Washington State Patrol
Forensic Laboratory Services

Washington State Patrol
Forensic DNA CE Day
June 18, 2015

NIST
National Institute of Standards and Technology

Expanding the Core CODIS Loci

Becky Hill and Michael Coble

NIST Applied Genetics Group
National Institute of Standards and Technology
Gaithersburg, Maryland

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.**

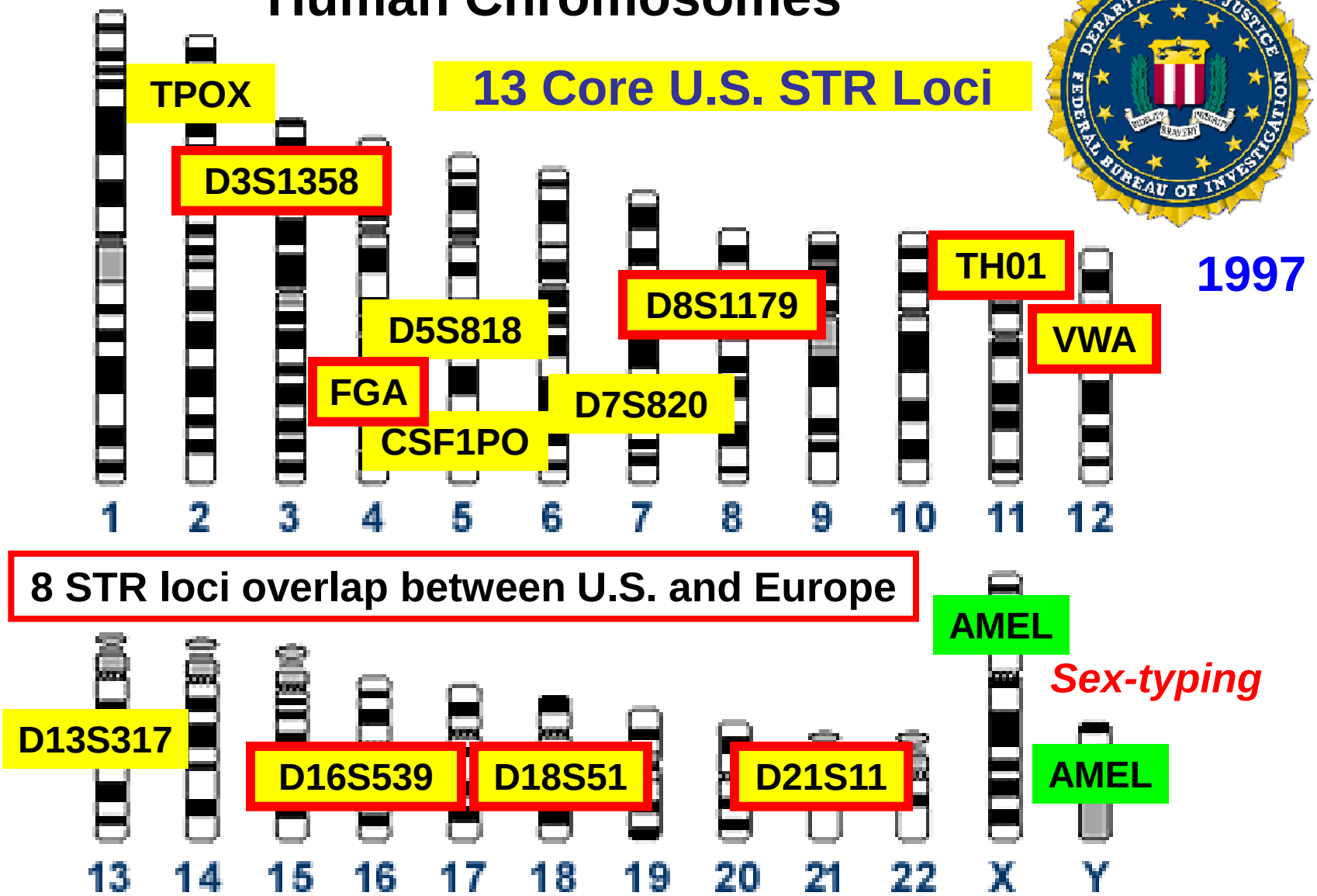
Presentation Outline

- 13 going on 20 – why do we need to increase the CODIS core loci?
- STR kits (including Fusion and GlobalFiler)
- Concordance study results
- New Y-STR kits

Additional STR Loci for the Future

- Will be needed for more complex kinship analyses and extended applications
  - Example: Y-STRs needed for familial searching
- Immigration testing needs more than 13 STRs
- Larger DNA databases will require more loci

Position of Forensic STR Markers on Human Chromosomes



Expanding the CODIS Core Loci

D.R. Hares (2012) Expanding the CODIS Core Loci in the United States. *Forensic Sci. Int. Genet.* 6: e52-e54
Addendum to expanding the CODIS core loci in the United States, Forensic Sci. Int. Genet. (2012) doi:10.1016/j.fsigen.2012.01.003

ELSEVIER

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Letter to the Editor

Expanding the CODIS core loci in the United States

CODIS Core Loci Working Group

Formed in May 2010 to make recommendations to FBI CODIS Unit

Douglas Hares (Chair) – FBI
John Butler – NIST
Cecelia Crouse – FL PBSO
Brad Jenkins – VA DFS
Ken Konzak – CA DOJ
Taylor Scott – IL SP

major reasons for expanding the CODIS core loci in the United States:

(1) To reduce the likelihood of adventitious matches [7] as the number of profiles stored at NDIS continues to increase each year (expected to total over 10 million profiles by the time of this publication). There are no signs that this trend will slow down as States expand the coverage of their DNA database programs and increase laboratory efficiency and capacity.

(2) To increase international compatibility to assist law enforcement data sharing efforts.

(3) To increase discrimination power to aid missing persons cases.

Possible scenarios for extending sets of genetic markers to be used in national DNA databases

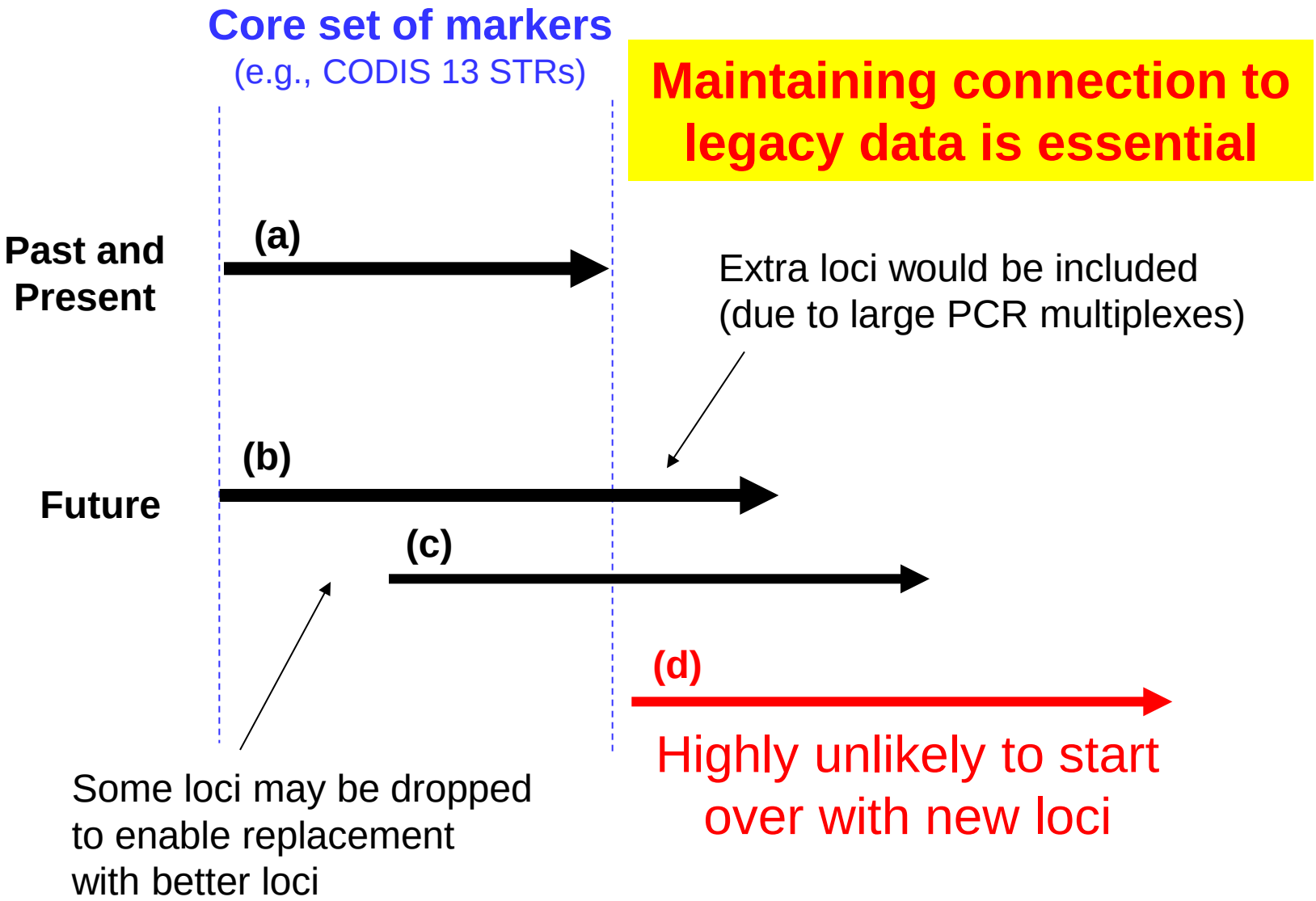


Figure 18.1 in forthcoming Butler (2009) *Fundamentals of Forensic DNA Typing*

Proposed Expanded CODIS Core Loci

D.R. Hares (2012) *Forensic Sci. Int. Genet.* 6(1):e52-e54

20 required loci
Amelogenin (for sex-typing)
18 autosomal STRs
1 Y-STR (DYS391)

