NIST PCR-Based DNA Profiling Standard (SRM 2391d)

Peter M. Vallone, Ph.D., Leader, Applied Genetics Group, NIST

Becky Steffen, Erica Romsos, Lisa Borsuk, Kevin Kiesler, Sarah, Riman, Katherine Gettings

12th Annual Meeting

Burlington, Vermont

August 2, 2019

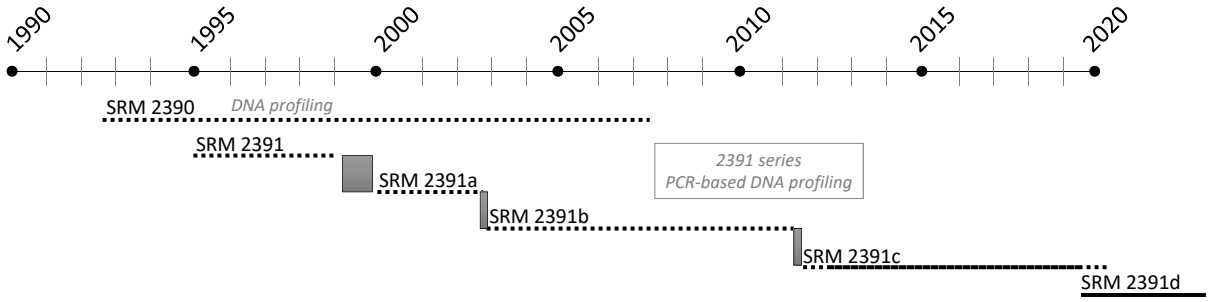
NIST APPLIED GENETICS

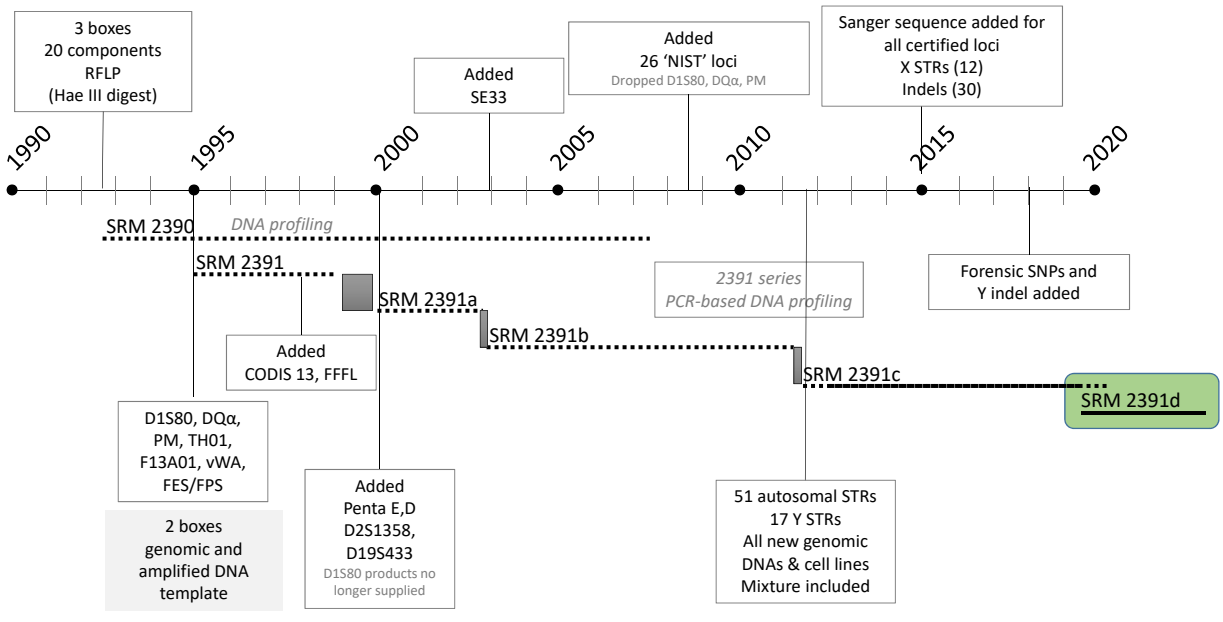
Today's Talk

- Overview of SRM2391d
- Materials
- Characterization methods
- Markers

On sale as of July 9, 2019

Brief history of SRM 2391 series…

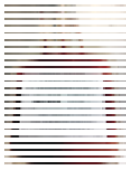




SRM 2391d: a group effort

- Led by Becky Steffen
  - Erica Romsos
  - Katherine Gettings
  - Kevin Kiesler
  - Lisa Borsuk
  - Sarah Riman

Materials (Five Components)



- Components A-D are genomic DNA extracted from purchased blood
  - Not from cell lines (challenges in getting use/permission from Coriell/NIGMS)
  - May be more commutable (similar to casework)
  - Different samples from 2391c
- Component E consists of cells spotted onto FTA paper
  - Two 6 mm punches; approximately 75,000 cells per punch
  - Toward the end of SRM 2391c profile degradation was observed for cells stored on 903 paper (cells on 903 paper not included in SRM 2391d)
  - Same cell line as used in 2391c (CRL-1486)

