

2391c: PCR-Based DNA Profiling Standard

Carolyn R. (Becky) Hill, Margaret C. Kline, Michael D. Coble, Peter M. Vallone

Email: becky.hill@nist.gov

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

P-294

The NIST Standard Reference Material (SRM) 2391c: PCR-Based DNA Profiling Standard was designed for use in the standardization of forensic and paternity quality assurance procedures for fragment-based typing short tandem repeat (STR) alleles generated by the polymerase chain reaction (PCR). The SRM is intended for law enforcement laboratories, non-clinical research purposes, and for assigning values to in-house control materials. Certified genotypes of the 6 components A-F were assigned for 24 autosomal and 17 Y-STR markers plus Amelogenin using concordance testing between commercial kits. Selected Sanger sequencing characterization was performed for the alleles of 11 STR markers when only one PCR primer set was available for fragment-based typing. Our goal is to characterize the STR loci in components A-C by Sanger sequencing methods for the STR repeat regions and adjacent flanking regions. Core U.S. and European STR autosomal loci were characterized first and will be followed by Y-STRs present in commercial kits, and non-core loci. Additional characterization of the SRM is intended to support the emerging interest in next-generation sequencing technologies for forensic typing applications. Sanger methods have characterized underlying polymorphisms (sequence, insertion-deletion, variation in complex motifs) typically not detected by fragment-based typing. The sequenced regions include the commercial or known PCR binding sites commonly implemented in fragment-based typing.

SRM 2391c: PCR-Based DNA Profiling Standard Reference Material

Table 1. Description of Components in SRM 2391c

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

^{a)}DNA concentrations and cell counts are nominal values and are not intended for use as quantitative standards. https://www.nist.gov/srmors/view_cert.cfm?sr=2391c

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.
- Below 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.

Concordance with STR Kits		DNA Sequencing of Alleles	
Autosomal STR Loci	Y-STR Loci	Autosomal STR Loci	Y-STR Loci
D2S1338	DYS19	D1S1656	DYS448
D2S441	DYS385a	D8S1115	DYS456
D3S1358	DYS385b	D12S391	DYS458
D5S818	DYS389I	Penta D	DYS635
D7S820	DYS390	Penta E	DY-GATA-H4
D8S1179	DYS389II	SE33	
D10S1248	DYS391		
D13S317	DYS392		
D16S539	DYS393		
D18S51	DYS437		
D19S433	DYS438		
D21S11	DYS439		
D22S1045	*Amelogenin		
CSF1PO			
FGA			
TH01			
TPOX			
vWA			

The genotypes for 41 STR markers + Amelogenin are certified; 26% were previously Sanger Sequenced

>2 STR Kits were tested for concordance

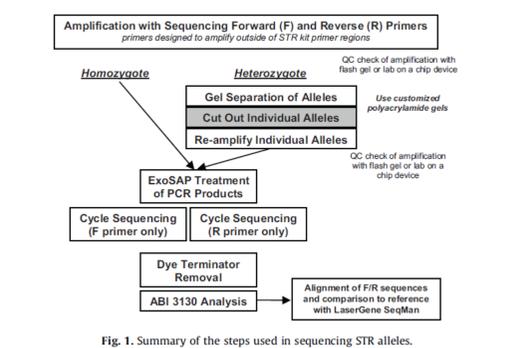
Methods for Sanger Sequencing

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

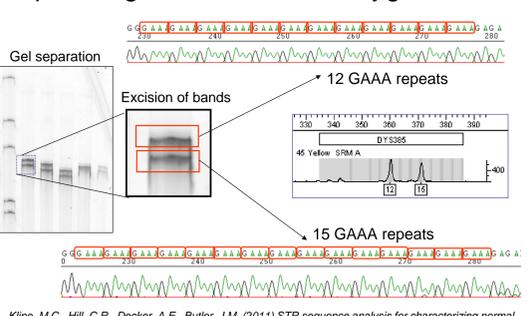


Short communication
STR sequence analysis for characterizing normal, variant, and null alleles
Margaret C. Kline¹, Carolyn R. Hill, Amy E. Decker¹, John M. Butler
National Institute of Standards and Technology, 100 Bureau Drive, MD 8312, Gaithersburg, MD 20899-1054

Sanger Sequencing Flow Chart



Sequencing Individual Heterozygous 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 here based on the May 2004 assembly of the human genome, build 35.

