Development of the **NIST SRM 2391d**: PCR-Based DNA Profiling Standard

Where Are We Now?

Becky Steffen, Erica Romsos, Kevin Kiesler, Katherine Gettings, Sarah Riman, Lisa Borsuk and Peter Vallone
Applied Genetics Group
National Institute of Standards and Technology

American Academy of Forensic Sciences
Baltimore, MD
February 23, 2019

NIST National Institute of Standards and Technology U.S. Department of Commerce

Official Disclaimer

The opinions and assertions contained herein are solely those of the author and are not to be construed as official or as views of the U.S. Department of Commerce.

Commercial equipment, instruments, software, or 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 U.S. Department of Commerce, nor does it imply that any of the materials, instruments, software or equipment identified are necessarily the best available for the purpose.

NIST National Institute of Standards and Technology U.S. Department of Commerce

Development of the Next PCR-Based DNA Profiling Standard

- As a successor to SRM 2391c
  - Inventory may be depleted by fall 2019
  - Develop SRM 2391d now to ensure availability when needed

Note: SRM 2391c can still be used until the expiration date (February 3, 2020) if stored properly.

- **Next Generation Sequencing** is used for certification in addition to **Capillary Electrophoresis** testing
  - Length- and sequence-based genotypes are provided
  - Information values are included for all commercially available forensic markers **beyond STR markers**

Goal: SRM 2391d will be the most comprehensive NIST forensic SRM to date

How are SRM 2391d values assigned?

- **NIST Certified Values** are assigned when multiple CE primer sets **AND** sequencing results are compared

***Highest confidence*; all sources of uncertainty and bias examined**

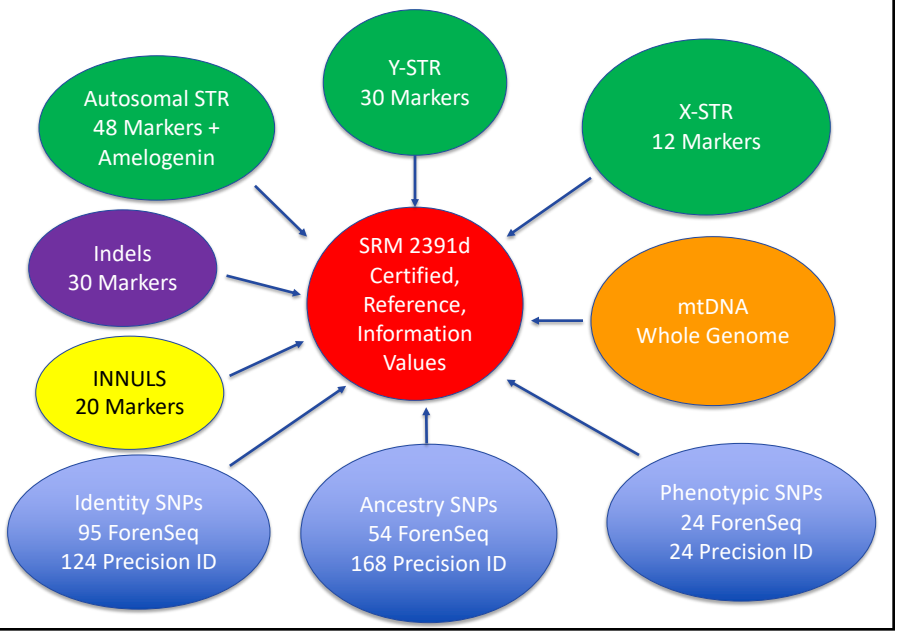
- **Reference Values** are assigned when multiple CE primer sets **OR** sequencing results are compared

***Fit for purpose*; not all sources of uncertainty have been examined**

- **Information Values** are assigned when only one primer set is used from either CE or sequencing

***For informational purposes*; no guarantees for uncertainty**

Markers included in the Certificate of Analysis



Which Autosomal STR Markers have Certified Values?