	Locus
Section A (required)	Amelogenin
	D18S51
	FGA
	D21S11
	D8S1179
	vWA
	D13S317
	D16S539
	D7S820
	TH01
	D3S1358
	D5S818
	CSF1PO
	D2S1338
	D19S433
	D1S1656
	D12S391
	D2S441
	D10S1248
	~~Penta E~~
	DYS391

	Locus
Section B (in order of preference)	TPOX
	D22S1045
	SE33
	~~Penta D~~

Current CODIS 13 loci in red font

Final Expanded CODIS Core Loci

D.R. Hares (2015) *Forensic Sci. Int. Genet. In press*

Letter to the Editor

Selection and implementation of expanded CODIS core loci in the United States

Dear Editor,

This letter is to update the DNA community on the Federal Bureau

	Locus
Section A (required)	Amelogenin
	D18S51
	FGA
	D21S11
	D8S1179
	vWA
	D13S317
	D16S539
	D7S820
	TH01
	D3S1358
	D5S818
	CSF1PO
	D2S1338
	D19S433
	D1S1656
	D12S391
	D2S441
	D10S1248
	DYS391
Section B (in order of preference)	TPOX
	D22S1045
	SE33

Locus
CSF1PO
D3S1358
D5S818
D7S820
D8S1179
D13S317
D16S539
D18S51
D21S11
FGA
TH01
TPOX
vWA
D1S1656
D2S441
D2S1338
D10S1248
D12S391
D19S433
D22S1045

Final Core 20 aSTR markers

Criteria for Acceptance of Additional Loci

D.R. Hares (2012) *Forensic Sci. Int. Genet.* 6(1):e52-e54

Considered only short tandem repeat (STR) loci due to need for compatibility to existing database of >10 million STR profiles

- **STR Loci**
  - **No known association to medical conditions or defects**
  - **Low mutation rate**
  - **High level of independence**
  - **High level of discrimination**
  - **Use by international forensic DNA community**
  - **Number of loci vs. discrimination factor**
  - **Compliance with Quality Assurance Standards (QAS)**

Criteria for Acceptance of Additional Loci

D.R. Hares (2012) *Forensic Sci. Int. Genet.* 6(1):e52-e54

Considered only short tandem repeat (STR) loci due to need for compatibility to existing database of >10 million STR profiles

- **Kit performance**
  - **Balance between loci**
  - **Reliable**
  - **Reproducible**
  - **Sensitive**
  - **Quality results**
  - **Adaptable for use by NDIS laboratories (# of amplifications, ability of kit manufacturers to produce)**
  - **QAS compliant (documentation and availability of validation requirements)**

Three major reasons for expanding the CODIS core loci in the United States

D.R. Hares (2012) *Forensic Sci. Int. Genet.* 6(1):e52-e54

- **To reduce the likelihood of adventitious matches** as the number of profiles stored at NDIS continues to increase each year

- **To increase international compatibility** to assist law enforcement data sharing efforts

- **To increase discrimination power to aid missing persons cases**

Adventitious Matches

- The only published account of a false match from a DNA database came in 1999 when the UK database then consisting of 660,000 profiles with only 6 STR loci (SGM assay) lead to a “hit” between two individuals whose 6-locus random match probability was 1 in 37 million (R. Willing, *USA Today*, Feb 8, 2000; “Mismatch calls DNA test into question”).

- Further testing with four additional STRs (SGM Plus loci) showed that the samples were from different individuals. **The UK expanded the number of core loci from 6 to 10 with the adoption of the SGM Plus kit to try and prevent another adventitious match**.

- The growth of DNA databases necessitates the inclusion of additional loci to avoid this problem.

For further information, see D.N.A. Box 8.3 in Butler, J.M. (2012) *Advanced Topics in Forensic DNA Typing: Methodology*, p. 251

An example in the Israeli Database

Forensic Science International: Genetics 7 (2013) 494–498

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

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

Case Report

Can brothers share the same STR profile?

Naomi Zaken [a,*], Uzi Motro [b], Reouven Berdugo [a], Liron Elkayam Sapir [a], Ashira Zamir [a]

[a] DNA Database Laboratory, Division of Identification and Forensic Science (DIFS), Israel Police, National H.Q., Jerusalem, Israel
[b] Department of Ecology, Evolution and Behavior and Department of Statistics, The Hebrew University of Jerusalem, Jerusalem, Israel

Zaken *et al.* (2013) *FSI-G*

- PP ESI kit (16 loci that include the European Core)
- Found two profiles that shared 31 of 32 alleles.
- Y-STR testing – same haplotype
- Brothers, not identical twins.

Table 2
Alleles of the different loci and kits used. The differential alleles are in bold.

Kit	Site	Brother 1	Brother 2
PowerPlex® ESI System Promega	AMEL	X,Y	X,Y
	D3S1358	17,18	17,18
	D19S433	14,15.2	14,15.2
	D2S1338	16,20	16,20
	D22S1045	15,15	15,15
	D16S539	11,12	11,12
	D18S51	12,**13**	12,**15**
	D1S1656	11,15	11,15
	D10S1248	14,15	14,15
	D2S441	11,14	11,14
	THO1	7,8	7,8
	vWA	15,16	15,16
	D21S11	30.2,32.2	30.2,32.2
	D12S391	18,23	18,23
	D8S1179	13,15	13,15
	FGA	20,24	20,24

Zaken *et al.* (2013) *FSI-G*

- In order to increase the discrimination power between related and unrelated individuals, we recommend that the DNA laboratories consider using a larger multiplex typing kit in cases like the one informed here.

International Comparability

Currently there are 24 autosomal STR markers present in commercial kits

ESS = European Standard Set

13 CODIS loci

7 ESS loci

U.S.	Europe
TPOX	
CSF1PO	
D5S818	
D7S820	
D13S317	
FGA	FGA
vWA	vWA
D3S1358	D3S1358
D8S1179	D8S1179
D18S51	D18S51
D21S11	D21S11
TH01	TH01
D16S539	D16S539
D2S1338	D2S1338
D19S433	D19S433
Penta D	
Penta E	
	D12S391
	D1S1656
	D2S441
	D10S1248
	D22S1045
	SE33

