Make it "SNPpy"
Updates to SRM 2391D: PCR-Based Profiling Standard

Carolyn R. Steffen, Erica L. Romsos, Kevin M. Kiesler, Lisa A. Borsuk, Katherine B. Gettings, and Peter M. Vallone

U.S. National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD 20899-8314, USA

Email: Becky.Steffen@nist.gov

P178

Standard Reference Material (SRM) 2391d: PCR-Based DNA Profiling Standard was released to the forensic community in 2019. Next Generation Sequencing (NGS) was used as the primary method of certification, where certified values were assigned when a high coverage sequence string was available for a marker. Using NGS to assign values has allowed for additional marker sets beyond short tandem repeat (STR) loci, including single nucleotide polymorphisms (SNPs) and mitochondrial DNA (mtDNA) whole genome sequences, to be included in the Certificate of Analysis (COA). Since the 2019 release, several commercial NGS panels have become available including the Verogen ForenSeq mtDNA Control Region, mtDNA Whole Genome, MainstAY, and Kintelligence Kits. In addition, three custom Ion AmpliSeq panels from Thermo Fisher (MH-74 Plex, VISAGE, and Y-SNP) are now available. While the mtDNA whole genome sequence for the components are already included and no new STR markers are introduced by MainstAY, the other recently released panels allow for the inclusion of >11,000 additional SNPs (e.g., identity, ancestry, phenotype, kinship, and X- and Y-SNPs) and 74 microhaplotypes to the COA for SRM 2391d in an update to be completed by Fall of 2022.

2019 Release

National Institute of Standards & Technology

Certificate of Analysis

Standard Reference Material® 2391d

PCR-Based DNA Profiling Standard

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

Materials (Five Components)

Components A-D are genomic DNA extracted from purchased blood

- Not from cell lines (challenges in obtaining 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 expiration 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)**

Table 1. Description of Components in SRM 2391d — **Information Values**

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.

Concentrations were assigned using droplet digital PCR (ddPCR)

How are SRM 2391d values assigned?

Certified Values are assigned when there is high coverage sequencing data available for a marker

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

Information Values are assigned when only one primer PCR set is used from CE testing and there is no supporting sequencing data

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

What Platforms Were Used for Testing?

Capillary Electrophoresis (CE) was performed with one instrument:
- 3500xL Genetic Analyzer (Thermo Fisher)

Next Generation Sequencing (NGS) was performed with two different instruments:
- MiSeq FGx (Verogen)
- Ion S5 XL (Thermo Fisher)

3500xl

MiSeq FGx

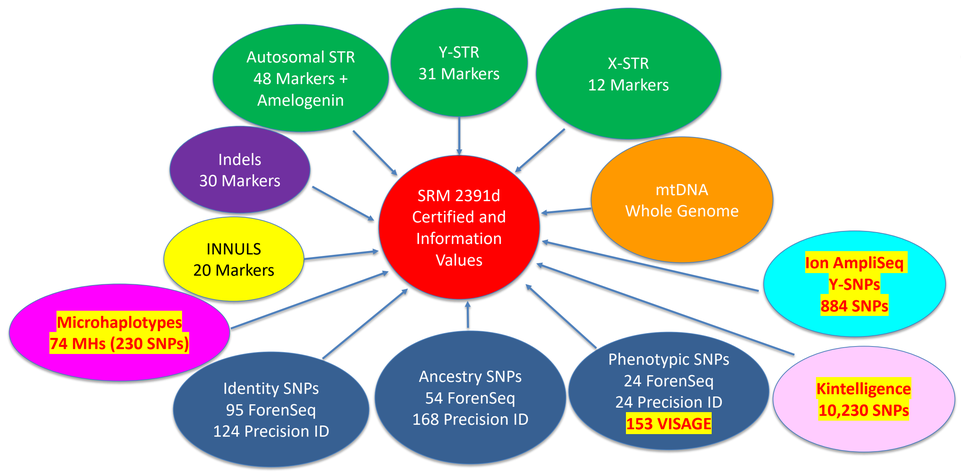
Ion S5 XL

Using SRM 2391d in a Forensic Lab

- To meet the FBI Quality Assurance Standards (Std 8.4)
- Validation Studies: instrument, commercial kit, and software
  - Developmental and Internal Validations
  - Known, ***well-characterized*** samples for all marker systems commercially available
- Make NIST traceable materials: http://ts.nist.gov/traceability/