Autosomal STR Markers
Thermo Fisher CE STR kits
Promega CE STR kits
Qiagen Investigator CE STR kits

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	ESSplex SE Plus	HDplex	24plex GO!	24plex QS	ForenSeq	Precision ID GF	PowerSeq 46GY	CODIS 20	European Standard Set	Certified Value	Reference Value	Information Value
D1S1656																																	X		
D1S1677																																		X
D4S2408																																	X	

24 Certified Autosomal STR Markers
1 Reference Autosomal STR Marker
23 Information Autosomal STR Markers

Which Y-STR Markers have Certified Values?

Y-STR Markers
Thermo Fisher CE STR kits
Promega CE STR kits
Qiagen Investigator CE STR kits
Verogen NGS kit
Thermo Fisher NGS kits
Promega NGS kits

23 Certified Y-STR Markers
0 Reference Y-STR Markers
7 Information Y-STR Markers

Y-STR Marker List	GlobalFiler	GlobalFiler Express	Yfiler	Yfiler Plus	PP Fusion	PP Fusion 6C	PowerPlex Y23	24plex GO!	24plex QS	ForenSeq	Precision ID GF	PowerSeq 46GY	Certified Value	Reference Value	Information Value
DYS19													X		
DYS385a/b													X		
DYS389I/II													X		
DYS390													X		
DYS391													X		
DYS392													X		
DYS393													X		
DYS437													X		
DYS438													X		
DYS439													X		
DYS448													X		
DYS449															X
DYS456													X		
DYS458													X		
DYS460													X		
DYS461															X
DYS481													X		
DYS505															X
DYS518															X
DYS522															X
DYS533													X		
DYS549													X		
DYS570													X		
DYS576													X		
DYS612															X
DYS627															X
DYS635													X		
DYS643													X		
Y-GATA-H4													X		
DYS387S1													X		

Which X-STR Markers have Certified Values?

X-STR Markers
Qiagen Investigator CE STR kit
Verogen NGS kit

7 Certified X-STR Markers
0 Reference X-STR Markers
5 Information X-STR Markers

X-STR Marker List	Argus X-12	ForenSeq	Certified Value	Reference Value	Information Value
DXS7132			X		
DXS7423			X		
DXS8378			X		
DXS10074			X		
DXS10079					X
DXS10101					X
DXS10103			X		
DXS10134					X
DXS10135			X		
DXS10146					X
DXS10148					X
HPRTB			X		

What Platforms Were Used for Testing?

- **Capillary Electrophoresis (CE)** was performed with one instrument:
  - 3500xL Genetic Analyzer (ThermoFisher)
- **Next Generation Sequencing (NGS)** was performed with two different instruments:
  - MiSeq FGx (Verogen)
  - Ion S5 XL (ThermoFisher)

3500xl

MiSeq FGx

Ion S5 XL

Commercial CE Kits that were tested (33 Kits Total)

Thermo Fisher (13)	Promega (13)	Qiagen (6)	InnoGenomics (1)
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 Y23		

Commercial NGS Kits that were tested (11 Kits Total)

AFDIL/MiSeq (1)	Verogen/MiSeq (1)	Thermo Fisher/Ion S5 XL (5)	Promega/MiSeq (2)	Qiagen/MiSeq (2)
mtDNA Whole Genome	ForenSeq Signature Prep Kit	Precision ID GlobalFiler NGS STR Panel v2	PowerSeq 46GY (prototype)	QIAseq mtDNA Whole Genome Panel
		Precision ID Ancestry Panel	PowerSeq CRM Nested System (mtDNA control region)	QIAseq SNP Panel
		Precision ID Identity Panel		
		Precision ID Phenotype Panel		
		Precision ID mtDNA Whole Genome Panel		

Ring *et al.* (2017)




What is Included in SRM 2391d?