5 loci adopted in 2009 to expand to 12 ESS loci

3 miniSTR loci developed at NIST

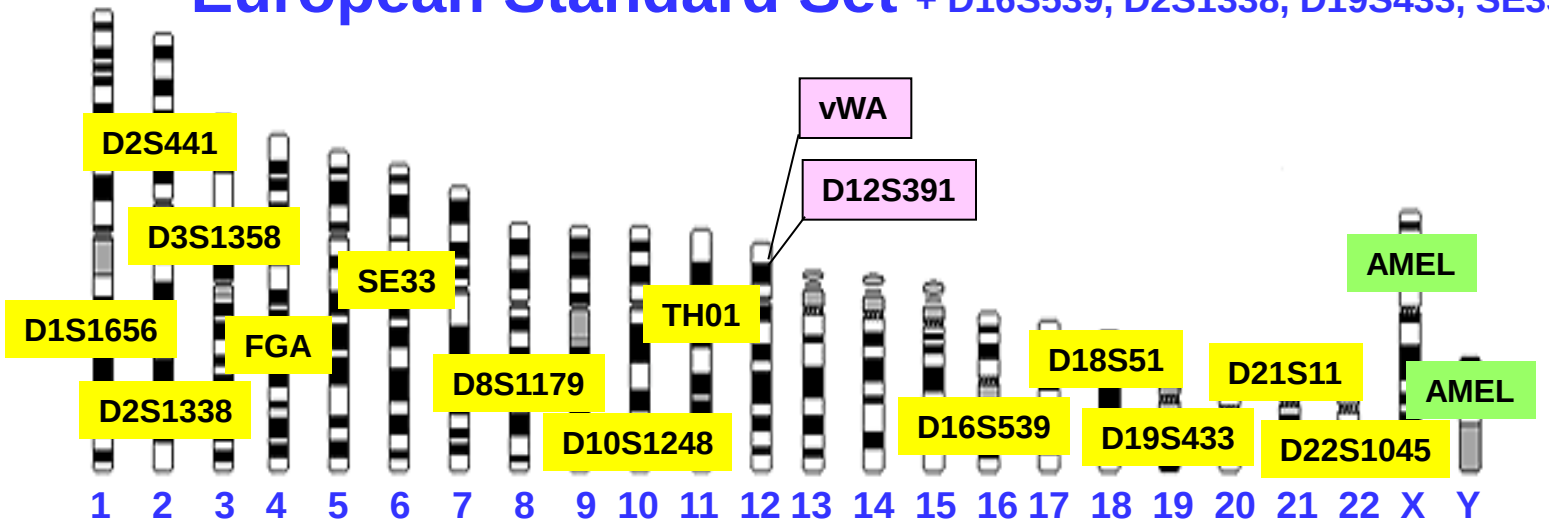
Locus used in China ← **D6S1043**

SE33 → Core locus for Germany

Common Forensic STR Loci

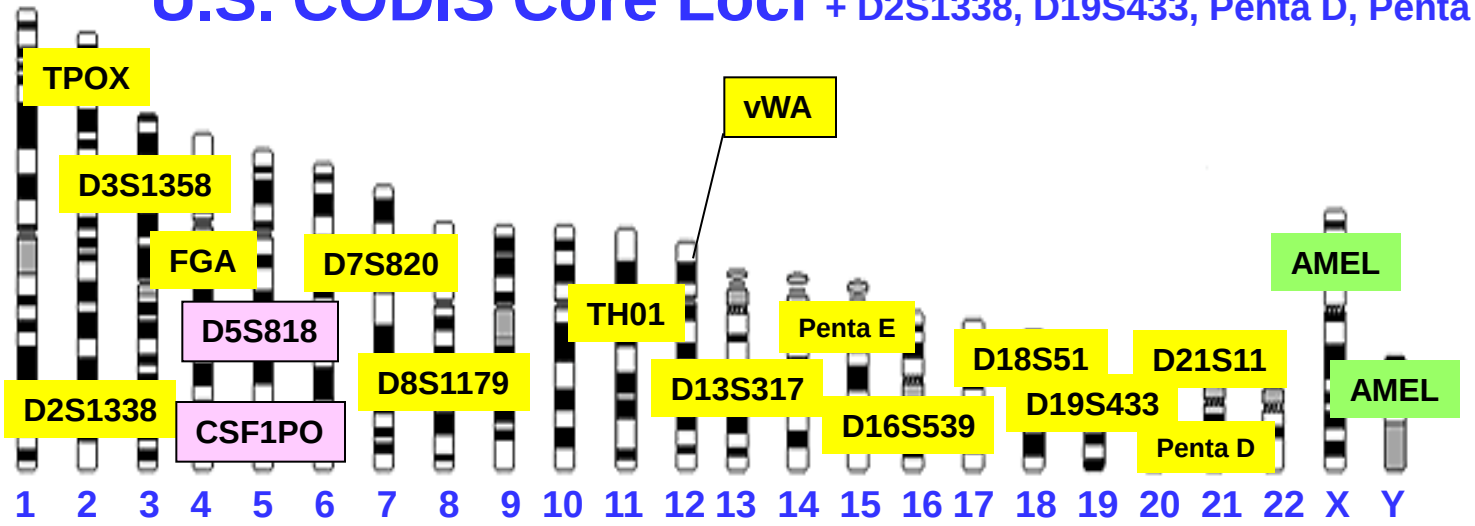
European Standard Set + D16S539, D2S1338, D19S433, SE33

Europe



U.S. CODIS Core Loci + D2S1338, D19S433, Penta D, Penta E

United States



What can we learn from European Standard Set (ESS) expansion experience?

PROFILES IN DNA

March 2009 issue

THE ESS LOCI

Expansion of the European Standard Set of DNA Database Loci—the Current Situation

Peter M. Schneider
Institute of Legal Medicine, University Hospital of Cologne, Germany

http://www.promega.com/resources/articles/profiles-in-dna/2009/expansion-of-the-european-standard-set/

ELSEVIER

Forensic Science International 119 (2001) 225–231

Forensic Science International

www.elsevier.com/locate/forsciint

A brief history of the formation of DNA databases in forensic science within Europe

Peter D. Martin[a,*], Hermann Schmitter[b], Peter M. Schneider[c]

[a]32 Oakfield Gardens, Kent, BR3 3AZ Beckenham, UK
[b]Serology Department, Bundeskriminalamt, Thaerstr. 11, 65193 Wiesbaden, Germany
[c]Institute of Legal Medicine, Johannes Gutenberg University, Am Pulverturm 3, 55131 Mainz, Germany

Received 10 September 2000; accepted 14 November 2000

Interpol European Working Party on DNA Profiling

Schmitter[3], Uwe Schleenbecker[3], Janos Woller[4], Paola Montagna[5],
en Wolf[7], Nils Moe[8], Jan Matusek[9], José Andradas Heranz[10],
Serra Utrilla[10], William Grahamslaw[11], Peter Fifka[11], Mark Branchflower[12]
uxelles, Belgium

Forensic Science International 119 (2001) 232–238

Forensic Science International

www.elsevier.com/locate/forsciint

Criminal DNA databases: the European situation

Peter M. Schneider[a,*], Peter D. Martin[b]

[a]Institute of Legal Medicine, Johannes Gutenberg University, Am Pulverturm 3, 55131 Mainz, Germany
[b]Former Deputy Director, Metropolitan Police Forensic Science Laboratory, London, UK

Received 10 September 2000; accepted 14 November 2000

DNA Databases for Offender Id
The Need for Technical, Legal an

Peter M. Schneider
Institute of Legal Medicine, Johannes Gutenberg University,
Am Pulverturm 3, 55131 Mainz, Germany

http://www.promega.com/products/pm/genetic-identity/ishi-conference-proceedings/2nd-eshi-oral-presentations/

Lessons from European ESS Expansion

- **Data studies should drive decisions**
  - Interlaboratory study with degraded DNA (Dixon et al. 2006 article was key)
- Casework capabilities are a primary goal
  - miniSTRs and desire for kits with ability to overcome inhibitors

EDNAP Study Showed Value of miniSTRs

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Forensic Science International 164 (2006) 33–44

Forensic Science International

www.elsevier.com/locate/forsciint

Analysis of artificially degraded DNA using STRs and SNPs—results of a collaborative European (EDNAP) exercise

L.A. Dixon [a,*], A.E. Dobbins [a], H.K. Pulker [a], J.M. Butler [b], P.M. Vallone [b], M.D. Coble [b], W. Parson [c], B. Berger [c], P. Grubwieser [c], H.S. Mogensen [d], N. Morling [d], K. Nielsen [d], J.J. Sanchez [d], E. Petkovski [e], A. Carracedo [f], P. Sanchez-Diz [f], E. Ramos-Luis [f], M. Briōn [f], J.A. Irwin [g], R.S. Just [g], O. Loreille [g], T.J. Parsons [g], D. Syndercombe-Court [h], H. Schmitter [i], B. Stradmann-Bellinghausen [j], K. Bender [j], P. Gill [a]

[a] *The Forensic Science Service, Research and Development, Trident Court, Birmingham, UK*
[b] *National Institute of Standards and Technology, Gaithersburg, MD, USA*
[c] *Institute of Legal Medicine, Innsbruck Medical University, Austria*

"Recently, there has been much debate about what kinds of genetic markers should be implemented as new core loci that constitute national DNA databases. The choices lie between conventional STRs, ranging in size from 100 to 450 bp; mini-STRs, with amplicon sizes less than 200 bp; and single nucleotide polymorphisms (SNPs)…Results were collated and analysed and, in general, **mini-STR systems were shown to be the most effective**…"

Lessons from European ESS Expansion

- **Data studies should drive decisions**
  - Interlaboratory study with degraded DNA (Dixon et al. 2006 article was key)
- Casework capabilities are a primary goal
  - miniSTRs and desire for kits with ability to overcome inhibitors
- **Initial locus selection announced through Letters to the Editor** of the leading forensic DNA journal (Gill et al. 2006a, 2006b)

Data Driven Decisions

ELSEVIER

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Forensic Science International 156 (2006) 242–244

Forensic Science International

www.elsevier.com/locate/forsciint

Short communication

The evolution of DNA databases—Recommendations for new European STR loci

Peter Gill [a,*], Lyn Fereday [b], Niels Morling [c], Peter M. Schneider [d]

