiniFiler	13, 13	656	4 bp del in the rev MiniFiler primer binding site [2]
D13S317	Pro+/ID/PP16/IDplex	5	9, 11	MiniFiler	11, 11	656	4 bp del in the rev MiniFiler primer binding site [2]
D13S317	Pro+/ID/PP16/IDplex	5	9, 12	MiniFiler	12, 12	656	4 bp del in the rev MiniFiler primer binding site [2]
D13S317	Pro+/ID/PP16/IDplex	5	9, 14	MiniFiler	14, 14	656	4 bp del in the rev MiniFiler primer binding site [2]
D13S317	Pro+/ID/IDplex	4	9,10	MiniFiler	10,10	481	4 bp del in the rev MiniFiler primer binding site [2]
D13S317	Pro+/ID/IDplex	4	10,11	MiniFiler	11,11	481	4 bp del in the rev MiniFiler primer binding site [2]
D13S317	Pro+/ID/IDplex	4	8,10	MiniFiler	8,8	481	4 bp del in the rev MiniFiler primer binding site [2]
D13S317	Pro+/ID/IDplex	4	10,12	MiniFiler	12,12	481	4 bp del in the rev MiniFiler primer binding site [2]
D16S539	ID/MiniFiler/ESI17/PP16/NGM/NGMs/SGM+/ESS/ESSplexSE/IDplex	11	12,13	ESX17	12,12	660	yet to be determined
D16S539	ID/MiniFiler/ESI17/ESX17/PP16/NGM/NGMs/SGM+/IDplex	11	9,11	ESS/ESSplexSE	11,11	653	C/G SNP 10 bp us from repeat
D16S539	ID/ESX17/ESI17/PP16/NGM/NGMs/SGM+/ESS/ESSplexSE/IDplex	11	9,11	MiniFiler	9,9	656	A/G SNP in MiniFiler primer binding site [2]
D16S539	ID/ESX17/ESI17/PP16/NGM/NGMs/SGM+/ESS/ESSplexSE/IDplex	11	11,12	MiniFiler	12,12	656	A/G SNP in MiniFiler primer binding site [2]
D16S539	ID/ESX17/ESI17/PP16/NGM/NGMs/SGM+/ESS/ESSplexSE/IDplex	11	9,11	MiniFiler	9,9	656	A/G SNP in MiniFiler primer binding site [2]
D16S539	ID/ESX17/ESI17/NGM/SGM+/ESS/IDplex	8	11,14	MiniFiler	14,14	481	A/G SNP in MiniFiler primer binding site [2]
D16S539	ID/ESX17/ESI17/NGM/SGM+/ESS/IDplex	8	9,11	MiniFiler	9,9	481	A/G SNP in MiniFiler primer binding site [2]
D16S539	ID/ESX17/ESI17/NGM/SGM+/ESS/IDplex	8	11,13	MiniFiler	13,13	481	A/G SNP in MiniFiler primer binding site [2]
D16S539	ID/ESX17/ESI17/NGM/SGM+/ESS/IDplex	8	11,12	MiniFiler	12,12	481	A/G SNP in MiniFiler primer binding site [2]
D16S539	ID/ESX17/ESI17/NGM/SGM+/ESS/IDplex	8	11,12	MiniFiler	12,12	481	A/G SNP in MiniFiler primer binding site [2]
D16S539	ID/ESX17/ESI17/NGM/SGM+/ESS/IDplex	8	9,12	MiniFiler	9,9	481	A/G SNP in MiniFiler primer binding site [2]
D18S51	ESX17/ESI17/PP16/MiniFiler/ESS/ESSplexSE/IDplex	12	13,15	ID/NGM/NGMs/Pro+/SGM+	15,15	660	G/A SNP 172 bp ds from repeat
D19S433	ESX17/ESI17	7	13,14	ID/NGM/SGM+/ESS/IDplex	14,14	785	G/A SNP 32 bp ds from repeat
D19S433	ESX17/ESI17	7	13,14.2	ID/NGM/SGM+/ESS/IDplex	14.2,14.2	785	G/A SNP 32 bp ds from repeat
D1S1656	ESI17/NGM/ESSplex	4	14,15.3	ESX17	15.3,15.3	786	C/T SNP 30 bp us from repeat
D22S1045	miniSTR/23plex/NGMs/ESI17/ESS/ESSplexSE/Hexaplex	8	15,17	ESX17*/NGM*	17,17	663	G/T SNP 15 bp us from repeat
D22S1045	miniSTR/23plex/NGMs/ESI17/ESS/ESSplexSE/Hexaplex	8	15,17	ESX17*/NGM*	17,17	663	G/T SNP 15 bp us from repeat
D22S1045	miniSTR/23plex/NGMs/ESI17/ESS/ESSplexSE/Hexaplex	8	15,17	ESX17*/NGM*	17,17	663	G/T SNP 15 bp us from repeat
D22S1045	23plex/ESI17/NGMs/ESS	6	15,16	ESX17*/NGM*	16,16	789	G/T SNP 15 bp us from repeat
D2S1338	PP18D	2	17,23	ID	23,23	50	G/A SNP 174 bp ds from repeat
D2S441	ESX17/ESI17/23plex/ESS/NGMs	6	9.1,11	NGM*	11,11	780	A ins adjacent to the 5' end of repeat
D2S441	ESX17/ESI17/23plex/ESS/NGMs	6	9.1,12	NGM*	12,12	780	A ins adjacent to the 5' end of repeat
D2S441	ESX17/ESI17/23plex/ESS/NGMs	6	9.1,10	NGM*	10,10	780	A ins adjacent to the 5' end of repeat
D2S441	ESX17/ESI17/23plex/ESS/NGMs	6	9.1,11	NGM*	11,11	780	A ins adjacent to the 5' end of repeat
D2S441	ESX17/ESI17/23plex/ESS/NGMs	6	9.1,11	NGM*	11,11	780	A ins adjacent to the 5' end of repeat
D2S441	ESX17/ESI17/23plex/ESS/NGMs	6	9.1,10	NGM*	10,10	780	A ins adjacent to the 5' end of repeat
D2S441	ESX17/ESI17/23plex/ESS/NGMs	6	9.1,11	NGM*	11,11	780	A ins adjacent to the 5' end of repeat
D3S1358	ESX17/ESI17/ESS/IDplex	8	14,17	ID/NGM/Pro+/SGM+	14,14	785	G/C SNP 12 bp us from repeat
D7S820	Pro+/ID	5	8, 9.3	PP16/MiniFiler/IDplex	8, 11	656	5 bp del 114 bp ds from repeat
D8S1179	ID/ESX17/ESI17/NGM/ESS/IDplex	8	14,15	Pro+/SGM+	14,14	773	A/G SNP 15 bp us from repeat
SE33	ESX17/ESI17/NGMs/ESSplexSE	5	26.2,27.2	SE33 Monoplex (SE33)	26.2,26.2	663	C/T SNP 110 bp us from repeat
SE33	ESX17/ESI17/NGMs/ESSplexSE	5	24.2,28.2	SE33 Monoplex (SE33)	28.2,28.2	663	C/T SNP 110 bp us from repeat
SE33	ESX17/ESI17/NGMs/ESSplexSE	5	21.2,26.2	SE33 Monoplex (SE33)	21.2,21.2	663	C/T SNP 110 bp us from repeat