2300 units were vialed for SRM 2391d

Materials

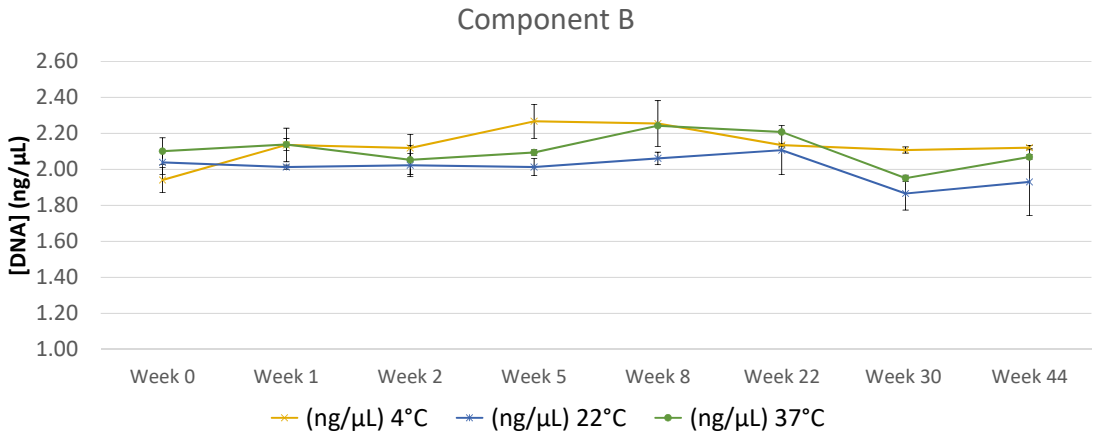
Table 1. Description of Components in SRM 2391d

Assigned using digital PCR

Component	Description	Volume	Concentration[a]
A	Anonymous single-source female genomic DNA in TE^{-4} buffer	55 µL	1.6 ± 0.5 ng/µL
B	Anonymous single-source male genomic DNA in TE^{-4} buffer	55 µL	1.7 ± 0.5 ng/µL
C	Anonymous single-source male genomic DNA in TE^{-4} buffer	55 µL	1.6 ± 0.2 ng/µL
D	Mixed-source, 3:1 (3 parts Component A and 1 part Component C) genomic DNA in TE^{-4} buffer	55 µL	1.5 ± 0.4 ng/µL
E	Anonymous single-source female cells spotted on FTA paper[b]	Two 6 mm punches	7.5×10^4 cells per punch

[a] DNA concentrations and cell counts are provided as Information Values.
[b] FTA paper cards contain chemicals that lyse cells, denature proteins and protect nucleic acids from nucleases, oxidation and UV damage. FTA cards rapidly inactivate organisms, including blood-borne pathogens, and prevent the growth of bacteria and other microorganisms.

Stability Testing (measuring quantity - dPCR)



Quantifiler Trio Degradation Index

Component	Week 0	Week 51		
		4°C	22°C	37°C
A	0.9	0.9	0.6	0.6
B	0.7	0.7	0.5	0.6
C	0.8	0.8	0.5	0.6
D	0.8	0.9	0.6	0.6
E	0.7	0.6	0.6	0.5

DI < 1 'Typically indicates that DNA is not degraded or inhibited'

Will continue to monitor degradation index over the lifetime of SRM 2391d

IPCCT flag triggered?	Degradation Index	Quality Index interpretation†
No	<1	Typically indicates that DNA is not degraded or inhibited.
	1 to 10	Typically indicates that DNA is slightly to moderately degraded. PCR inhibition is also possible, however not enough to significantly suppress IPC amplification.
	>10 or blank (no value)	Typically indicates that DNA is significantly degraded. PCR inhibition is also possible, however not enough to significantly suppress IPC amplification. Highly degraded samples that cannot be recovered by STR can be analyzed with HID-Ion AmpliSeq™ Panels and the Ion Personal Genome™ (PGM™) System (see "Degraded sample studies: GlobalFiler™ STR kit and HID-Ion AmpliSeq™ Identity and Ancestry Panel" on page 103).
Yes	<1	Although theroretically possible, this result is unlikely because PCR inhibitors in sufficient concentration to trigger the IPCCT flag typically would affect the large autosomal target as well.
	>1 or blank (no value)	Typically indicates that the DNA is affected by degradation and/or PCR inhibition.

† These are general guidelines that may not apply to all samples depending on the inhibitors present, the varying quantity of contributor DNA in mixed samples and the STR kit used. (STR kits are For Forensic or Paternity Use Only.)