[a] *Forensic Science Service, Birmingham, UK*
[b] *Forensic Science Service, London, UK*
[c] *Department of Forensic Genetics, Institute of Forensic Medicine, University of Copenhagen, Denmark*
[d] *Institute of Legal Medicine, University of Cologne, Germany*

Received 25 May 2005; accepted 26 May 2005
Available online 5 July 2005

"Following a recent meeting by the ENFSI and EDNAP groups on the 4–5 April, 2005, in Glasgow, UK, it was unanimously agreed that the process of standardization within Europe **should take account of recent work that unequivocally demonstrated that chance of obtaining a result from a degraded sample was increased when small amplicons (mini-STRs) were analysed**…"

Lessons from European ESS Expansion

- **Data studies should drive decisions**
  - Interlaboratory study with degraded DNA (Dixon et al. 2006 article was key)
- Casework capabilities are a primary goal
  - miniSTRs and desire for kits with ability to overcome inhibitors
- **Initial locus selection announced through Letters to the Editor** of the leading forensic DNA journal (Gill et al. 2006a, 2006b)
- Companies responded with prototype kits for evaluation
- Expanded ESS loci were selected and voted upon after data review by ENFSI labs **(4 years after initial recommendations were made)**
- EU adopted recommendations of ENFSI
- Commercial kits became available to meet expanded ESS requirements
- Population data gathered and software developed
- European labs must be compliant by Nov 30, 2011 (**2 years after adoption**)
- Casework capabilities not familial searching potential were the intent of the core loci selection

Characterizing New STR Loci

Main Points:

- In April 2011, the FBI announced plans to expand the core loci for the U.S. beyond the current 13 CODIS STRs

- NIST has collected and characterized U.S. population data on new loci to aid understanding of various marker combinations

- We are collecting all available information from the literature on the 24 commonly used autosomal STR loci

Publications:

- Hares (2012) Expanding the U.S. core loci… *FSI Genetics* 6(1): e52-e54
- Butler & Hill (2012) *Forensic Sci Rev* 24(1): 15-26
- Butler et al. (2012) Profiles in DNA – introduces NIST 1036 data set
- Hill et al. (2013) FSI Genetics – 29 autosomal STRs in PowerPlex CS7 and other kits

Article in the January 2012 issue of *Forensic Science Review*

Available at **http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm**

Biology and Genetics of New Autosomal STR Loci Useful for Forensic DNA Analysis

REFERENCE: Butler JM, Hill CR: Biology and genetics of new autosomal STR loci useful for forensic DNA analysis; *Forensic Sci Rev* 24:15; 2012.

ABSTRACT: Short tandem repeats (STRs) are regions of tandemly repeated DNA segments found throughout the human genome that vary in length (through insertion, deletion, or mutation) with a core repeated DNA sequence. Forensic laboratories commonly use tetranucleotide repeats, containing a four base pair (4-bp) repeat structure such as GATA. In 1997, the Federal Bureau of Investigation (FBI) Laboratory selected 13 STR loci that form the backbone of the U.S. national DNA database. Building on the European expansion in 2009, the FBI announced plans in April 2011 to expand the U.S. core loci to as many as 20 STRs to enable more global DNA data sharing. Commercial STR kits enable consistency in marker use and allele nomenclature between laboratories and help improve quality control. The STRBase website, maintained by the U.S. National Institute of Standards and Technology (NIST), contains helpful information on STR markers used in human identity testing.

Key Words: Autosomal genetic markers, CODIS STRs, core loci, DNA typing, European Standard Set, expanded U.S. core loci, short tandem repeat (STR), STR kits.

Discusses the 24 autosomal STR loci available in commercial kits

FBI Consortium Validation Project

- Purpose of the CVP
  - To have several crime labs across the U.S. validate these two new kits
  - To determine if these new loci are appropriate to add to the CODIS core set
  - To implement the new loci into the U.S. National DNA Database

Implementation Plan

Plan Published in February 2012

http://www.fbi.gov/about-us/lab/codis/planned-process-and-timeline-for-implementation-of-additional-codis-core-loci

Determination of Additional CODIS Core Loci

D.R. Hares (2012) Expanding the CODIS Core Loci in the United States. *Forensic Sci. Int. Genet.* 6: e52-e54
Addendum to expanding the CODIS core loci in the United States, Forensic Sci. Int. Genet. (2012) doi:10.1016/j.fsigen.2012.01.003

What	Why	Who/How	When
Form a Working Group (WG) to discuss initial selection	Establishes target goals	CODIS Core Loci Working Group with FBI Chair and 5 members; Web meetings	May 2010 - present
Announce proposed additional CODIS core loci	Sets desired target goals and informs manufacturers	WG Chair; Publish proposed listing of CODIS core loci	April 2011 online (published Jan 2012)
Ongoing Progress Reports	Provides updates for DNA community	WG Chair; Present updates on status of CODIS Core Loci project at meetings	2010-2012
Implementation Considerations & Strategy	Identify issues for implementation and timeline	WG	June 2011 - present
Manufacturers develop prototype kits	Creates tools to meet target goals	Manufacturers; Provide status reports to WG for timeline	2011-2012
Test and validate prototype kits	Examines if target goals can be met	Validation Laboratories; Follow QAS compliant validation plan	Beginning in 2012
Review and evaluate data from validation	Evaluates if desired performance is obtained	NIST, SWGDAM and FBI; Provide feedback, if any, to Manufacturers	In conjunction with and at the conclusion of validation
Selection of new CODIS core loci	Allows protocols to be established	FBI; seek input from DNA community and stakeholders; Notify Congress	After evaluation of validation data and kit production factors
Implementation of new CODIS core loci at the National DNA Index System	Enables target goals to be met	All NDIS-participating labs	~ 24 months after selection of new CODIS core loci

http://www.fbi.gov/about-us/lab/codis/planned-process-and-timeline-for-implementation-of-additional-codis-core-loci

Implementation Plan

Plan Published in February 2012

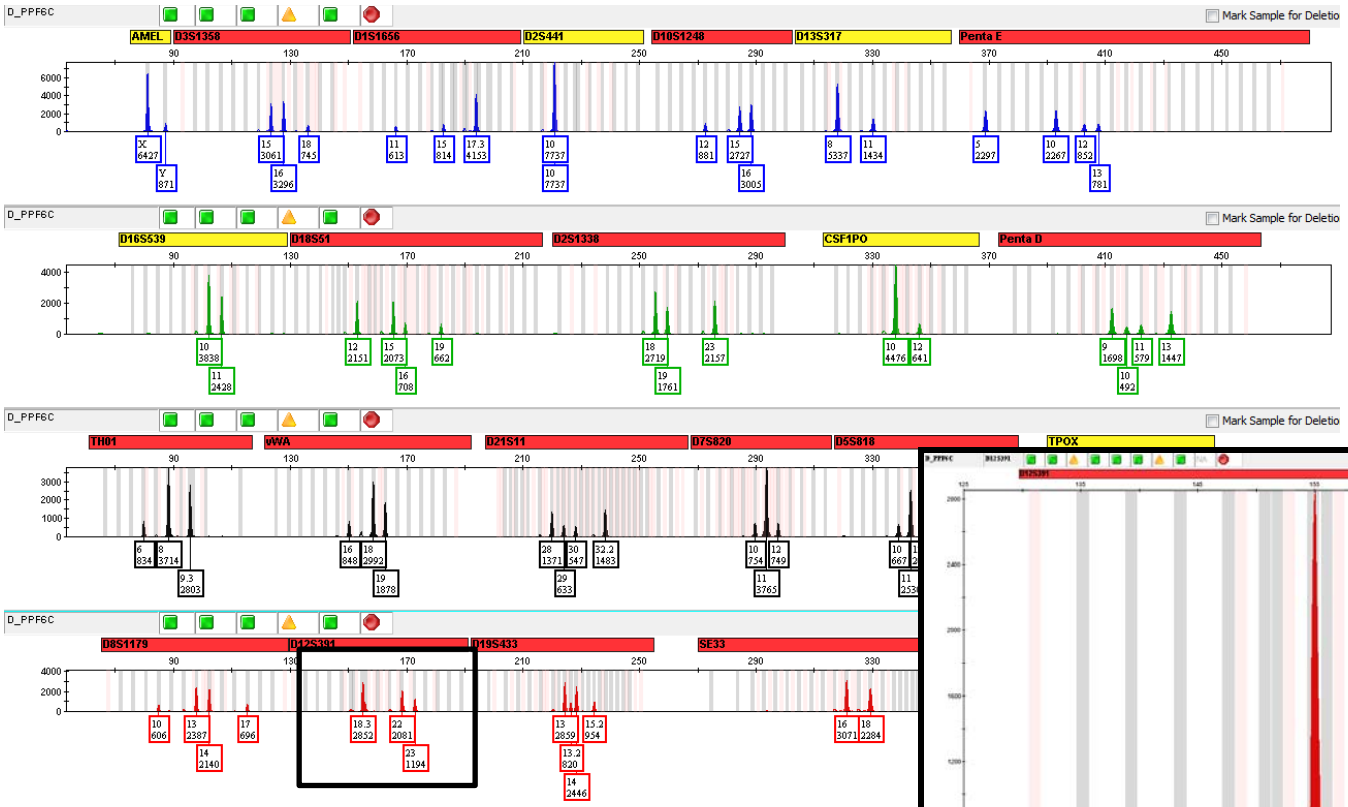
http://www.fbi.gov/about-us/lab/codis/planned-process-and-timeline-for-implementation-of-additional-codis-core-loci