SE33	ESX17/ESI17/NGMs/ESSplexSE	5	24.2,25.2	SE33 Monoplex (SE33)	24.2,24.2	663	C/T SNP 110 bp us from repeat
SE33	SE33/NGMs	5	20,29.2	ESX17/ESI17/ESSplexSE	20,28.3	663	TTG del 28 bp ds from repeat
SE33	SE33/NGMs	3	19,25.2	ESX17	19,19	663	C/T SNP 60 bp ds from repeat
SE33	SE33/NGMs	4	19,25.2	ESI17*/ESSplexSE	19,25.3	663	C/T SNP 60 bp ds from repeat
SE33	SE33/ESX17/NGMs	5	13.2,18	ESI17*/ESSplexSE	13.3,18	50	G/A SNP 68 bp ds from the repeat
SE33	SE33/ESX17/NGMs	5	15.2,19	ESI17*/ESSplexSE	15.3,19	50	G/A SNP 68 bp ds from the repeat
SE33	SE33/ESX17/NGMs	5	15.2,28.2	ESI17*/ESSplexSE	15.3,28.2	50	G/A SNP 68 bp ds from the repeat
SE33	SE33/ESX17/NGMs	5	19,23.2	ESI17*/ESSplexSE	19,23.3	663	G/A SNP 68 bp ds from the repeat
SE33	SE33/ESX17/NGMs	5	16.2,18	ESI17*/ESSplexSE	16.3,18	663	G/A SNP 68 bp ds from the repeat
SE33	SE33/ESX17/NGMs	5	12.2,27.2	ESI17*/ESSplexSE	12.3,27.2	663	G/A SNP 68 bp ds from the repeat
SE33	SE33/ESX17/NGMs	5	16.2,21	ESI17*/ESSplexSE	16.3,21	663	G/A SNP 68 bp ds from the repeat
SE33	SE33/ESX17/NGMs	5	16.2,28.2	ESI17*/ESSplexSE	16.3,28.2	663	G/A SNP 68 bp ds from the repeat
SE33	SE33/ESX17/NGMs	5	13.2,19	ESI17*/ESSplexSE	13.3,19	663	G/A SNP 68 bp ds from the repeat
SE33	SE33/ESX17/NGMs	5	13.2,20	ESI17*/ESSplexSE	13.3,20	789	G/A SNP 68 bp ds from the repeat
SE33	SE33/ESX17/NGMs	5	15.2,27.2	ESI17*/ESSplexSE	15.3,27.2	789	G/A SNP 68 bp ds from the repeat
SE33	ESX17/ESI17/SE33/NGMs	5	13,18	ESSplexSE	18,18	662	AAAA del 85 bp us of the repeat
SE33	ESX17/ESI17/SE33/NGMs	5	15,17	ESSplexSE	15,15	662	AAAA del 85 bp us of the repeat
SE33	ESX17/ESI17/SE33/NGMs	5	16,17	ESSplexSE	17,17	662	AAAA del 85 bp us of the repeat
SE33	ESX17/ESI17/SE33/NGMs	5	15,19	ESSplexSE	19,19	662	AAAA del 85 bp us of the repeat
TH01	ID/MiniFiler/ESI17/ESX17/PP16/NGM/NGMs/SGM+/Hexaplex	12	6,7	ESSplex/ESSplexSE/IDplex	6,6	653	G/A SNP 37 bp ds from the repeat
vWA	ID/MiniFiler/ESI17/ESX17/PP16/NGM/NGMs/SGM+	11	12,16	ESS/ESSplexSE/IDplex	16,16	653	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/PP16/NGM/NGMs/SGM+	11	13,18	ESS/ESSplexSE/IDplex	18,18	653	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/PP16/NGM/NGMs/SGM+	11	14,15	ESS/ESSplexSE/IDplex	15,15	653	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/PP16/NGM/NGMs/SGM+	11	14,15	ESS/ESSplexSE/IDplex	15,15	653	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/PP16/NGM/NGMs/SGM+	11	14,16	ESS/ESSplexSE/IDplex	16,16	653	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/PP16/NGM/NGMs/SGM+	11	14,17	ESS/ESSplexSE/IDplex	17,17	653	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/PP16/NGM/NGMs/SGM+	11	14,17	ESS/ESSplexSE/IDplex	17,17	653	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/PP16/NGM/NGMs/SGM+	11	14,17	ESS/ESSplexSE/IDplex	17,17	653	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/PP16/NGM/NGMs/SGM+	11	14,19	ESS/ESSplexSE/IDplex	19,19	653	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/PP16/NGM/NGMs/SGM+	11	14,19	ESS/ESSplexSE/IDplex	19,19	653	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/PP16/NGM/NGMs/SGM+	11	14,19	ESS/ESSplexSE/IDplex	19,19	653	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/PP16/NGM/NGMs/SGM+	11	15,16	ESS/ESSplexSE/IDplex	16,16	653	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/PP16/NGM/NGMs/SGM+	11	15,17	ESS/ESSplexSE/IDplex	17,17	653	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/NGM/SGM+	8	13,15	ESS/IDplex	15,15	780	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/NGM/SGM+	8	13,15	ESS/IDplex	15,15	780	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/NGM/SGM+	8	13,16	ESS/IDplex	16,16	780	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/NGM/SGM+	8	13,17	ESS/IDplex	17,17	780	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/NGM/SGM+	8	13,19	ESS/IDplex	19,19	780	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/NGM/SGM+	8	14,16	ESS/IDplex	16,16	780	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/NGM/SGM+	8	14,17	ESS/IDplex	17,17	780	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/NGM/SGM+	8	14,17	ESS/IDplex	17,17	780	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/NGM/SGM+	8	14,18	ESS/IDplex	18,18	780	ATCCATCC del 4 bp ds from repeat
vWA	ID/MiniFiler/ESI17/ESX17/NGM/SGM+	8	14,18	ESS/IDplex	18,18	780	ATCCATCC del 4 bp ds from repeat