Characterization of the material

Digital PCR for quantification

QX200

Capillary Electrophoresis

3500xL

Next Generation Sequencing (NGS)

MiSeq FGx

Ion S5 XL

CE-based assays used in characterization

Table 2. Capillary Electrophoresis (CE) typing kits used for SRM 2391d testing

Thermo Fisher	Promega	Qiagen	InnoGenomics
MiniFiler	PowerPlex S5	Investigator ESSplex SE Plus	InnoTyper 21
Identifiler	PowerPlex CS7	Investigator HDplex	
Identifiler Plus	PowerPlex 16	Investigator 24plex QS	
Identifiler Direct	PowerPlex 16 HS	Investigator 24plex GO!	
NGM	PowerPlex 18D	Investigator Argus X-12	
NGM SElect	PowerPlex 21	Investigator DIPplex	
NGM Detect	PowerPlex ESX 17		
VeriFiler Express	PowerPlex ESX 17 Fast		
VeriFiler Plus	PowerPlex ESI 17 Pro		
GlobalFiler	PowerPlex ESI 17 Fast		
GlobalFiler Express	PowerPlex Fusion		
Yfiler	PowerPlex Fusion 6C		
Yfiler Plus	PowerPlex VersaPlex 27PY		
	PowerPlex Y23		

In house CE-typing
D1S1677
D2S1776
D3S4529
D4S2408
D5S2800
D9S1122
D12ATA63
D14S14343
D17S1301
D20S482
DYS461
DYS505
DYS522
DYS612

NGS-based assays used in characterization

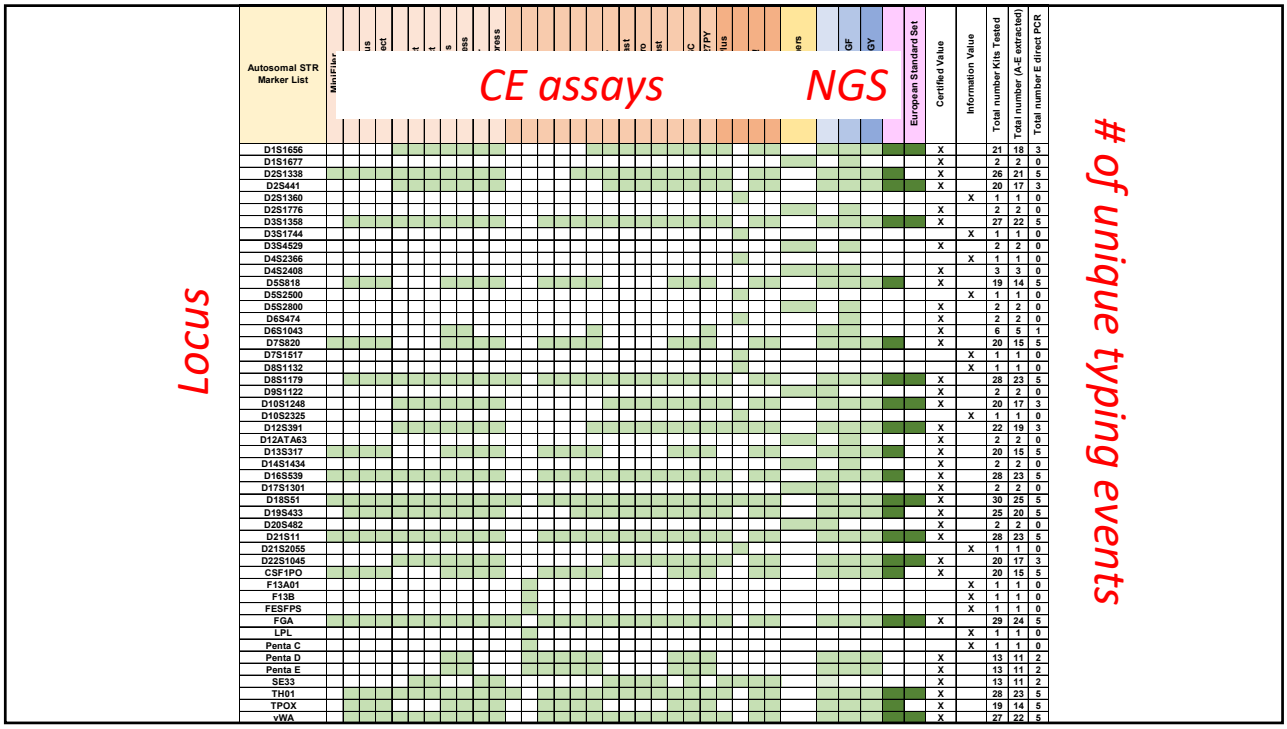
Table 3. Next Generation Sequencing (NGS) methods used for SRM 2391d testing

AFDIL MiSeq FGx	Verogen MiSeq FGx	Thermo Fisher Ion S5 XL	Promega MiSeq FGx	Qiagen MiSeq FGx
AFDIL mtGenome protocol [4] (mtDNAWhole Genome)	ForenSeq Signature Prep Kit	Precision ID GlobalFiler NGS STR Panel v2	PowerSeq 46GY System (prototype)	Human Mitochondrial Panel (mtDNA Whole Genome)
Literature-based assay		Precision ID Ancestry Panel	PowerSeq CRM Nested System (mtDNA Control Region)	
		Precision ID Identity Panel		
		Ion Ampliseq DNA Phenotype Panel		
		Precision ID mtDNA Whole Genome Panel		