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	WVA	M25858	DYS439	AC002992	
D5S818	AC008512	D21S11	AP000433	DYS19	AC017019	DYS448	AC025227	
D7S820	AC004848	D22S1045	AL022314	DYS385	AC022486	DYS456	AC010166.2	
D8S1179	AF216671	CSF1PO	X14720	DYS389	AF140635	DYS458	AC010902	
D8S1115	AC090739	Penta E	FGA	M64982	DYS390	AC011289	DYS635	AC004772
D10S1248	AL391869	Penta D	AP001752	DYS391	AC011302	Y-GATA H4	AC011751	
D12S391	G08921		AC027004	DYS392	AC06152			

Sequences were aligned de novo using LaserGene SeqMan software and compared to SeqBuilder maps based on the listed GenBank reference sequences.

Sanger Sequencing Repeat Structure and Motifs

Sequence in gray is not counted towards the allele call

SRM 2391c - Component A					
Marker	Certified Genotype	Sanger Result	Repeat Structure - Allele 1	Repeat Structure - Allele 2	
D1S1656	17.3,17.3	17.3,17.3	[TAGA] ₁ TGA [TAGA] ₁₂ TAGG [TG] ₅	[TAGA] ₁ TGA [TAGA] ₁₂ TAGG [TG] ₅	
D2S1338	18.23	18.23	[TGCC] ₂ [TTCC] ₁₂	[TGCC] ₂ [TTCC] ₁₂ GTCC [TTCC] ₂	
D2S441	10.10	10.10	[TCTA] ₁₀	[TCTA] ₈ TCTG TCTA	
D3S1358	15.16	15.16	TCTA [TCTG] ₃ [TCTA] ₁₂	TCTA [TCTG] ₃ [TCTA] ₁₂	
D5S818	11.12	11.12	[AGAT] ₁₁	[AGAT] ₁₂	
D7S820	11.11	11.11	[GATA] ₁₁	[GATA] ₁₁	
D8S1179	13.14	13.14	[TCTA] ₁₃	[TCTA] ₂ TCTG [TCTA] ₁₁	
D8S1115	15.16	15.16	[ATT] ₁₅	[ATT] ₁₆	
D10S1248	15.16	15.16	[GGAA] ₁₅	[GGAA] ₁₆	
D12S391	18.3,22	18.3,22	AGAT GAT[AGAT] ₃ [AGAC] ₇ AGAT	[AGAT] ₁₃ [AGAC] ₈ AGAT	
D13S317	8.8	8.8	[TATC] ₈	[TATC] ₃	
D16S539	10.11	10.11	[GATA] ₁₀	[GATA] ₁₁	
D18S51	12.15	12.15	[AGAA] ₁₂	[AGAA] ₁₅	
D19S433	13.14	13.14	[AAGG] AAAG [AAGG] TAGG [AAGG] ₁₁	[AAGG] AAAG [AAGG] TAGG [AAGG] ₁₂	
D21S11	28,32.2	28,32.2	[TCTA] ₄ [TCTG] ₆ [TCTA] ₃ TA [TCTA] ₃ TCA	[TCTA] ₄ [TCTG] ₆ [TCTA] ₃ TA [TCTA] ₃ TCA	
D22S1045	15.15	15.15	[ATT] ₁₂ ACT [ATT] ₂	[ATT] ₁₂ ACT [ATT] ₂	
CSF1PO	10.10	10.10	[AGAT] ₁₀	[AGAT] ₁₀	
FGA	21.23	21.23	[TTTC] ₂ TTTT TTCT [CTTT] ₁₅ CTCC [TTCC] ₂	[TTTC] ₂ TTTT TTCT [CTTT] ₁₅ CTCC [TTCC] ₂	
Penta D	9.13	9.13	[AAAGA] ₉	[AAAGA] ₁₀	
Penta E	5.10	5.10	[AAAGA] ₅	[AAAGA] ₁₀	
SE33	16.18	16.18	[AAAG] ₂ AG [AAAG] ₃ AG [AAAG] ₁₆ G [AAAG] ₃	[AAAG] ₂ AG [AAAG] ₃ AG [AAAG] ₁₆ G	
TH01	8.9,3	8.9,3	[AATG] ₈	[AATG] ₈ ATG [AATG] ₃	
TPOX	8.8	8.8	[AATG] ₈	[AATG] ₈	
vWA	18.19	18.19	TCTA [TCTG] ₄ [TCTA] ₁₃	TCTA [TCTG] ₄ [TCTA] ₁₄	