- All up-to-date results can be found on STRBase:
  - ISFG poster (Vienna, Austria), 8/31-9/2, 2011, "Concordance Testing Comparing STR Multiplex Kits with a Standard Data Set"
  - Promega ISHI (National Harbor, MD), 10/4-10/5, 2011, "Concordance Testing Comparing STR Multiplex Kits with a Standard Data Set"

SRM 2391b/2391c
PCR-Based Profiling Standard

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

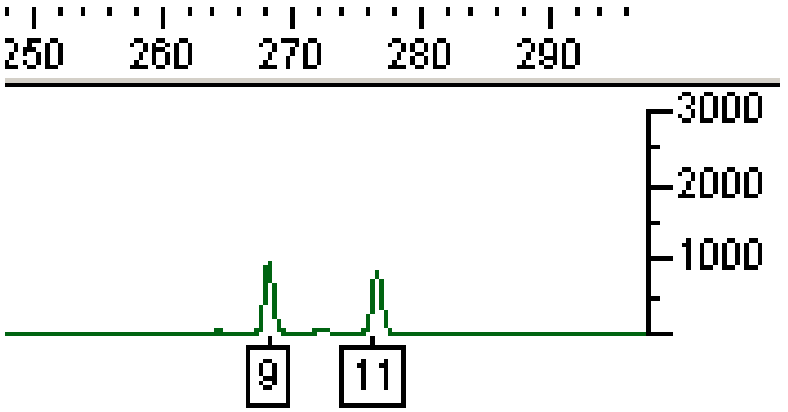
- All new kits tested have been completely concordant with the certified values of all markers for each component for SRM 2391b and 2391c