Ring, J.D., et al., A performance evaluation of Nextera XT and KAPA HyperPlus for Rapid Illumina Library Preparation of Long-Range Mitogenome Amplicons; (2017) FSI G 29: 174–180

Interlaboratory participation

- Promega
  - CE (7) ESI 17 Fast, ESX 17 Fast, PPY23, PP21, Fusion 6C, Fusion, VersaPlex 27PY
  - NGS (3) PowerSeq 46GY, Mito control region, Mito whole genome
- Verogen
  - NGS ForenSeq (STRs and SNPs)
- Qiagen
  - CE (1) kit (HDplex); Mito whole genome
- AFDIL
  - Analysis of FASTQ files from NIST runs



Autosomal STR Marker List	MiniFiler	Identifiler	Identifiler Plus	Identifiler Direct	NGM	NGM SElect	NGM Detect	Verifiler Plus	Verifiler Express	GlobalFiler	GlobalFiler Express	PP S5	PP CS7	PP 16	PP 16 HS	PP 18D	PP 21	PP ESX 17	PP ESX 17 Fast	PP ESI 17 Pro	PP ESI 17 Fast	PP Fusion	PP Fusion 6C	PP VersaPlex 27PY	ESSplex SE Plus	HDplex	24plex GO!	24plex QS	In-house Primers	ForenSeq	Precision ID GF	PowerSeq 46GY	CODIS 20	European Standard Set	Certified Value	Information Value	Total number Kits Tested	Total number (A-E extracted)	Total number E direct PCR
D1S1656																																			X		21	18	3
D1S1677																																			X		2	2	0
D2S1338																																			X		26	21	5
D2S441																																			X		20	17	3
D2S1360																																				X	1	1	0
D2S1776																																			X		2	2	0
D3S1358																																			X		27	22	5
D3S1744																																				X	1	1	0
D3S4529																																			X		2	2	0

Component	Genotype agreement observed for 18 CE-based typing kits
A	15.3, 18.3
B	13, 15.3
C	15, 16
D	15, 15.3, 16, 18.3
E	11, 16.3

Example D1S1656

Autosomal STR Marker List	MiniFiler	Identifiler	Identifiler Plus	Identifiler Direct	NGM	NGM SElect	NGM Detect	Verifiler Plus	Verifiler Express	GlobalFiler	GlobalFiler Express	PP S5	PP CS7	PP 16	PP 16 HS	PP 18D	PP 21	PP ESX 17	PP ESX 17 Fast	PP ESI 17 Pro	PP ESI 17 Fast	PP Fusion	PP Fusion 6C	PP VersaPlex 27PY	ESSplex SE Plus	HDplex	24plex GO!	24plex QS	In-house Primers	ForenSeq	Precision ID GF	PowerSeq 46GY	CODIS 20	European Standard Set	Certified Value	Information Value	Total number Kits Tested	Total number (A-E extracted)	Total number E direct PCR
D1S1656																																			X		21	18	3
D1S1677																																			X		2	2	0
D2S1338																																			X		26	21	5
D2S441																																			X		20	17	3
D2S1360																																				X	1	1	0
D2S1776																																			X		2	2	0
D3S1358																																			X		27	22	5
D3S1744																																				X	1	1	0
D3S4529																																			X		2	2	0

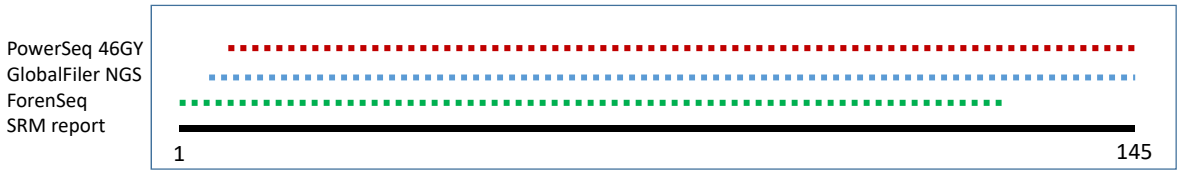
Component	Sequenced by three NGS kits
A	15.3, 18.3
B	13, 15.3
C	15, 16
D	15, 15.3, 16, 18.3
E	11, 16.3

Example D1S1656

Example: Component A, D1S1656, 15.3 allele

D1S1656 15.3 CCTA [TCTA]11 TCA [TCTA]3 rs4847015			
Bases reported	Start	End	
SRM 2391d certificate	1	145	*145*
Repeat region	62	124	*63*
PowerSeq 46GY	6	145	*140*
GlobalFiler NGS	5	145	*141*
ForenSeq	1	124	*124*
STRSeq accession number			MH174852.1