SRM 2391c - Component B					
Marker	Certified Genotype	Sanger Result	Repeat Structure - Allele 1	Repeat Structure - Allele 2	
D1S1656	11.14	11.14	[TAGA] ₁₁ [TG] ₅	[TAGA] ₁₄ [TG] ₅	
D2S1338	17.17	17.17	[TGCC] ₁ [TTCC] ₁₁	[TGCC] ₁ [TTCC] ₁₁	
D2S441	10.14	10.14	[TCTA] ₁₀	[TCTA] ₁₁ TTTA [TCTA] ₂	
D3S1358	15.19	15.19	TCTA [TCTG] ₃ [TCTA] ₁₁	TCTA [TCTG] ₃ [TCTA] ₁₅	
D5S818	12.13	12.13	[AGAT] ₁₂	[AGAT] ₁₃	
D7S820	10.10	10.10	[GATA] ₁₀	[GATA] ₁₀	
D8S1179	10.13	10.13	[TCTA] ₁₀	[TCTA] ₁₃	
D8S1115	15.17	15.17	[ATT] ₁₅	[ATT] ₁₇	
D10S1248	13.13	13.13	[GGAA] ₁₃	[GGAA] ₁₃	
D12S391	19.24	19.24	[AGAT] ₁₂ [AGAC] ₃ AGAT	[AGAT] ₁₅ [AGAC] ₃ AGAT	
D13S317	9.12	9.12	[TATC] ₉	[TATC] ₁₂	
D16S539	10.13	10.13	[GATA] ₁₀	[GATA] ₁₃	
D18S51	13.16	13.16	[AGAA] ₁₃	[AGAA] ₁₆	
D19S433	16,16.2	16,16.2	[AAGG] AAAG [AAGG] TAGG [AAGG] ₁₄	[AAGG] AA [AAGG] TAGG [AAGG] ₁₅	
D21S11	32,32.2	32,32.2	[TCTA] ₄ [TCTG] ₆ [TCTA] ₃ TA [TCTA] ₃ TCA	[TCTA] ₄ [TCTG] ₆ [TCTA] ₃ TA [TCTA] ₃ TCA	
D22S1045	15.17	15.17	[ATT] ₁₄ ACT [ATT] ₂	[ATT] ₁₄ ACT [ATT] ₂	
CSF1PO	10.11	10.11	[AGAT] ₁₀	[AGAT] ₁₁	
FGA	20.23	20.23	[TTTC] ₂ TTTT TTCT [CTTT] ₁₅ CTCC [TTCC] ₂	[TTTC] ₂ TTTT TTCT [CTTT] ₁₅ CTCC [TTCC] ₂	
Penta D	8.12	8.12	[AAAGA] ₈	[AAAGA] ₁₂	
Penta E	7.15	7.15	[AAAGA] ₇	[AAAGA] ₁₅	
SE33	17.18	17.18	[AAAG] ₂ AG [AAAG] ₃ AG [AAAG] ₁₇ G [AAAG] ₃	[AAAG] ₂ AG [AAAG] ₃ AG [AAAG] ₁₈ G [AAAG] ₃	
TH01	6.9,3	6.9,3	[AATG] ₆	[AATG] ₆ ATG [AATG] ₃	
TPOX	8.11	8.11	[AATG] ₆	[AATG] ₆	
vWA	17.18	17.18	TCTA [TCTG] ₄ [TCTA] ₁₂	TCTA [TCTG] ₄ [TCTA] ₁₃	
DYS19	14	14	[TAGA] ₃ TAGG [TAGA] ₁₁	[TAGA] ₁₁	
DYS385a	13	13	[GAAA] ₁₃	[GAAA] ₁₃	
DYS385b	17	17	[GAAA] ₁₇	[GAAA] ₁₇	
DYS389I	13	13	[TCTG] ₃ [TCTA] ₁₀	[TCTG] ₃ [TCTA] ₁₀	
DYS389II	31	31	[TCTG] ₆ [TCTA] ₁₂ [TCTG] ₃ [TCTA] ₁₀	[TCTG] ₆ [TCTA] ₁₂ [TCTG] ₃ [TCTA] ₁₀	
DYS390	23	23	[TCTG] ₃ [TCTA] ₁₀ TCTG [TCTA] ₄	[TCTG] ₃ [TCTA] ₁₀ TCTG [TCTA] ₄	
DYS391	10	10	[TCTA] ₁₀	[TCTA] ₁₀	
DYS392	11	11	[TAT] ₁₁	[TAT] ₁₁	
DYS393	12	12	[AGAT] ₁₂	[AGAT] ₁₂	
DYS437	14	To Be Determined	To Be Determined	To Be Determined	
DYS438	10	To Be Determined	To Be Determined	To Be Determined	
DYS439	11	11	[AGAT] ₁₁	[AGAT] ₁₁	
DYS448	20	20	[AGAGAT] ₁₂ N ₁₂ [AGAGAT] ₈	[AGAGAT] ₁₂	
DYS456	15	15	[AGAT] ₁₅	[AGAT] ₁₅	
DYS458	17.2	17.2	[GAAA] ₁₅ AA [GAAA] ₂	[GAAA] ₁₅ AA [GAAA] ₂	
DYS635	20	20	[TCTA] ₄ [TGTA] ₂ [TCTA] ₂ [TGTA] ₂ [TCTA] ₁₀	[TCTA] ₄ [TGTA] ₂ [TCTA] ₂ [TGTA] ₂ [TCTA] ₁₀	
Y-GATA H4	11	11	[TAGA] ₁₁	[TAGA] ₁₁	