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

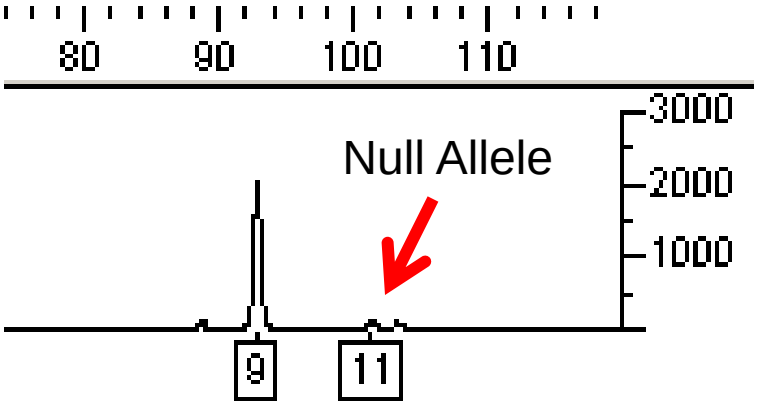
SRM 2391b Genomic 8 with D16S539

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

Identifiler

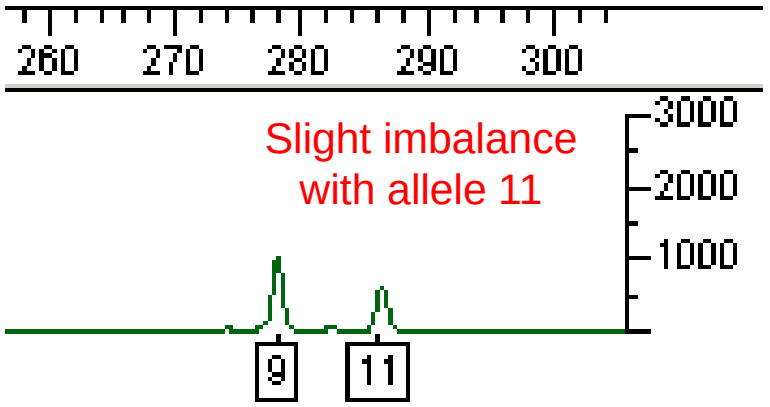


MiniFiler



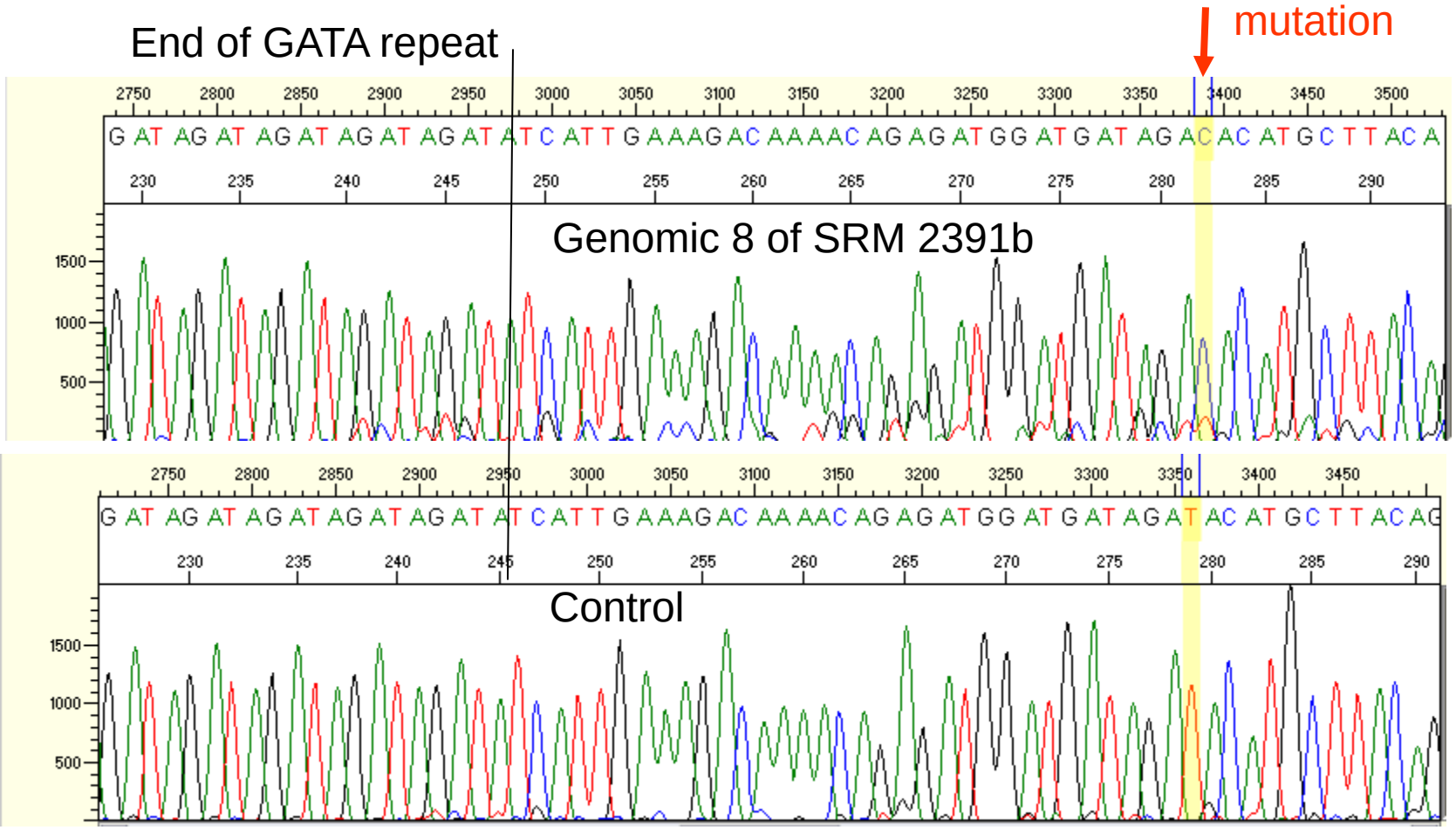
***Due to primer binding site mutation**

PowerPlex 16



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.

Concordance Testing at NIST

- 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&c concordance
  - Several null alleles have been fixed before the final release of new STR multiplex kits

Characterization of STR Loci

Available in Commercial Kits

The 10 STR Loci Beyond the CODIS 13

STR Locus	Location	Repeat Motif	Allele Range*	# Alleles*
D2S1338	2q35	TGCC/TTCC	10 to 31	40
D19S433	19q12	AAGG/TAGG	5.2 to 20	36
Penta D	21q22.3	AAAGA	1.1 to 19	50
Penta E	15q26.2	AAAGA	5 to 32	53
D1S1656	1q42	TAGA	8 to 20.3	25
D12S391	12p13.2	AGAT/AGAC	13 to 27.2	52
D2S441	2p14	TCTA/TCAA	8 to 17	22
D10S1248	10q26.3	GGAA	7 to 19	13
D22S1045	22q12.3	ATT	7 to 20	14
SE33	6q14	AAAG‡	3 to 49	**178**

5 new European loci

**Allele range and number of observed alleles from Appendix 1, J.M. Butler (2011) Advanced Topics in Forensic DNA Typing: Methodology; ‡SE33 alleles have complex repeat structure*

25 Alleles Reported in the Literature for **D1S1656**

15 NIST observed alleles circled in red

Allele (Repeat #)	Promega ESX 17	Promega ESI 17	ABI NGM	Repeat Structure $[TAGA]_4[TGA]_{0-1}[TAGA]_nTAGG[TG]_5$	Reference
8	133 bp	222 bp	171 bp	$[TAGA]_8[TG]_5$	Phillips *et al.* (2010)
9	137 bp	226 bp	175 bp	$[TAGA]_9[TG]_5$	Phillips *et al.* (2010)
10 (a)	141 bp	230 bp	179 bp	$[TAGA]_{10}[TG]_5$	Lareu *et al.* (1998)
10 (b)	141 bp	230 bp	179 bp	$[TAGA]_{10}TAGG[TG]_5$	Phillips *et al.* (2010)
11	145 bp	234 bp	183 bp	$[TAGA]_{11}[TG]_5$	Lareu *et al.* (1998)
12 (a)	149 bp	238 bp	187 bp	$[TAGA]_{12}[TG]_5$	Lareu *et al.* (1998)
12 (b)	149 bp	238 bp	187 bp	$[TAGA]_{11}TAGG[TG]_5$	Lareu *et al.* (1998)
13 (a)	153 bp	242 bp	191 bp	$[TAGA]_{12}TAGG[TG]_5$	Lareu *et al.* (1998)
13 (b)	153 bp	242 bp	191 bp	$[TAGA]_{13}[TG]_5$	Phillips *et al.* (2010)
13.3	156 bp	245 bp	194 bp	$[TAGA]_1TGA[TAGA]_{11}TAGG[TG]_5$	Phillips *et al.* (2010)
14 (a)	157 bp	246 bp	195 bp	$[TAGA]_{13}TAGG[TG]_5$	Lareu *et al.* (1998)
14 (b)	157 bp	246 bp	195 bp	$[TAGA]_{14}[TG]_5$	Phillips *et al.* (2010)
14.3	160 bp	249 bp	198 bp	$[TAGA]_4TGA[TAGA]_9TAGG[TG]_5$	Phillips *et al.* (2010)
15	161 bp	250 bp	199 bp	$[TAGA]_{14}TAGG[TG]_5$	Lareu *et al.* (1998)
15.3	164 bp	253 bp	202 bp	$[TAGA]_4TGA[TAGA]_{10}TAGG[TG]_5$	Lareu *et al.* (1998)
16	165 bp	254 bp	203 bp	$[TAGA]_{15}TAGG[TG]_5$	Lareu *et al.* (1998)
16.3	168 bp	257 bp	206 bp	$[TAGA]_4TGA[TAGA]_{11}TAGG[TG]_5$	Lareu *et al.* (1998)
17	169 bp	258 bp	207 bp	$[TAGA]_{16}TAGG[TG]_5$	Lareu *et al.* (1998)
17.1	170 bp	259 bp	208 bp	Not published	Schröer *et al.* (2000)
17.3	172 bp	261 bp	210 bp	$[TAGA]_4TGA[TAGA]_{12}TAGG[TG]_5$	Lareu *et al.* (1998)
18	173 bp	262 bp	211 bp	$[TAGA]_{17}TAGG[TG]_5$	Phillips *et al.* (2010)
18.3	176 bp	265 bp	214 bp	$[TAGA]_4TGA[TAGA]_{13}TAGG[TG]_5$	Lareu *et al.* (1998)
19	177 bp	266 bp	215 bp	Not published	Asamura *et al.* (2008)
19.3	180 bp	269 bp	218 bp	$[TAGA]_4TGA[TAGA]_{14}TAGG[TG]_5$	Lareu *et al.* (1998)
20.3	184 bp	273 bp	222 bp	Not published	Gamero *et al.* (2000)

from Appendix 1, J.M. Butler (2011) Advanced Topics in Forensic DNA Typing: Methodology