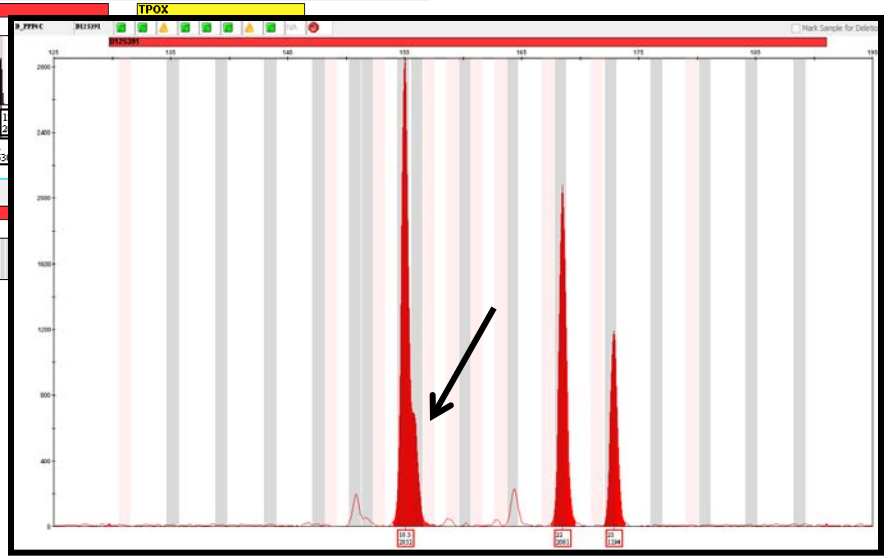
- Selection of laboratories to participate in validation studies
- **Validation of proposed new CODIS core loci**
- Selection of CODIS core loci
  - Announced Nov 2014 at the CODIS meeting
- Implementation of new CODIS core loci into NDIS operations
  - January 2017

Validation Plan

- CODIS Core Loci Group and NIST are currently working with U.S. laboratories participating in the consortium validation effort
  - Manufacturer grade kits were provided for validation effort
  - Specific plans distributed to Database, Casework and Missing Person Laboratories
  - 11 laboratories involved in the consortium effort
  - Manuscripts in preparation

Component D Resolution Issue at 12S391 PowerPlex Fusion

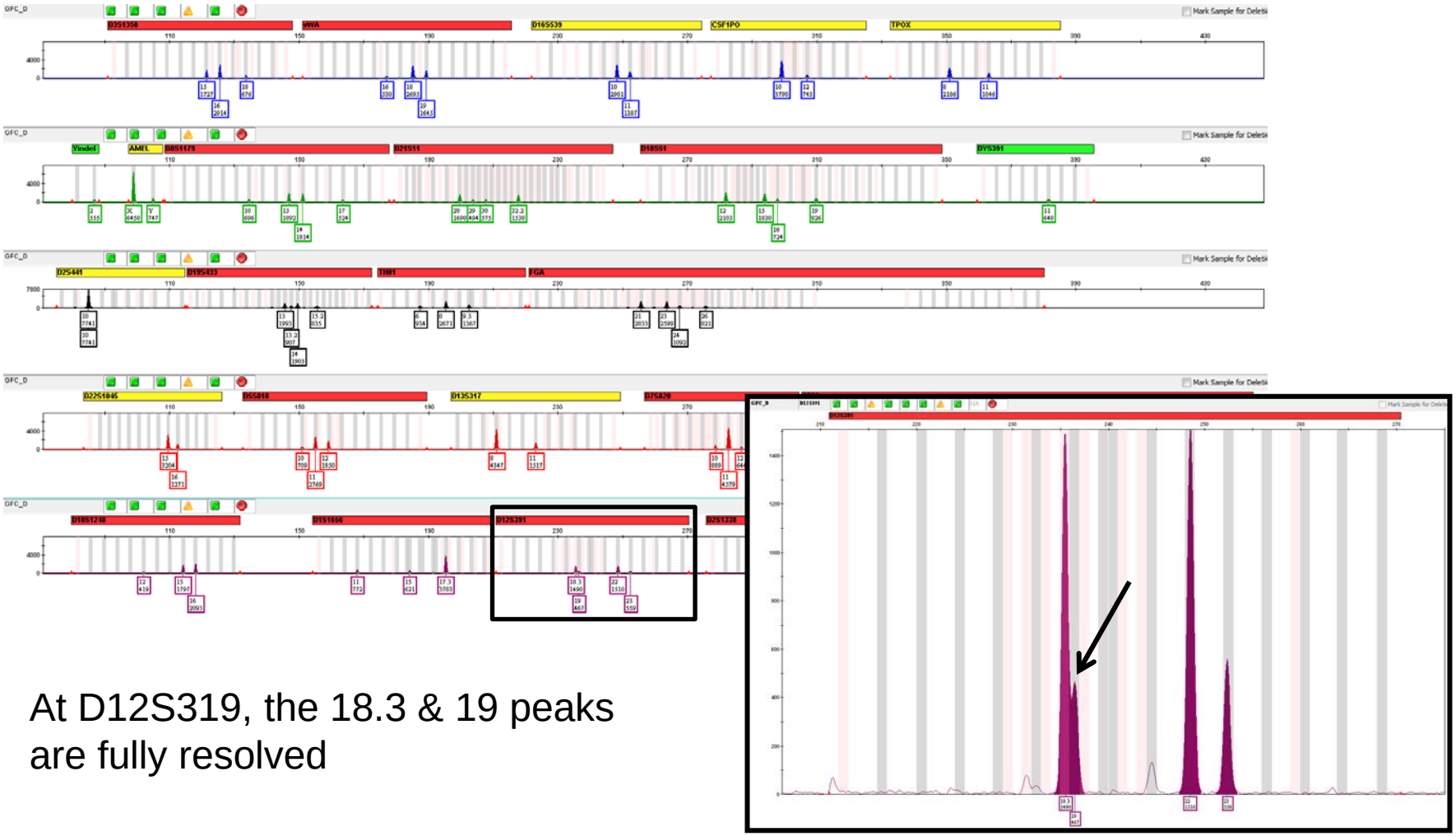




At D12S319, the 18.3 & 19 peaks are not fully resolved so the 19 peak is not called

1 ng DNA, 30 cycles

GlobalFiler Component D at D12S391



At D12S319, the 18.3 & 19 peaks are fully resolved

1 ng DNA, 29 cycles

Additional Loci

More Loci are Useful
in Situations Involving Relatives

- **Missing Persons** and Disaster Victim Identification (kinship analysis)
- Immigration Testing (often limited references)
  - Recommendations for 25 STR loci
- Deficient Parentage Testing
  - often needed if only one parent and child are tested

Relationship testing labs are being pushed to answer more difficult genetic questions…and **we want to make sure the right tools are in place**

In February 25, 2011 issue of *Forensic Science International*…

Forensic Science International 205 (2011) 64–68

Contents lists available at ScienceDirect

Forensic Science International

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

Examples of kinship analysis where Profiler Plus™ was not discriminatory enough for the identification of victims using DNA identification☆