- **Sample format**:
  - 4 extracted DNA samples
    - 3 single source and 1 mixed sample at a 3:1 ratio (female:male)
  - 1 cell line (female) spotted onto FTA paper as intact cells
  - 5 samples total: Components A-E
- **Concentration of the samples is ~1.5 ng/µL DNA for the extracted DNA (A-D) and 7.5 x 10^4 cells spotted on FTA paper (E)**
  - The concentrations were determined by droplet digital PCR (ddPCR)
  - The concentrations will NOT be certified values – they will be reported as information values

Component	ng/µL	U(ng/µL)
A	1.5	0.3
B	1.7	0.3
C	1.6	0.2
D	1.5	0.3

Components A-D will have *different profiles* from SRM 2391c
Component E will have *the same profile* as SRM 2391c

Where Are We Now?

- All components have been diluted (A-D), spotted (E), bottled and labeled
- Homogeneity Testing has been completed
- Stability Testing up to 30 weeks has been completed (4ºC, 22ºC, and 37ºC)

Statistical Engineering Division (SED) approved

- All CE testing has been completed (33 kits)
- NGS testing is almost complete (11 kits)
- Data review process and concordance evaluations are in progress
- Reports of Analysis and Certificate of Analysis are in progress

SRM 2391d will be released in summer 2019

Autosomal STR Marker Certified Values (24)

	Component A		Component B		Component C		Component D				Component E	
Locus	Allele 1	Allele 2	Allele 1	Allele 2	Allele 1	Allele 2	Allele 1	Allele 2	Allele 3	Allele 4	Allele 1	Allele 2
AMEL	X	X	X	Y	X	Y	X	Y			X	X
D1S1656	15.3	18.3	13	15.3	15	16	15	15.3	16	18.3	11	16.3
D2S1338	25	25	17	23	23	24	23	24	25		19	20
D2S441	11	11	11	11	11	14	11	14			10	10
D3S1358	17	17	15	17	14	18	14	17	18		14	15
D5S818	10	11	12	12	13	15	10	11	13	15	11	13
D6S1043	12	19	13	18	11	18	11	12	18	19	11	11
D7S820	8	10	10	10	9	10	8	9	10		8	10
D8S1179	12	13	12	15	12	15	12	13	15		11	13
D10S1248	14	15	12	15	12	16	12	14	15	16	14	14
D12S391	21	24	19	20	17	18	17	18	21	24	17	22
D13S317	9	12	11	11	12	14	9	12	14		8	12
D16S539	12	13	9	11	9	12	9	12	13		11	12
D18S51	14	15	17	18	16	18	14	15	16	18	14	17
D19S433	13	15	11	16.2	13	15	13	15			14	14
D21S11	29	30	28	29	29	31	29	30	31		29	30
D22S1045	14	16	12	15	14	15	14	15	16		16	17
CSF1PO	12	14	12	12	10	11	10	11	12	14	10	11
FGA	21	24	24	26	22	23	21	22	23	24	20	23
Penta D	8	9	11	13	9	13	8	9	13		14	14
Penta E	13	14	5	7	12	14	12	13	14		13	19
SE33	17	28.2	17	28.2	17	18	17	18	28.2		22	30.2
TH01	7	9.3	7	7	8	9.3	7	8	9.3		6	9.3
TPOX	8	9	8	12	8	10	8	9	10		8	11
vWA	17	19	15	17	14	17	14	17	19		17	18