SRM 2391c - Component C					
Marker	Certified Genotype	Sanger Result	Repeat Structure - Allele 1	Repeat Structure - Allele 2	
D1S1656	11.15	11.15	[TAGA] ₁₁ [TG] ₅	[TAGA] ₁₁ TAGG [TG] ₅	
D2S1338	19.19	19.19	[TGCC] ₇ [TTCC] ₁₂	[TGCC] ₇ [TTCC] ₁₂	
D2S441	10.10	10.10	[TCTA] ₁₀ TCTG TCTA	[TCTA] ₈ TCTG TCTA	
D3S1358	16.18	16.18	TCTA [TCTG] ₃ [TCTA] ₁₂	TCTA [TCTG] ₃ [TCTA] ₁₄	
D5S818	10.11	10.11	[AGAT] ₁₀	[AGAT] ₁₁	
D7S820	10.12	10.12	[GATA] ₁₀	[GATA] ₁₂	
D8S1179	10.17	10.17	[TCTA] ₁₀	[TCTA] ₂ TCTG [TCTA] ₁₄	
D8S1115	9.9	9.9	[ATT] ₉	[ATT] ₉	
D10S1248	12.16	12.16	[GGAA] ₁₂	[GGAA] ₁₆	
D12S391	19.23	19.23	[AGAT] ₁₃ [AGAC] ₃ AGAT	[AGAT] ₁₂ [AGAC] ₁₀ AGAT	
D13S317	11.11	(12,12)*	[TATC] ₁₂ del [ATCA] 6 bp ds	[TATC] ₁₂ del [ATCA] 6 bp ds	
D16S539	10.10	10.10	[GATA] ₁₀	[GATA] ₁₀	
D18S51	16.19	16.19	[AGAA] ₁₆	[AGAA] ₁₉	
D19S433	13.2,15.2	13.2,15.2	[AAGG] AA [AAGG] TAGG [AAGG] ₁₂	[AAGG] AA [AAGG] TAGG [AAGG] ₁₄	
D21S11	29,30	29,30	[TCTA] ₄ [TCTG] ₆ [TCTA] ₃ TA [TCTA] ₃ TCA	[TCTA] ₄ [TCTG] ₆ [TCTA] ₃ TA [TCTA] ₃ TCA	
D22S1045	16.16	16.16	[ATT] ₁₃ ACT [ATT] ₂	[ATT] ₁₃ ACT [ATT] ₂	
CSF1PO	10.12	To Be Determined	To Be Determined	To Be Determined	
FGA	24.26	24.26	[TTTC] ₂ TTTT TTCT [CTTT] ₁₅ CTCC [TTCC] ₂	[TTTC] ₂ TTTT TTCT [CTTT] ₁	