https://www.ncbi.nlm.nih.gov/bioproject/?term=strseq



Example: Component A, D1S1656, 15.3 allele

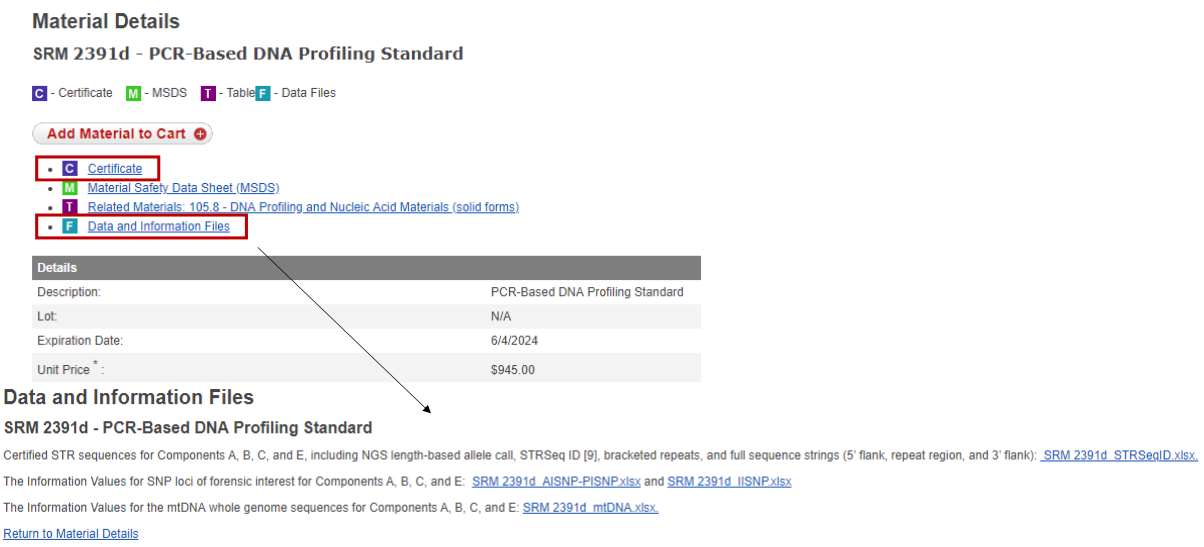
D1S1656 15.3 CCTA [TCTA]11 TCA [TCTA]3 rs4847015			
Bases reported	Start	End	
SRM 2391d certificate	1	145	*145*
Repeat region	62	124	*63*
ForenSeq	1	124	*124*
PowerSeq 46GY	6	145	*140*
GlobalFiler NGS	5	145	*141*
STRSeq accession number			MH174852.1

5′TTCAGAGAAATAGAATCACTAGGGAACCAAATATATATACATACAATTAAACACACAC
ACA[CCTA][TCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTA][TCA]
[TCTATCTATCTA]CATCA**T**ACAGTTGACCCTTGA 3′

rs4847015 (A/T)
@130 bp

Reported string (allele) is a combination from the three kits

https://www-s.nist.gov/srmors/view_detail.cfm?srm=2391d



Material Details

SRM 2391d - PCR-Based DNA Profiling Standard

C - Certificate M - MSDS T - Table F - Data Files

Add Material to Cart

- C Certificate
- M Material Safety Data Sheet (MSDS)
- T Related Materials: 105.8 - DNA Profiling and Nucleic Acid Materials (solid forms)
- F Data and Information Files

Details	
Description:	PCR-Based DNA Profiling Standard
Lot:	N/A
Expiration Date:	6/4/2024
Unit Price * :	$945.00

Data and Information Files

SRM 2391d - PCR-Based DNA Profiling Standard

Certified STR sequences for Components A, B, C, and E, including NGS length-based allele call, STRSeq ID [9], bracketed repeats, and full sequence strings (5' flank, repeat region, and 3' flank): SRM 2391d_STRSeqID.xlsx.

The Information Values for SNP loci of forensic interest for Components A, B, C, and E: SRM 2391d_AISNP-PISNP.xlsx and SRM 2391d_IISNP.xlsx

The Information Values for the mtDNA whole genome sequences for Components A, B, C, and E: SRM 2391d_mtDNA.xlsx.