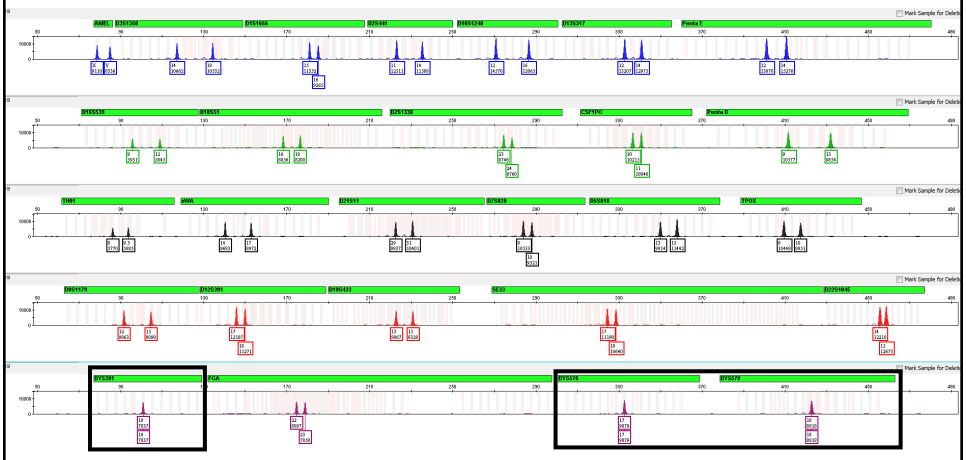
Y-STR Marker Certified Values (23)

	Component B		Component C		Component D	
Locus	Allele 1	Allele 2	Allele 1	Allele 2	Allele 1	Allele 2
DYF387S1	36	38	36	39	36	39
DYS19	15		16		16	
DYS385	15	16	16	17	16	17
DYS389I	12		12		12	
DYS389II	30		31		31	
DYS390	21		21		21	
DYS391	11		10		10	
DYS392	11		11		11	
DYS393	13		13		13	
DYS437	14		14		14	
DYS438	11		11		11	
DYS439	13		12		12	
DYS448	21		22		22	
DYS456	15		15		15	
DYS458	17		18		18	
DYS460	10		10		10	
DYS481	26		28		28	
DYS533	11		11		11	
DYS549	11		12		12	
DYS570	20		18		18	
DYS576	15		17		17	
DYS635	21		21		21	
DYS643	15		14		14	
YGATAH4	13		12		12	

*Components A and E are females and therefore do not have a Y chromosome

Example Data: Component C

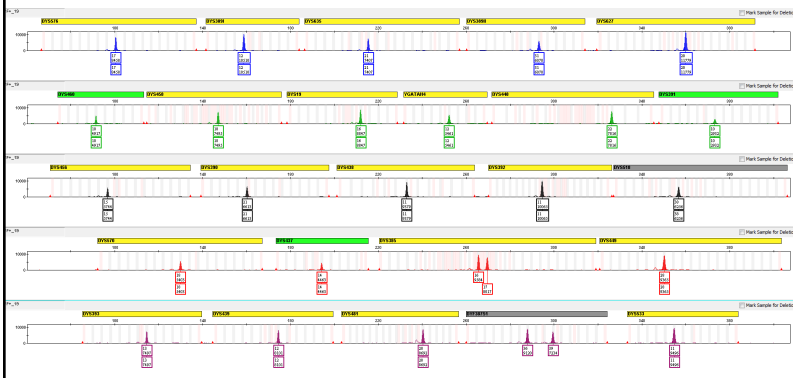
Autosomal STR CE Profile: PowerPlex Fusion 6C



*Y-STR Markers

Fully Heterozygous Profile with PowerPlex Fusion 6C

Y-STR CE Profile: Yfiler Plus



YHRD: No matches in 188,209 Haplotypes
(using Minimal Haplotype)
https://yhrd.org

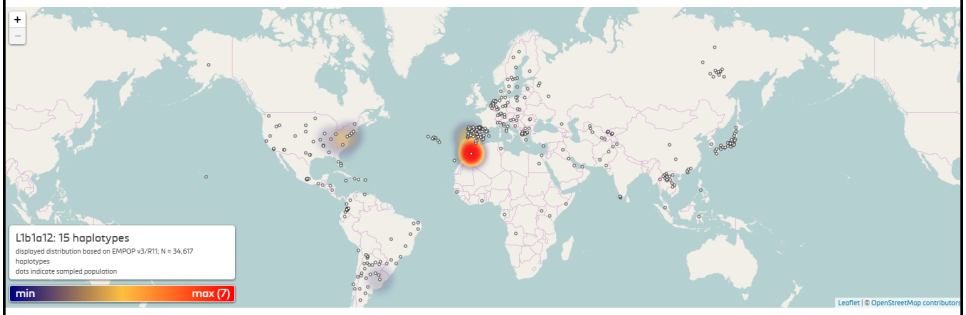
Whit Athey's Haplogroup Predictor: E1b1a
http://www.hprg.com/hapest5/hapest5a/hapest5.htm?order=num