NIST **U.S. Population Allele Frequencies**

D1S1656 (15 different alleles)

15 different alleles

Allele	African American (n=342)	Asian (n=97)	Caucasian (n=361)	Hispanic (n=236)
10	0.0146	0.0000	0.0028	0.0064
11	0.0453	0.0309	0.0776	0.0275
12	0.0643	0.0464	0.1163	0.0890
13	0.1009	0.1340	0.0665	0.1144
14	**0.2573**	0.0619	0.0789	0.1165
14.3	0.0073	0.0000	0.0028	0.0042
15	0.1579	**0.2784**	**0.1496**	0.1377
15.3	0.0292	0.0000	0.0582	0.0508
16	0.1096	0.2010	0.1357	**0.1758**
16.3	0.1023	0.0155	0.0609	0.0508
17	0.0278	0.0722	0.0471	0.0424
17.3	0.0497	0.0876	0.1330	0.1483
18	0.0029	0.0155	0.0055	0.0064
18.3	0.0234	0.0515	0.0499	0.0254
19.3	0.0073	0.0052	0.0152	0.0042

N=1036

(only unrelated samples used; fathers removed from this sample set)

D1S1656 Characteristics

- **15 alleles** observed
- **93 genotypes** observed
- **>89% heterozygotes** (heterozygosity = 0.8890)
- **0.0224 Probability of Identity** (P_I)

$$P_I = \sum (genotype\ \ frequencies)^2$$

These values have been calculated for all 29 STR loci across the U.S. population samples examined

Loci sorted on Probability of Identity (P_I) values

Locus	Alleles Observed	Genotypes Observed	Het (obs)	P_I Value n=1036
SE33	52	304	0.9353	0.0066
Penta E	23	138	0.8996	0.0147
D2S1338	13	68	0.8793	0.0220
D1S1656	15	93	0.8890	0.0224
D18S51	22	93	0.8687	0.0258
D12S391	24	113	0.8813	0.0271
FGA	27	96	0.8745	0.0308
D6S1043	27	109	0.8494	0.0321
Penta D	16	74	0.8552	0.0382
D21S11	27	86	0.8330	0.0403
D8S1179	11	46	0.7992	0.0558
D19S433	16	78	0.8118	0.0559
vWA	11	39	0.8060	0.0611
F13A01	16	56	0.7809	0.0678
D7S820	11	32	0.7944	0.0726
D16S539	9	28	0.7761	0.0749
D13S317	8	29	0.7674	0.0765
TH01	8	24	0.7471	0.0766
Penta C	12	49	0.7732	0.0769
D2S441	15	43	0.7828	0.0841
D10S1248	12	39	0.7819	0.0845
D3S1358	11	30	0.7519	0.0915
D22S1045	11	44	0.7606	0.0921
F13B	7	20	0.6911	0.0973
CSF1PO	9	31	0.7558	0.1054
D5S818	9	34	0.7297	0.1104
FESFPS	12	36	0.7230	0.1128
LPL	9	27	0.7027	0.1336
TPOX	9	28	0.6902	0.1358

Better for mixtures (more alleles seen)

29 STR Loci present in STR kits

rank ordered by their variability

N=1036

(only unrelated samples used)

There are several loci more polymorphic than the **current CODIS 13 STRs**

361 Caucasians
342 African Americans
236 Hispanics
97 Asians

Better for kinship (low mutation rate)

Probability of Identity Combinations
(assuming unrelated individuals)

STR Kit or Core Set of Loci	Total N=1036	Caucasians (n=361)	African Am. (n=342)	Hispanics (n=236)	Asians (n=97)
CODIS 13	5.02E-16	2.97E-15	1.14E-15	1.36E-15	1.71E-14
Identifiler	6.18E-19	6.87E-18	1.04E-18	2.73E-18	5.31E-17
PowerPlex 16	2.82E-19	4.24E-18	6.09E-19	1.26E-18	2.55E-17
PowerPlex 18D	3.47E-22	9.82E-21	5.60E-22	2.54E-21	7.92E-20
ESS 12	3.04E-16	9.66E-16	9.25E-16	2.60E-15	3.42E-14
ESI 16 / ESX 16 / NGM	2.80E-20	2.20E-19	6.23E-20	4.03E-19	9.83E-18
ESI 17 / ESX 17 / NGM SElect	1.85E-22	1.74E-21	6.71E-22	3.97E-21	1.87E-19
CODIS 20	9.35E-24	7.32E-23	6.12E-23	8.43E-23	4.22E-21
GlobalFiler	7.73E-28	1.30E-26	3.20E-27	2.27E-26	1.81E-24
PowerPlex Fusion	6.58E-29	2.35E-27	1.59E-28	2.12E-27	1.42E-25
All 29 autosomal STRs	2.24E-37	7.36E-35	3.16E-37	2.93E-35	4.02E-32
29 autoSTRs + DYS391	1.07E-37	3.26E-35	1.77E-37	1.29E-35	2.81E-32

~8-13 orders of magnitude improvement for total P_I (n=1036)

NIST U.S. Population Data

- The data from our 1036 U.S. population samples is currently available on STRBase:

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

- A summary of the NIST 1036 data set has been published in Profiles in DNA for autosomal and YSTR loci

- Population Data announcements have been published in FSI: Genetics for
  - 29 autosomal STR loci (*Hill et al*)
  - 23 Y-STR loci (*Coble et al*)


Forensic Science International: Genetics 7 (2013) e82–e83

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

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

Letter to the Editor

U.S. population data for 29 autosomal STR loci

run and population statistics were confirmed using the Power-Marker v3.25 statistics program [10].