Return to Material Details

Excel files containing information for STRs, SNPs and mitochondrial sequencing

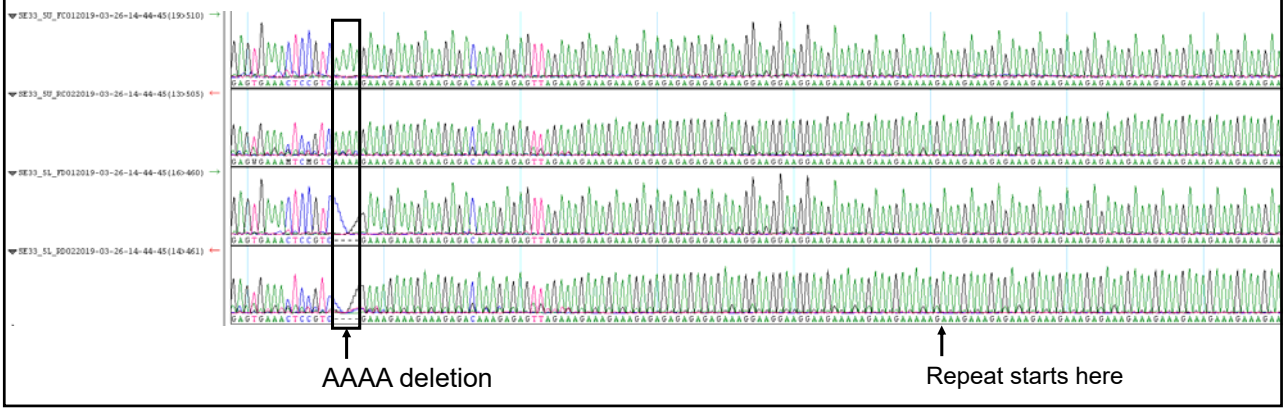
Locus	NGS Allele	STRSeq ID	Bracket	Description	Sequence
CSF1PO	12	MH085189.2	[ATCT]12	CSF1PO 12 [ATCT]12	CTGTGTCAGACCCTGTTCTAAGTACTTCCTATCTATCTATCTATCTATCTATCTATCTATCTATCT
CSF1PO	14	MH085191.2	[ATCT]14	CSF1PO 14 [ATCT]14	CTGTGTCAGACCCTGTTCTAAGTACTTCCTATCTATCTATCTATCTATCTATCTATCTATCTATCT
D10S1248	14	MH167062.2	[GGAA]14	D10S1248 14 [GGAA]14	TAGCCCCAGGACCAATCTGGTCACAAACATATTAATGAATTGAACAAATGAGTGAGTGGAAGG
D10S1248	15	MH167063.2	[GGAA]15	D10S1248 15 [GGAA]15	TAGCCCCAGGACCAATCTGGTCACAAACATATTAATGAATTGAACAAATGAGTGAGTGGAAGG
D12ATA63	13	MK990405.1	[TTG]3 [TTA]10	D12ATA63 13 [TTG]3 [TTA]10	AGCAATTTAAAAATGTTGTTGTTGTTATTATTATTATTATTATTATTATTATTACTTGAGACAGGG
D12ATA63	17	MK990407.1	[TTG]3 [TTA]14	D12ATA63 17 [TTG]3 [TTA]14	AGCAATTTAAAAATGTTGTTGTTGTTATTATTATTATTATTATTATTATTATTATTATTATTAC
D12S391	21	MH167160.1	[AGAT]12 [AGAC]8 AGAT	D12S391 21 [AGAT]12 [AGAC]8 AGAT	CAGAGAGAAAGAATCAACAGGATCAATGGATGCATAGGTAGATAGATAGATAGATAGATAGATAGAT
D12S391	24	MH167187.1	[AGAT]15 [AGAC]8 AGAT	D12S391 24 [AGAT]15 [AGAC]8 AGAT	CAGAGAGAAAGAATCAACAGGATCAATGGATGCATAGGTAGATAGATAGATAGATAGATAGATAGAT
D13S317	9	MH167206.1	[TATC]9	D13S317 9 [TATC]9	TTCTTTAGTGGGCATCCGTGACTCTCTGGACTCTGACCCATCTAACGCCTATCTGTATTTACAAA
D13S317	12	MH167226.1	[TATC]12	D13S317 12 [TATC]12	TTCTTTAGTGGGCATCCGTGACTCTCTGGACTCTGACCCATCTAACGCCTATCTGTATTTACAAA

Some interesting examples…

Some interesting examples: SE33

- Component B typed as (17,28.2) with *thirteen* CE-STR kits
- Sequencing revealed 18 repeats with 4 bp AAAA deletion 85 upstream from the repeat

$[AAAG]_2$ AG $[AAAG]_3$ AG **$[AAAG]_{18}$** G $[AAAG]_3$ AG = 18 repeats



Some interesting examples: D6S474

- Complex motif [**AGAT**][**GATA**] as described in Hill et al. 2008
- Reported as such in SRM2391b and 2391c
- Precision ID GlobalFiler NGS reporting [**AGAT**][**GATA**]
- Investigator HDplex reporting one repeat less [**TAGA**]TGA[**TAGA**]

Component	SRM certificate	Investigator HDplex kit
A	16,18	15,17
B	14,16	13,15
C	14,18	13,17
D	14,16,18	13,15,17
E	14,16	13,15