Markers Included in the Certificate of Analysis

All new marker sets and kits added are highlighted in yellow



Summary of Values Assigned (2022)

Marker Type	Number of Certified Loci	Number of Information Loci
Autosomal STR	35	13
Y-STR	28	3
X-STR	7	5
Mitochondrial DNA	-	Full mtGenome
Indel/Innuls	-	50
SNPs	-	323 + 153 + 884 + 10,230 → 11,590
Microhaplotypes	-	74

Capillary Electrophoresis Kits tested (40 Kits Total)

Thermo Fisher (15)	Promega (13)	Qiagen (11)	InnoGenomics (1)
Minifiler	PowerPlex CS7	Investigator ESSplex SE Plus	InnoTyper 21
Identifiler	PowerPlex 16	Investigator ESSplex SE QS	
Identifiler Plus	PowerPlex 16 HS	Investigator IDplex Plus	
Identifiler Direct	PowerPlex 18D	Investigator IDplex GO!	
NGM	PowerPlex 21	Investigator HDplex	
NGM SElect	PowerPlex ESX 17	Investigator 24plex QS	
NGM Select Express	PowerPlex ESX 17 Fast	Investigator 24plex GO!	
NGM Detect	PowerPlex ESI 17 Pro	Investigator 26plex QS	
Verifiler Express	PowerPlex ESI 17 Fast	Investigator Argus Y-28 QS	
Verifiler Plus	PowerPlex Fusion	Investigator Argus X-12 QS	
*GlobalFiler	PowerPlex Fusion 6C	Investigator DIPplex	
*GlobalFiler IQC	VersaPlex 27PY		
*GlobalFiler Express	PowerPlex Y23		
Yfiler			
Yfiler Plus			

*GlobalFiler, GlobalFiler IQC, and GlobalFiler Express Y indel will be included as an Information Value in the 2022 Update

Sequencing Kits/Methods tested (18 Kits Total)

AFDIL MiSeq FGx (1)	Verogen MiSeq FGx (5)	Thermo Fisher Ion S5 XL (9)	Promega MiSeq FGx (2)	Qiagen MiSeq FGx (1)
AFDIL mtGenome protocol (mtDNAWhole Genome)	ForenSeq Signature Prep Kit	Precision ID GlobalFiler NGS STR Panel v2	*PowerSeq 46GY System	Human Mitochondrial Panel (mtDNA Whole Genome)
	ForenSeq MainstAY Kit	Precision ID Ancestry Panel	PowerSeq CRM Nested System, Custom (mtDNA Control Region)	
	ForenSeq mtDNA Whole Genome Kit	Precision ID Identity Panel		
	ForenSeq mtDNA Control Region Kit	Precision ID mtDNA Whole Genome Panel		
	ForenSeq Kintelligence Kit	Precision ID mtDNA Control Region Panel		
		Ion Ampliseq DNA Phenotype Panel		
		Ion Ampliseq MH-74 Plex Panel		
		Ion Ampliseq VISAGE Panel		
		Ion Ampliseq Y-SNP Panel		

Ring *et al.* (2017)

*Prototype PowerSeq 46GY was tested previously; however, PowerSeq 46GY is now available commercially, so it was retested with SRM 2391d Components

>300 units sold to date since June of 2019

References:
[1] SRM 2391d: PCR-Based DNA Profiling Standard Certificate of Analysis (2019). Available online at https://www-s.nist.gov/srmors/certificates/2391d.pdf. Accessed August 16, 2022.
[2] Thompson, A.; Taylor, B.N.; Guide for the Use of the International System of Units (SI); NIST Special Publication 811; U.S. Government Printing Office: Washington, DC (2008); available at https://www.nist.gov/physical-measurement-laboratory/special-publication-811. Accessed August 16, 2022.

Funding: This work was supported by the NIST Special Programs Office: Forensic Genetics.

Acknowledgements: We would like to thank Tunde Huszar for her insight and data analysis with some of the new marker systems for this update.

2022 Update

Why are we doing an update to SRM 2391d?