Purpose of Sequencing SRM 2391c

- To further characterize Components A-C, determine interesting genomic characteristics within STR fragments (SNPs, insertions/deletions, etc.)
- Presented as a poster (ISFG 2013)

http://www.cstl.nist.gov/biotech/strbase/pub_pres/Hill-ISFG2013-SRM2391c.pdf

- **To support Next Generation Sequencing of Components A-C**

SRM 2391c: PCR-Based DNA Profiling Standard

- Includes 6 components:

Table 1. Description of Components in SRM 2391c

Component		Description	Amount	Concentration[a]
A	Single-source	Anonymous single-source female genomic DNA in TE^{-4} buffer	50 µL	1.1 – 2.1 ng/µL
B		Anonymous single-source male genomic DNA in TE^{-4} buffer	50 µL	1.1 – 2.1 ng/µL
C		Anonymous single-source male genomic DNA in TE^{-4} buffer	50 µL	1.1 – 2.1 ng/µL
D	Mixture	Mixed-source (Components A and C) genomic DNA in TE^{-4} buffer	50 µL	1.1 – 2.1 ng/µL
E	Stain	Anonymous single-source female cells spotted on 903 paper	Two 6 mm punches	7.5×10^4 cells per punch
F	Stain	Anonymous single-source male cells spotted on FTA paper	Two 6 mm punches	7.5×10^4 cells per punch

[a]DNA concentrations and cell counts are nominal values and are not intended for use as quantitative standards.

https://www-s.nist.gov/srmors/view_cert.cfm?srm=2391C

Sequencing SRM 2391c

- Components A, B, C (genomic, single source DNA samples – one female, two males)

- All certified values were assigned after concordance checks were performed (different primer set results were compared)

- Sequencing has been performed on loci where limited primer sets were available:
  - D1S1656, D12S391, Penta D, Penta E, SE33, D8S1115, DYS448, DYS456, DYS458, DYS635, DYGATAH4

NIST Certified Values

- A NIST certified value is a value for which NIST has the highest confidence in its accuracy in that all known or suspected sources of bias have been investigated or taken into account.
- There are 41 STR markers plus Amelogenin that have certified genotypes assigned by electrophoretic match to previously sequenced alleles (30) or by direct sequencing (11).
- The remaining 30 markers were Sanger sequenced for Components A-C to further characterize the repeat structure and flanking sequence.

Certified Genotypes

Concordance with STR Kits

Autosomal STR Loci	Y-STR Loci
D2S1338	DYS19
D2S441	DYS385a
D3S1358	DYS385b
D5S818	DYS389I
D7S820	DYS389II
D8S1179	DYS390
D10S1248	DYS391
D13S317	DYS392
D16S539	DYS393
D18S51	DYS437
D19S433	DYS438
D21S11	DYS439
D22S1045	*Amelogenin
CSF1PO	
FGA	
TH01	
TPOX	
vWA	

DNA Sequencing of Alleles

Autosomal STR Loci	Y-STR Loci
D1S1656	DYS448
D8S1115	DYS456
D12S391	DYS458
Penta D	DYS635
Penta E	DY-GATA-H4
SE33	

41 STR Markers + Amelogenin are certified
26% have been Sanger Sequenced

>2 STR Kits were tested for concordance

Methods for Sanger Sequencing

- NIST DNA sequencing procedures and all sequencing primers were published in 2011 (see S1)
- Note: alternative primers were designed for D19S433

Forensic Science International: Genetics 5 (2011) 329–332

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Short communication

STR sequence analysis for characterizing normal, variant, and null alleles

Margaret C. Kline *, Carolyn R. Hill, Amy E. Decker [1], John M. Butler

National Institute of Standards and Technology, 100 Bureau Drive, M/S 8312, Gaithersburg, MD 20899, USA