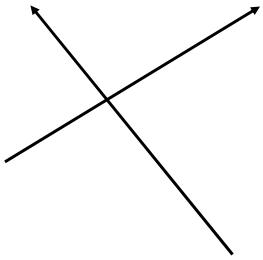
Some interesting examples: DYS612

Six repeat difference

Component	SRM certificate	ForenSeq report
A	34	28
B	34	28

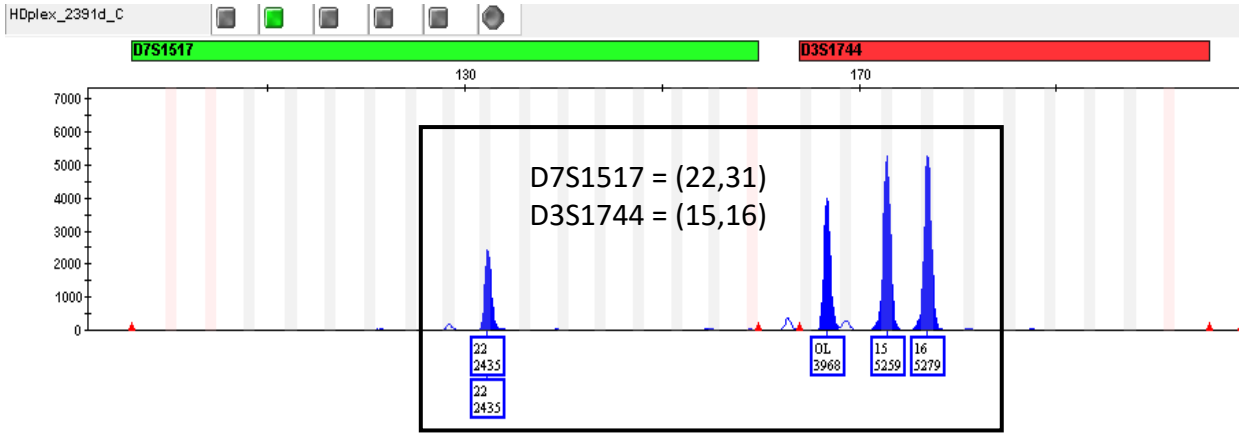
Kayser, et al. Am. J. Hum. Genet. 2004 (*first described*)
(CCT)5(CTT)1(TCT)4(CCT)1(TCT)n

D'Amato, et al. FSI Genetics 2010 (*a proposed nomenclature)*
(CCT)5(CTT)1(TCT)4(CCT)1(TCT)n



Ballantyne et al. Human Mutation 2014 (*updated to comply with ISFG guidelines)*
(CCT)5(CTT)1(TCT)4(CCT)1(TCT)n

Some interesting examples: Qiagen Investigator HDplex – Component C



Allele calls confirmed by Sanger sequencing

Identity SNP markers

- 101 autosomal SNPs reported
  - ForenSeq (94)
  - Precision ID Identity Panel (90) + 34 Y-SNPs

83 identity autosomal SNPs in common

- Forward strand genotype reported

	Locus	2391d_A	Coverage FS	Coverage ID	ACR FS	ACR ID	2391d_B	Coverage FS	Coverage ID	ACR FS	ACR ID	2391d_C	Coverage FS	Coverage ID	ACR FS	ACR ID	2391d_E	Coverage FS	Coverage ID	ACR FS	ACR ID
ForenSeq Only	rs1294331	CT			●		TT					CT			◕		CT			◕	
	rs13182883	GG					AG			●		AG			●		AG			●	
	rs1336071	CC					CC					CT			◕		CC				
	rs2107612	AA					AG			●		AG			●		AA				
	rs2399332	GG					TT					GT			●		GT			●	
	rs279844	AT			●		AT			●		AA					AT			●	
	rs2920816	AG			◕		AA					AG			●		AA				
	rs4606077	CC					CC					CT			●		CC				
	rs763869	AG			●		AG			●		AG			●		AG			●	
	rs8037429	CC	●				TT	●				CT	●		●		CC				
	rs8078417	CC	●				TT	●				CT	●		●		CT	●		●	

Coverage

FS	ID
> 2000	> 30,000
2000 < = 1000	30,000 < = 20,000
1000 < = 200	20,000 < = 10,000
< 200	< 10,000

ACR
> 0.8
0.8 < = 0.6
0.6 < = 0.4
0.4 < = 0.2
< 0.2

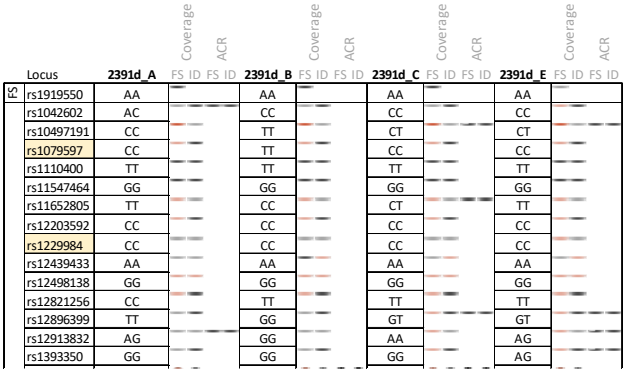
Provided as information values

Ancestry and Phenotype SNP markers

- Ancestry/Phenotype SNPs 188 total
  - ForenSeq (78)
  - Precision ID Ancestry (165) and Phenotype Panel (24)

77 SNPs in common

- Forward strand genotype reported



Coverage

FS	ID
> 2000	> 20,000
2000 < = 1000	20,000 < = 10,000
1000 < = 200	10,000 < = 5,000
< 200	< 5,000

ACR

> 0.8
0.8 < = 0.6
0.6 < = 0.4
0.4 < = 0.2
< 0.2

Genotypes are reported not the application (eye color, ancestry)