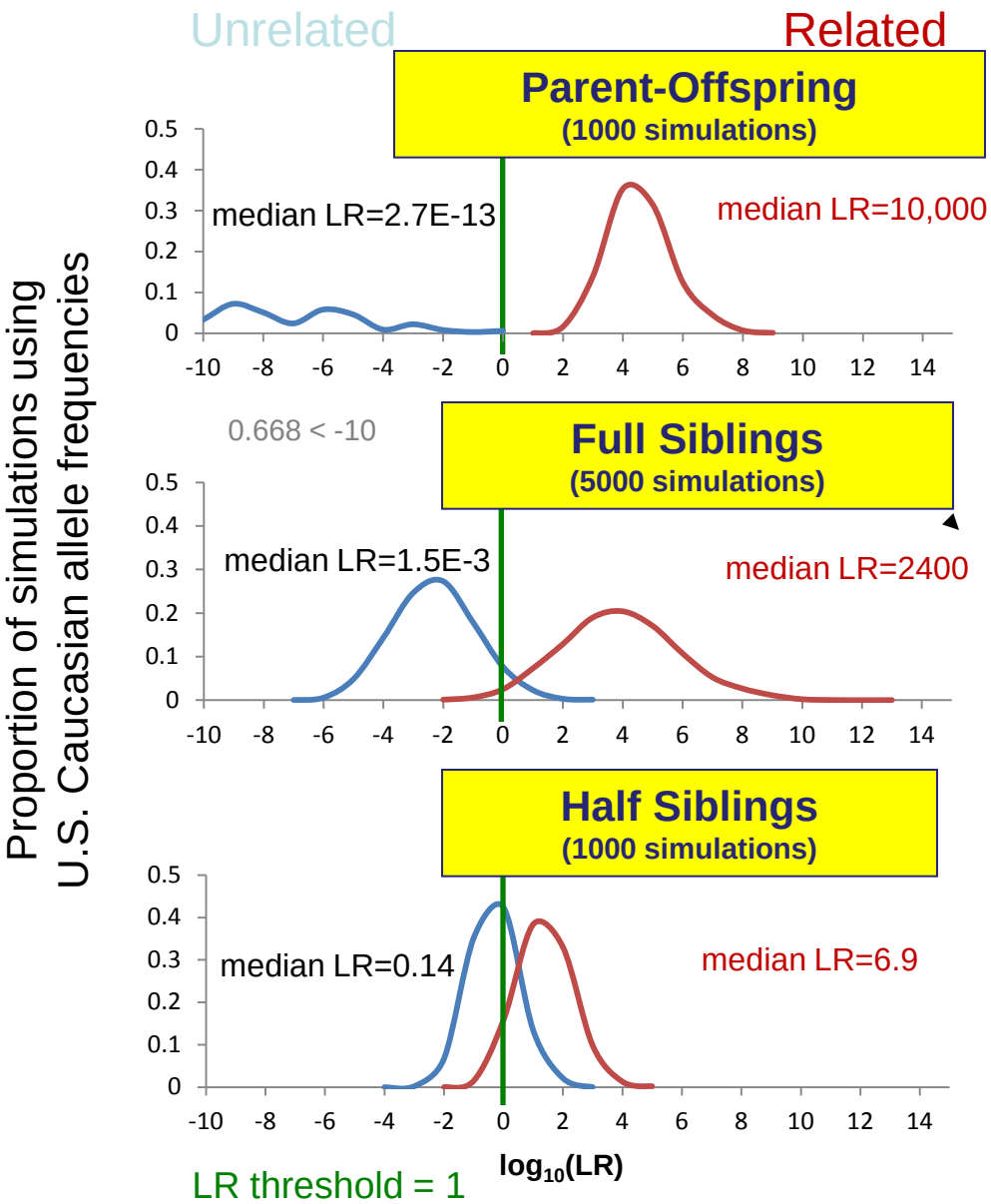
D. Hartman[a,b,*], L. Benton[a], L. Morenos[a], J. Beyer[a], M. Spiden[a], A. Stock[a]

[a] Victorian Institute of Forensic Medicine, 57-83 Kavanagh St, Southbank, VIC 3006, Australia
[b] Department of Forensic Medicine, Monash University, 57-83 Kavanagh St, Southbank, VIC 3006, Australia

Disaster victim identification from the 2009 Victorian bushfires relied on DNA (82% involved kinship associations rather than direct matching)

They advocate additional autosomal STR loci to aid kinship associations

How do 13 loci perform for kinship analysis?



The degree of overlap corresponds with possible values for false positive or false negative results.

<u>Parent-offspring comparisons:</u>

No overlap between unrelated and related LR distributions

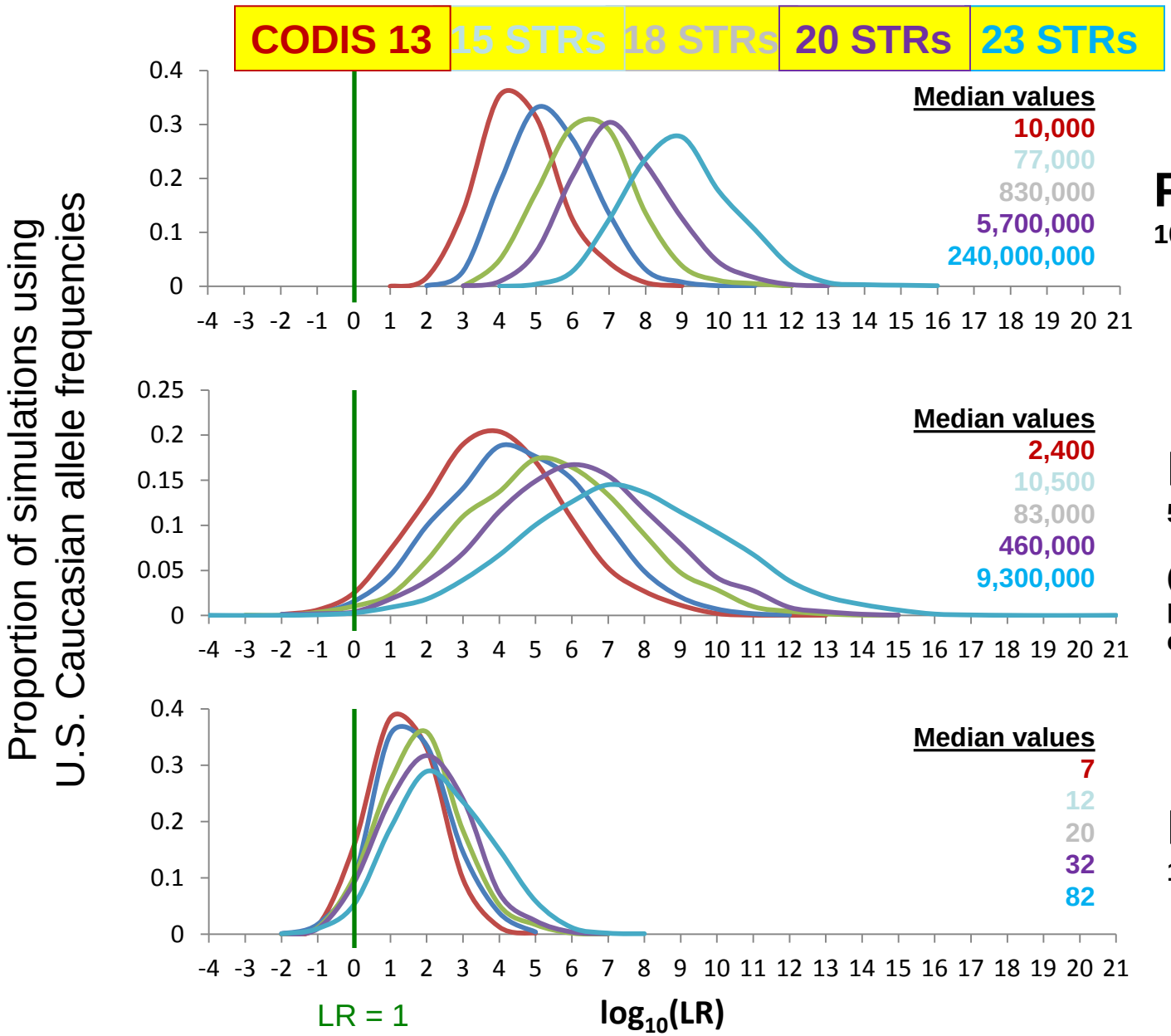
<u>Full sibling comparisons:</u>

False positive rate = 0.027
False negative rate = 0.033

<u>Half sibling comparisons:</u>

False positive rate = 0.155
False negative rate = 0.168

Do additional loci improve kinship determination?



Parent-offspring
1000 simulations

Full siblings
5000 simulations

(additional simulations performed for smoother curves)

Half siblings
1000 simulations

		Appearing as				
Configuration	N contributor mixture	1	2	3	4	5
CODIS 13	6	1.75E-40	6.34E-09	1.61E-01	0.9457	0.9999
GlobalFiler		0	9.59E-21	5.32E-05	0.1881	0.8599
Proposed CODIS without SE33		1.61E-66	1.67E-16	3.81E-03	0.6798	0.9969
Fusion		8.32E-76	2.07E-19	8.62E-04	0.5759	0.9941
CODIS 13	5	9.78E-33	2.10E-06	4.14E-01	0.9897	
GlobalFiler		6.36E-61	7.01E-15	4.84E-03	0.6101	
Proposed CODIS without SE33		5.48E-54	9.04E-12	5.92E-02	0.9229	
Fusion		3.75E-61	7.80E-14	2.70E-02	0.8886	
CODIS 13	4	7.02E-25	5.15E-04	7.85E-01		
GlobalFiler		3.50E-46	3.49E-09	1.65E-01		
Proposed CODIS without SE33		6.28E-41	3.14E-07	4.32E-01		
Fusion		2.01E-46	1.76E-08	3.37E-01		
CODIS 13	3	8.42E-17	0.0595			
GlobalFiler		5.77E-31	0.0004			
Proposed CODIS without SE33		1.65E-27	0.0031			
Fusion		3.53E-31	0.0010			
CODIS 13	2	1.70E-08				
GlobalFiler		2.05E-15				
Proposed CODIS without SE33		1.03E-13				
Fusion		1.61E-15				