Sequencing Flow Chart

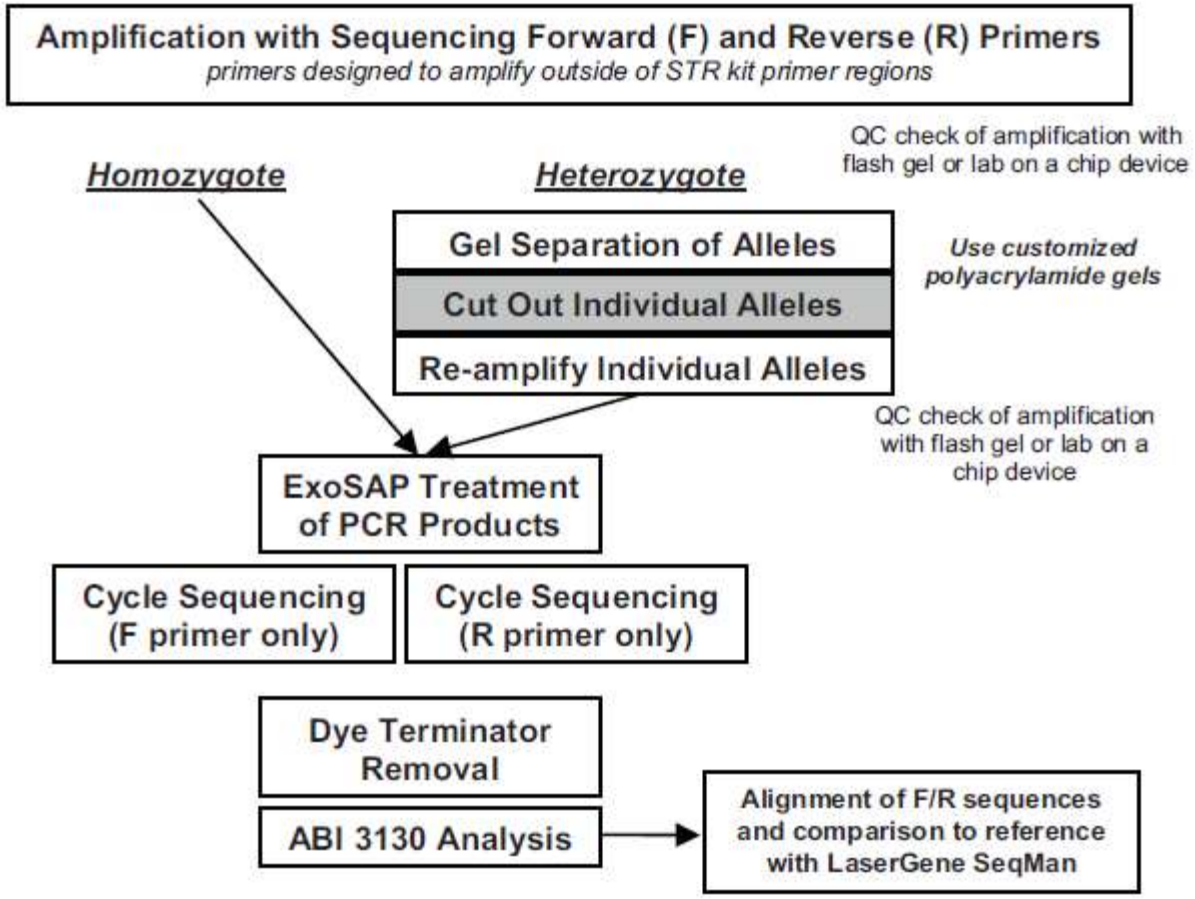
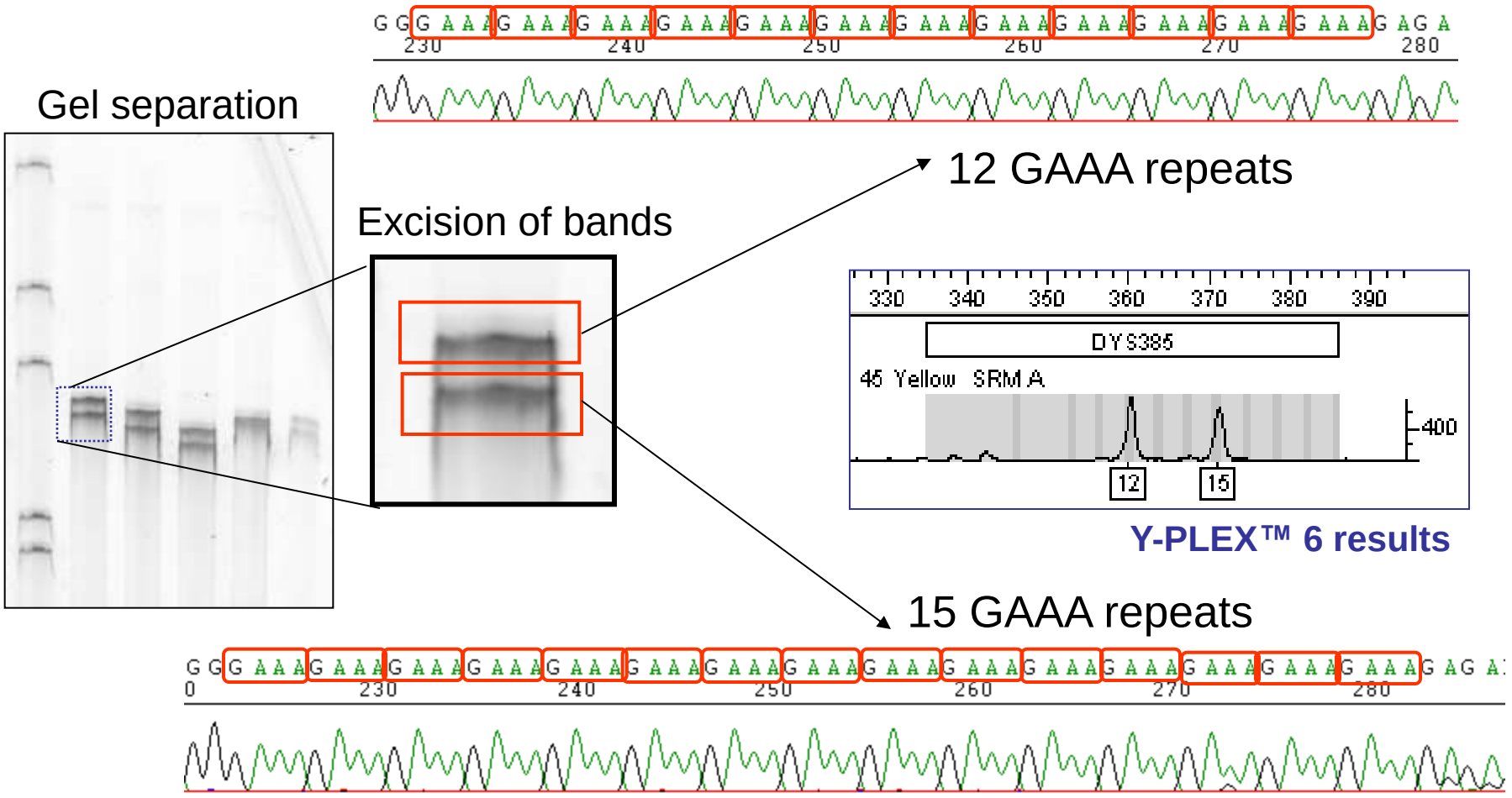


Fig. 1. Summary of the steps used in sequencing STR alleles.

Sequencing Individual Heterozygous (DYS385) Alleles



Kline, M.C., Hill, C.R., Decker, A.E., Butler, J.M. (2011) STR sequence analysis for characterizing normal, variant, and null alleles. Forensic Sci. Int. Genet. 5(4): 329-332

GenBank Reference Sequences

- The GenBank Accession numbers and reference alleles were obtained were based on the May 2004 assembly of the human genome, build 35.
- Sequences were aligned de novo using LaserGene SeqMan software and compared to SeqBuilder maps based on the listed GenBank reference sequences.