Summary of ancestry and phenotype marker estimates

Component	ForenSeq			Precision ID				Mito	Y SNP
	Ancestry	*Hair*	*Eye*	*Ancestry*	*Hair*	*Hair*	*Eye*		
A	European	0.68	0.66	European	0.66	1.00 light	0.67	T2b3	-
B	African	0.69	0.86	African	0.66	0.93 light	0.85	L1c1a	E
C	African	0.84	1.00	African	0.68	1.00 dark	1.00	L1b1a	E
E	European	0.61	0.71	European/SW Asian	0.69	0.72 light	0.72	T2a3	-

Predictions made using vendor tools

Mitochondrial genome sequencing

NGS platform	Assay	Data Analysis Method
Ion Torrent S5XL	Precision ID mtDNA Whole Genome Panel	HID Genotyper Plugin v2.1 Converge v2.1
MiSeq FGx	Qiagen Human Mitochondrial Panel	GeneGlobe
MiSeq FGx	AFDIL mtGenome protocol	GeneMarker HTS v1.2.2 CLC bio Genomics Workbench v7/AQME
MiSeq FGx	Promega CRM Nested & WG	GeneMarker HTS v1.2.2

Heteroplasmy calling filter set at 5%

Can this material replace the SRM 2392 and 2392-I (Mitochondrial DNA standards?)

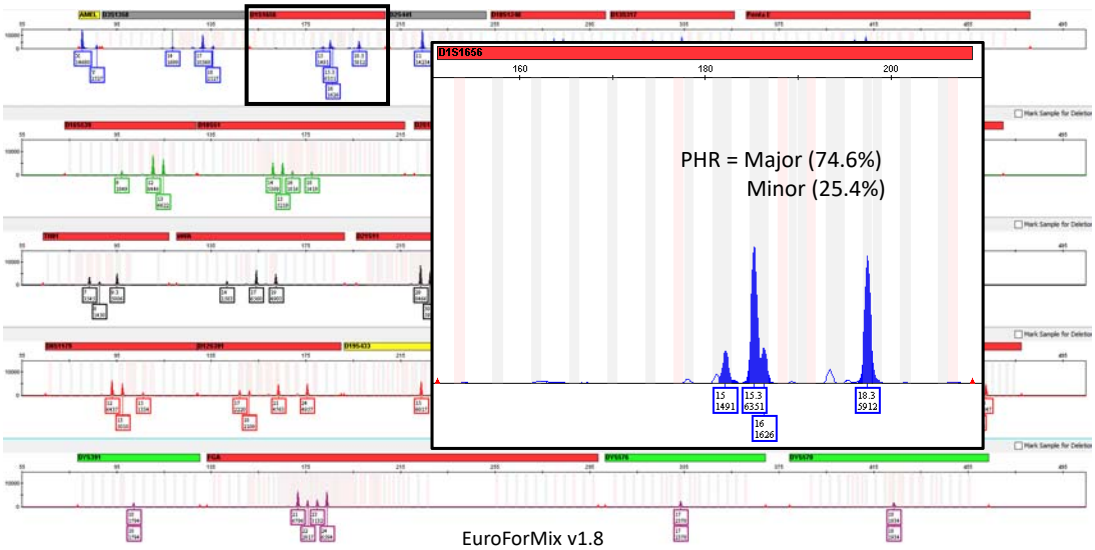
Your feedback is needed!

SRM 2391d Component E

Site	rCRS	Component E	Comment
73	A	G	
195	T	C	
263	A	G	
309.1	:	C	Insertion
315.1	:	C	Insertion
709	G	A	
750	A	G	
1438	A	G	
1888	G	A	
2706	A	G	
3107	C	:	Deletion
4216	T	C	
4464	G	A	
4769	A	G	
4917	A	G	
7028	C	T	
8697	G	A	
8860	A	G	
9100	A	G	
10463	T	C	
11251	A	G	
11719	G	A	
11812	A	G	
12408	T	C	
13368	G	A	
13965	T	C	
14233	A	G	
14766	C	T	
14905	G	A	
15326	A	G	
15452	C	A	
15607	A	G	
15928	G	A	
16126	T	C	
16294	C	T	
16519	T	C	

Mito Hg
A T2b3+151
B L1c1a+@198
C L1b1a
E T2a3

Component D - 3:1 mixture of A (female) and C (male)



EuroForMix v1.8

	Mixture Proportion	
	Average	Stdev
Mix-prop. C1	73.5%	0.5%
Mix-prop. C2	26.5%	0.5%

Summary

- The SRM 2391 series will continue to support the FBI-QAS and the validation and implementation of forensic marker systems

Marker Type	Number of Certified loci	Number of loci with Information values
Autosomal STR	35	13
Y-STR	28	3
X-STR	7	5
Mitochondrial DNA	-	Full mtGenome
Indel/Innuls	-	50
SNPs	-	323

Certified allele calls supported by sequence data and CE-length based measurements

Thank you for your attention! Questions?

Contact: Peter.Vallone@nist.gov

- **Disclaimer** - Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Commerce.
- 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 NIST, nor does it imply that any of the materials, instruments, or equipment identified are necessarily the best available for the purpose.
- All work presented has been reviewed and approved by the NIST Human Subjects Protections Office.

Margaret Kline
Dr. Dave Duewer