Results Table

Haplo-group	Fitness score	Proba-bility (%)
E1b1a	58	100.0
E1b1b	18	0.0
G2a	20	0.0
G2c	5	0.0
H	25	0.0
I1	7	0.0
I2a (xI2a1)	20	0.0
I2a1	3	0.0
I2b (xI2b1)	7	0.0
I2b1	14	0.0
J1	11	0.0
J2a1b	5	0.0
J2a1h	6	0.0
J2a1 x J2a1-bh	11	0.0
J2b	9	0.0
L	13	0.0
N	2	0.0
Q	17	0.0
R1a	11	0.0
R1b	4	0.0
T	16	0.0

mtDNA Whole Genome Sequencing

AFDIL mtDNA Whole Genome Sequencing protocol (MiSeq), Ring *et al.*, 2017



EMPOP results:
https://empop.online/haplotypes#matches_details

Haplogroup	Ancestry	Match
L1b1a12	African	unique

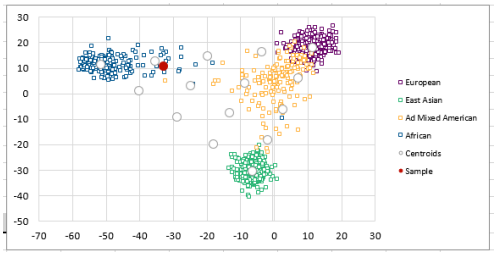
Ring, J.D., Sturk-Angreaggi, K., Peck, M.A., Marshall, C. (2017) A performance evaluation of Nextera XT and KAPA HyperPlus for rapid Illumina library preparation of long-range mitogenome amplicons. Forensic Sci Int Genet 29:174-180.

SNP Phenotype and Ancestry Estimation (ForenSeq DNA Signature Prep Kit)

Hair Color Results	
Brown	0.16
Red	0.00
Black	0.84
Blond	0.00

Eye Color Results	
Intermediate	0.00
Brown	1.00
Blue	0.00

Biogeographical Ancestry Results	
Distance to Nearest Centroid	3.36



Population(Region, sampleSize 2N)	Probability of Genotype in each Population	Likelihood Ratio
Somali(Africa,40)	1.576E-15	
African American(ASW)(Africa,122)	3.044E-16	5.18
Sandawe(Africa,80)	1.824E-16	8.64
Ethiopian Jews(Africa,64)	1.032E-16	15.3
African Americans(Africa,182)	7.118E-17	22.1
Masai(Africa,44)	8.17E-18	193.0
Chagga(Africa,90)	1.289E-18	1220.0
Luhya(LWK)(Africa,198)	4.072E-20	38700.0
Lisongo(Africa,16)	3.211E-20	49100.0
Hausa(Africa,78)	4.487E-21	351000.0