- To confirm stability of the Components A-E
  - Digital PCR (dPCR) to confirm quantitation
  - Quantitative PCR (qPCR) to check degradation
  - CE testing to check profiles
- To include new CE kits and NGS kits that have been released since 2019
  - Queried vendors about which kits to include
    - Thermo Fisher, Promega, QIAGEN, and Verogen
  - 8 total CE kits
  - 9 total NGS panels
- To extend the June 4, 2024 expiration date by 5 years

New Marker Sets Included in 2022 Update

- **ForenSeq Kintelligence Kit (Verogen, MiSeq FGx)**
  - **10,230** Ancestry, Identity, Kinship, Phenotype, X-SNPs, and Y-SNPs
  - Used in Forensic Genetic Genealogy (FGG)
  - Data compatible with GEDmatch PRO
- **Ion AmpliSeq VISAGE Panel (Thermo Fisher, Ion S5XL)**
  - Panel developed by the VISAGE Consortium
  - **153** externally visible characteristics (EVCs) and biogeographical ancestry (BGA) SNPs
  - 41 EVCs SNPs for eye, hair and skin color from HIrisPlex-S and 115 BGA SNPs (three overlap with the EVCs SNP set)
- **Ion Ampliseq Y-SNP Panel (Thermo Fisher, Ion S5XL)**
  - Panel developed by Arwin Ralf at Erasmus MC
  - **884** Y-SNPs, inferring of 640 Y haplogroups
- **Ion AmpliSeq MH-74 Panel (Thermo Fisher, Ion S5XL)**
  - Panel developed by GWU, Ken Kidd, and Thermo Fisher
  - **74** Microhaplotypes (230 SNPs)

ForenSeq Kintelligence Kit

Table 2: Kit SNP content

Category	Number of SNPs	Percentage of Total
Ancestry SNPs	56	0.5%
Identity SNPs	94	1%
Kinship SNPs	9867	96%
Phenotype SNPs*	22	0.2%
X-SNPs	106	1.2%
Y-SNPs	85	0.9%

* Two SNPs overlap the ancestry and phenotype categories and are counted in the phenotype category only.

GEDmatch PRO → The largest database of voluntarily submitted DNA profiles for forensic comparisons

Kintelligence Data – Component C

2391d_C_KIN_R2 Kinship SNP Report

Sample Name	2391d_C_KIN_R2
Project Name	Kintelligence_2391d
Analysis Method	Verogen Kintelligence Analysis Method
Run Name	Kintelligence_2391d_run2
Gender	XY
Contributor Status	Single-Source
Created	12/09/2021 1:41:20 PM
User	admin@nist.gov
Kinship SNP Coverage Information	
Loci Typed	9654 / 9867

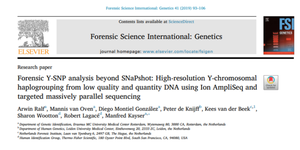
97.8 % Coverage

Locus	Genotype	Allele Name	Typed Allele?	Reads	QC Indicators
rs6690515	A,A	G	No	0	
rs6690515		A	Yes	635	
rs28635343	C,C	C	Yes	1910	
rs28635343		T	No	0	
rs16824588	T,C	T	Yes	147	
rs16824588		C	Yes	226	
rs424079	C,A	C	Yes	380	
rs424079		A	Yes	522	
rs2052204	G,A	G	Yes	1151	
rs2055204		A	Yes	1062	
rs4648384	C,A	C	Yes	949	
rs4648384		A	Yes	822	
rs16823228	G,G	A	No	0	
rs16823228		G	Yes	3158	
rs2817178	C,C	T	No	0	
rs2817178		C	Yes	138	
rs731031	T,T	T	Yes	450	
rs731031		C	No	0	
rs2455107	T,G	T	Yes	714	
rs2455107		G	Yes	983	
rs2483274	G,G	A	No	0	
rs2483274		G	Yes	1585	
rs868688	C,T	C	Yes	58	
rs868688		T	Yes	40	
rs2493272	C,C	T	No	0	
rs2493272		C	Yes	145	

X-SNPs | Y-SNPs | Ancestry SNPs | Phenotype SNPs | Identity SNPs | Kinship SNPs | Sample History | Settings

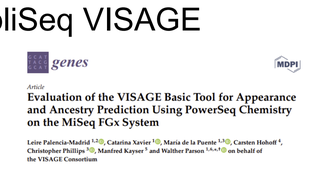
Ion AmpliSeq MH-74

Ion AmpliSeq Y-SNP



Ion AmpliSeq VISAGE

VISAGE — VISIBLE ATTRIBUTES THROUGH GENOMICS



11,267 SNPs added in 2022 Update

- The SRM 2391 series will continue to support the FBI-QAS for the validation and implementation of forensic marker systems.
- Certified allele calls are supported by sequence data and CE-length based measurements.

Poster available for download from STRBase:
https://strbase.nist.gov/pub_pres/ISFG2022_Steffen_SRM2391dUpdate.pdf