Contributors
Coble, Bright, Buckleton,Curran

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

29 STR Loci present in STR kits

rank ordered by their variability

Better for mixtures (more alleles seen)

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

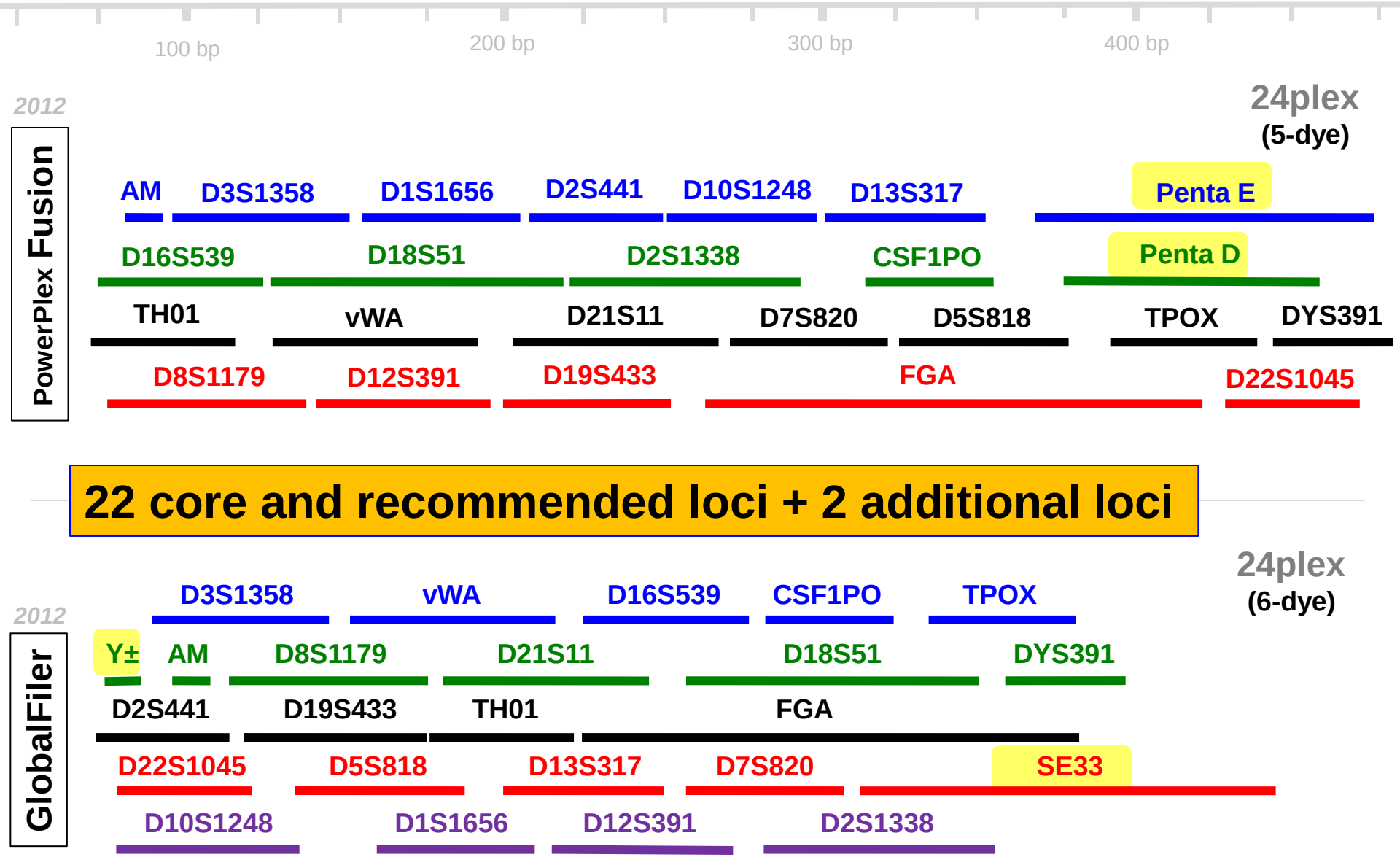
U.S. population data for 29 autosomal STR loci

run and population statistics were confirmed using the PowerMarker v3.25 statistics program [10].

- Added to CODIS PopStats software in Sept 2013

STR Kits

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™
- Resources
- Human Identification Home
- GlobalFiler™ Express Kit
- GlobalFiler™ Kit

Go faster. Go further. Go global.

GlobalFiler®

Powered by 6-Dye™

Introducing the world's most powerful STR kits

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

NDIS Approved

Both the GlobalFiler® and GlobalFiler® Express kits are approved for use by laboratories generating DNA profiles for inclusion in the US National DNA Index System (NDIS) CODIS database.

Order now

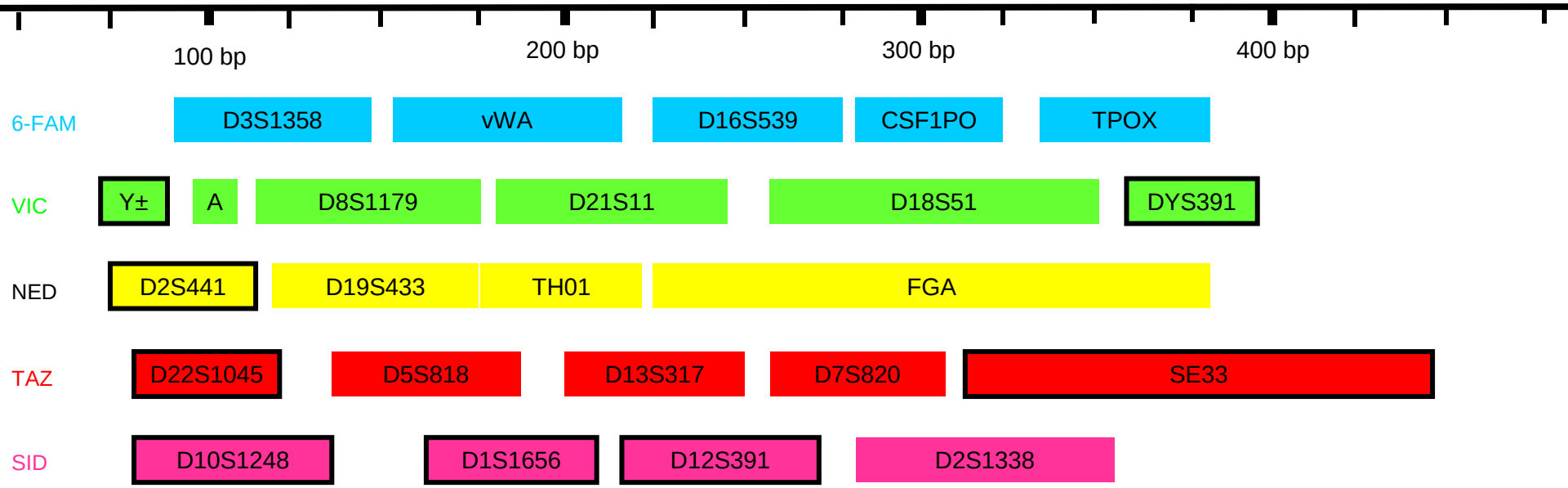
Request a demo

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

http://www.lifetechnologies.com/us/en/home/industrial/human-identification/globalfiler-str-kit.html

Life Technologies **GlobalFiler**

24plex



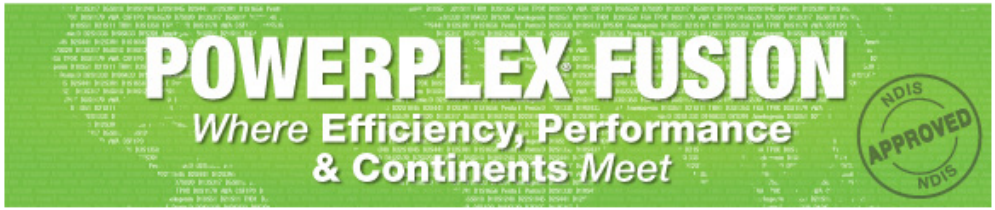
- 24 STR loci in 6 dyes (3500 use or 3130 upgrade required)
  - Includes SE33 and a Y-indel
- GlobalFiler Express: direct amplification capabilities
  - Single source samples: 40 min amplification
- GlobalFiler Casework
  - Casework samples: 80 min amplification
- GlobalFiler gives ~12 orders of magnitude improvement using the NIST 1036 data set

Two separate kits

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

PowerPlex Fusion

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.