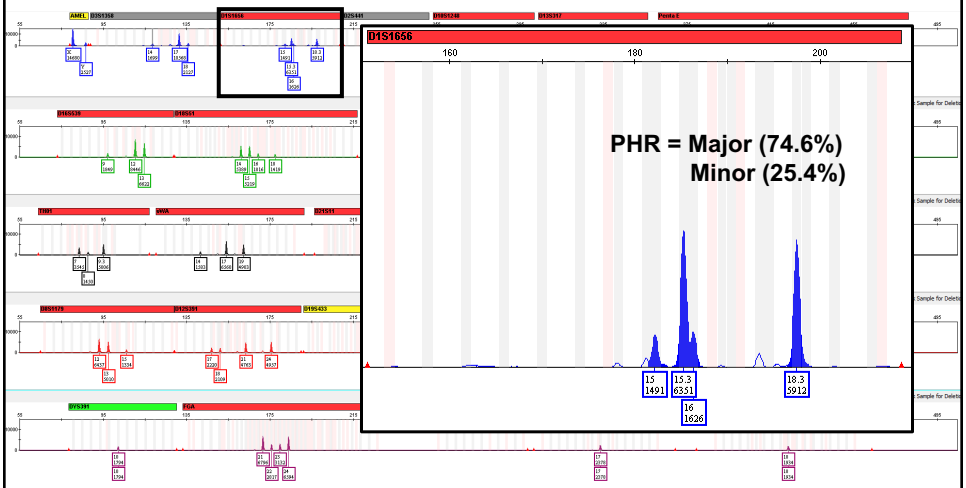
Indicates the values are within an order of magnitude of the highest likelihood.

KiddLab – Set of 55 AISNPs
Population likelihoods based on 55 SNPs and 139 reference populations for this DNA profile
http://frog.med.yale.edu/FrogKB/

Other Markers Determined: X-STRs, Indels, INNULS, and other SNP Panels

Example Data: Component D (3:1 Mixture)

Autosomal STR CE Profile: PowerPlex Fusion 6C



Component D:
3:1 mixture of Components A + C
(female to male)

EuroForMix v1.8

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

What can you use SRM 2391d for?

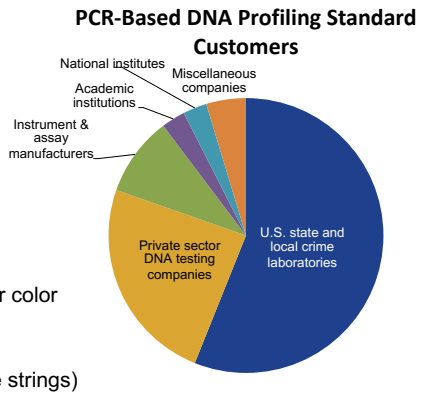
- To meet the FBI Quality Assurance Standards: QAS 9.5.5

> 9.5.5 The laboratory shall check its DNA procedures annually or whenever substantial changes are made to a procedure against an appropriate and available NIST standard reference material or standard traceable to a NIST standard.

- Validation Studies: instrument, commercial kit, and software
  - Developmental and Internal Validations
  - Known, ***well-characterized*** samples for all systems commercially available
- Make NIST traceable materials: http://ts.nist.gov/traceability/

Support to the Forensic Community

- PCR-Based DNA Profiling Standard Customers
  - U.S. state and local crime laboratories
  - Private sector DNA testing companies
  - Instrument and assay manufacturers
  - Academic institutions
  - National institutes
  - Miscellaneous companies/industry



- **Emerging Forensic Technology**
  - New Markers
    - CODIS 13 → CODIS 20: January 1, 2017
    - New SNP markers for ancestry and eye/hair color predictions
  - New Methods
    - Next Generation Sequencing (full sequence strings)
    - New CE instruments and STR kits

Summary and Final Thoughts

- The next PCR-Based DNA Profiling Standard is being developed as the most ***comprehensive*** forensic SRM yet
  - STR genotypes and haplotypes
  - Information from commercially available forensic markers beyond the STR markers
- Capillary Electrophoresis and Next Generation Sequencing has been performed to assign certified, reference, and information values to the final components
- SRM 2391d can be used for validation studies and to support the forensic community as new technologies emerge
- SRM 2391d will be available in summer 2019

Thank you for your attention!

NIST APPLIED GENETICS

Questions?

becky.steffen@nist.gov

1-301-975-4275

Acknowledgements

Margaret Kline

David Duewer

All work presented has been reviewed and approved by the NIST Human Subjects Protections Office

Funding: NIST Special Programs Office: *Forensic DNA*

A copy of this presentation is available at: http://strbase.nist.gov/NISTpub.htm#Presentations