Marker	GenBank Accession Number	Marker	GenBank Accession Number	Marker	GenBank Accession Number	Marker	GenBank Accession Number
D1S1656	G07820	D13S317	AL353628.2	SE33	V00481	DYS393	AC006152
D2S1338	AC010136	D16S539	AC024591.3	TH01	D00269	DYS437	AC002992
D2S441	AC079112	D18S51	AP001534	TPOX	M68651	DYS438	AC002992
D3S1358	AC099539	D19S433	AC008507.6	vWA	M25858	DYS439	AC002992
D5S818	AC008512	D21S11	AP000433	DYS19	AC017019	DYS448	AC025227
D7S820	AC004848	D22S1045	AL022314	DYS385	AC022486	DYS456	AC010106.2
D8S1179	AF216671	CSF1PO	X14720	DYS389	AF140635	DYS458	AC010902
D8S1115	AC090739	FGA	M64982	DYS390	AC011289	DYS635	AC004772
D10S1248	AL391869	Penta D	AP001752	DYS391	AC011302	Y GATA H4	AC011751
D12S391	G08921	Penta E	AC027004	DYS392	AC06152		

Sequencing Results

- All sequencing results of Components A-C for 41 STR markers, including repeat structures of individual alleles, can be found on the following poster:

http://www.cstl.nist.gov/biotech/strbase/pub_pres/Hill-ISFG2013-SRM2391c.pdf

Marker	Component	Allele	Allele Repeat Structure
D8S1179	C	17	$[TCTA]_2$ TCTG $[TCTA]_{14}$
D12S391	A	22	$[AGAT]_{13}$ $[AGAC]_8$ AGAT
D12S391	C	19	$[AGAT]_{13}$ $[AGAC]_5$ AGAT
D12S391	C	23	$[AGAT]_{12}$ $[AGAC]_{10}$ AGAT
D21S11	B	32	$[TCTA]_4$ $[TCTG]_6$ {$[TCTA]_3$ TA $[TCTA]_3$ TCA $[TCTA]_2$ TCCATA} $[TCTA]_{14}$
SE33	C	31.2	$[AAAG]_2$ AG $[AAAG]_3$ AG $[AAAG]_9$ AAAAAG $[AAAG]_{21}$ G AAGG$[AAAG]_2$ AG
DYS389II	B	31	$[TCTG]_6$ $[TCTA]_{12}$ $[TCTG]_3$ $[TCTA]_{10}$
DYS458	B	17.2	$[GAAA]_{15}$ AA $[GAAA]_2$
DYS635	B	20	$[TCTA]_4$ $[TGTA]_2$ $[TCTA]_2$ $[TGTA]_2$ $[TCTA]_{10}$
DYS635	C	21	$[TCTA]_4$ $[TGTA]_2$ $[TCTA]_2$ $[TGTA]_2$ $[TCTA]_{11}$

Novel repeat motifs that were not listed in Butler J.M. (2012) or STRBase fact sheets

SNPs Found in Repeat Flanking Regions

- Multiple SNPs were found in the DNA sequence in the repeat flanking regions. Primers that bind on SNPs can result in null alleles when STR typing.
- Note that the variants characterized in this work are constrained by the size of the original PCR amplicon generated (Kline et al. 2011).

Marker	Component	Allele	Flanking Region Variants
D5S818	A	12	T→C 13 bp us of the repeat
D5S818	B	13	T→C 13 bp us of the repeat
D5S818	B	13	G→T 4 bp ds of the repeat
D5S818	C	10	T→C 13 bp us of the repeat
D5S818	C	11	T→C 13 bp us of the repeat
D7S820	C	10	T→G 65 bp ds of the repeat
D13S317	C	11	A→C 115 bp ds of the repeat
D16S539	A	10	A→C 16 bp ds of the repeat
D16S539	A	10	C→A 95 bp us of the repeat
D16S539	A	11	C→A 95 bp us of the repeat
D16S539	B	10	C→A 95 bp us of the repeat
D16S539	C	10	C→A 95 bp us of the repeat
Penta E	A	10	G→A 123 bp us of the repeat
Penta E	A	10	A→G 268 bp us of the repeat
Penta E	A	10	A→C 280 bp us of the repeat
Penta E	B	7	G→A 123 bp us of the repeat
Penta E	B	7	A→G 268 bp us of the repeat
Penta E	B	7	A→C 280 bp us of the repeat
Penta E	B	15	G→A 123 bp us of the repeat
Penta E	B	15	A→G 268 bp us of the repeat
Penta E	B	15	A→C 280 bp us of the repeat
TPOX	A	8	T→G 148 bp ds of the repeat
TPOX	B	8	T→G 148 bp ds of the repeat

Abbreviations: bp = base pairs, us = upstream, ds = downstream

Other Candidates for Sequencing

- Additional non-core autosomal STR markers
  - D6S1043 (Sinofiler, PowerPlex 21)
  - 22 miniSTR loci (not including D2S441, D10S1248, D22S1045, D8S1115)
  - Penta C
  - FFFL loci (F13A01, F13B, FESFPS, LPL)
- Y-STR markers to sequence
  - DYS460, DYS481, DYS533, DYS549, DYS643 (PowerPlex Y23, Yfiler Plus)
- Rapidly mutating (RM) Y-STRs
  - 13 total (**DYF387S1a/b**, DYF399S1, DYF403S1a/b, DYF404S1, **DYS449**, **DYS518**, DYS526a/b, DYS547, **DYS570**, **DYS576**, DYS612, DYS626, **DYS627**)

Future Directions

- Sequencing of Components A-C will be completed for all remaining autosomal and Y-STR loci, including non-core loci to raise all reference and informational genotypes to a certified level.
- Sequencing will also be completed for all autosomal and Y-STR markers for Components E and F (Component D is a mixture of Components A and C).
- Once sequencing is complete, the SRM 2391c Certificate of Analysis will be updated with this new information.
- This work supports the high throughput next generation sequencing technologies at NIST for forensic typing applications.
- SRM 2391c has replaced SRM 2395 for Y-STR typing.

Summary

- Additional STR loci are important as DNA databases grow larger each year: the power of discrimination increases as new loci are added
  - Adding seven new loci (CODIS 13 vs CODIS 20) adds approximately 8 orders of magnitude improvement

- Commercial companies are continuing to release larger STR multiplex kits to meet the needs of the forensic community

- NIST has a set of 1036 unrelated U.S. population samples that have been used to fully characterize 29 autosomal STR loci available in commercial STR multiplex kits

Acknowledgments

$ National Institute of Justice and NIST OLES
Promega, Life Technologies, Qiagen for kits

John Butler (NIST Office of Special Programs)
Mike Coble (NIST Applied Genetics Group)
Pete Vallone (NIST Applied Genetics Group)

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

Final version of this presentation will be available at:
http://www.cstl.nist.gov/strbase/NISTpub.htm