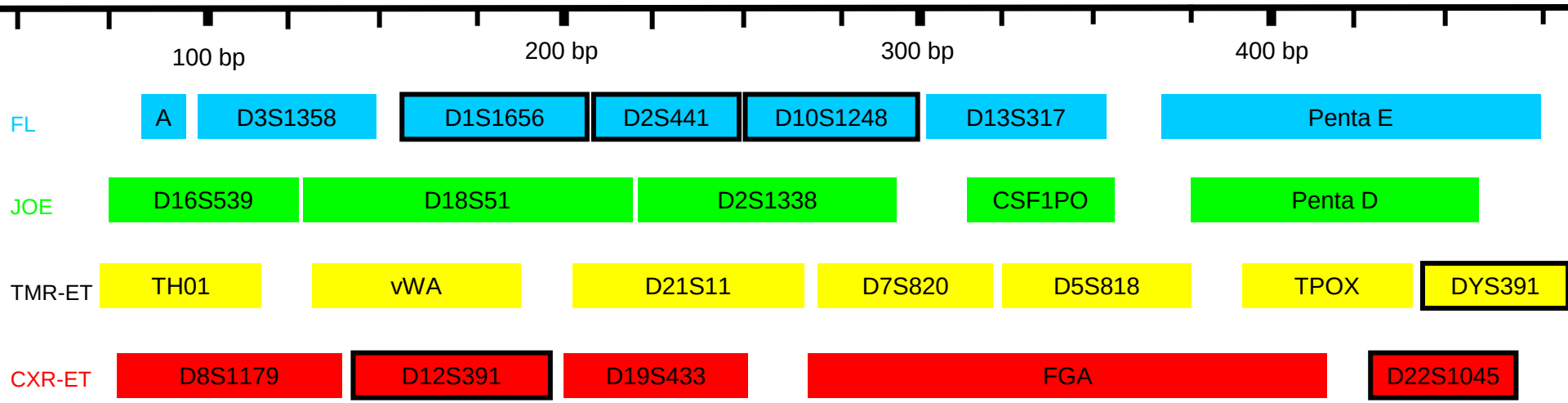
The PowerPlex® Fusion System provides all of the materials needed for co-amplification and five-color fluorescent detection of 24 loci (23 STR loci and Amelogenin), including the CODIS core loci and the European Standard Set (ESS) loci. With 24 loci, the system offers the most STR loci and highest discrimination 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.

Figure 1. Configuration of the PowerPlex® Fusion System. The PowerPlex® Fusion System allows co-amplification and four-color detection of 24 loci, including all CODIS and ESS loci: D3S1358, D1S1656, D2S441, D10S1248, D13S317, D16S539, D18S51, D2S1338, CSF1PO, TH01, vWA, D21S11, D7S820, D5S818, TPOX, D8S1179, D12S391, D19S433, FGA, D22S1045, plus Penta E, Penta D, DYS391 and Amelogenin.

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

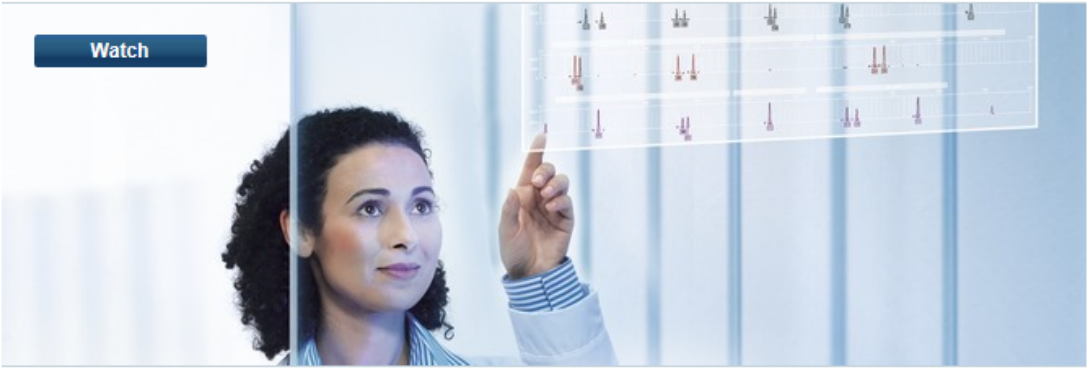
http://www.promega.com/products/pm/genetic-identity/powerplex-fusion/

Qiagen Investigator 24plex

Available worldwide from Oct 2014 and not before June 2015 in the U.S.

There's More to a Sample than a Profile — Global STR Analysis Including Quality Control

Print Bookmark Share



The FBI CODIS Core Loci Working Group have published the recommendation that the CODIS core loci should be expanded, and a combination of STR markers from the Combined DNA Index System (CODIS), the European Network of Forensic Science Institutes (ENFSI), and the European DNA Profiling Group (EDNAP) should be used to improve the accuaracy of forensic testing.

QIAGEN has developed new kits — Investigator 24plex Kits — that coamplify all 23 recommended markers. The kits use novel 6-dye technology to keep the amplicon length of markers short while avoiding overlapping of markers. Kits are available for purified DNA from casework and for reference samples.

This webinar will outline the benefits of the Investigator 24plex Kits and highlight the improved workflow, which saves unnecessary re-runs of a sample by use of an internal "Quality Sensor" control.

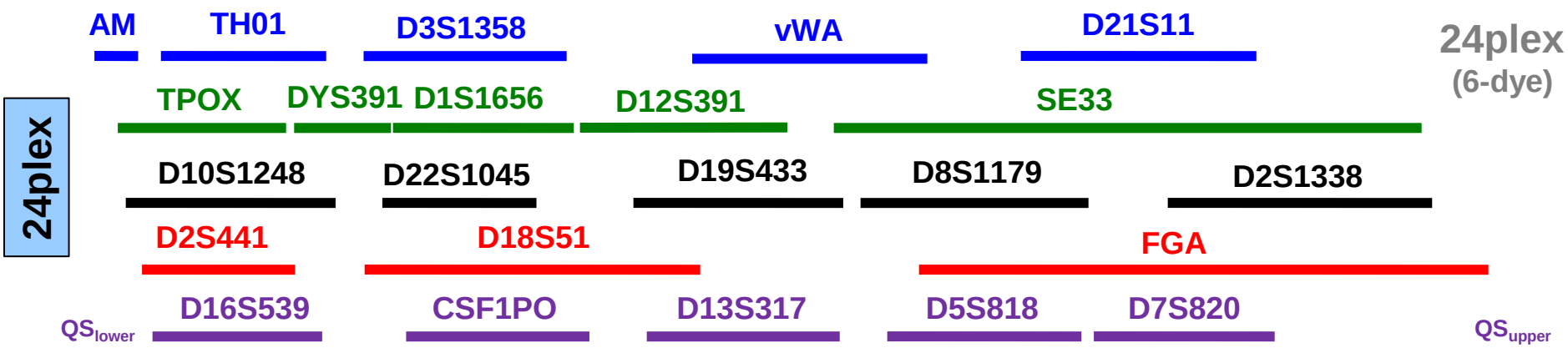
This unique Quality Sensor can distinguish:

- Successful amplification
- Degraded DNA
- Inhibited DNA
- No DNA
- Failed PCR amplification

Join us to learn how to improve results and streamline your forensic STR analysis with the new Investigator 24plex Kits.

http://www.qiagen.com/resources/e-learning/webinars/human-identity-and-forensics/multiplex-pcr-global-str/

Qiagen Investigator **24plex**



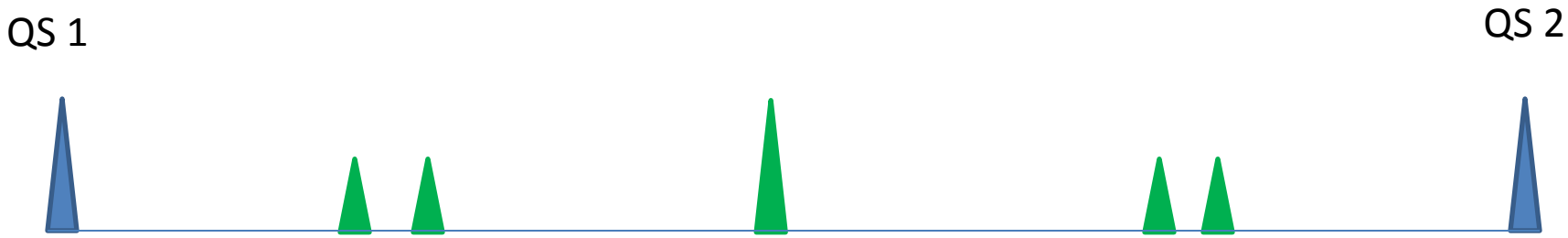
- 23 STR loci in 6 dyes (3500 use or 3130 upgrade required)
  - Includes SE33
- Includes Quality Sensors for detecting degraded and/or inhibited DNA
- 24plex GO!: direct amplification capabilities
  - Single source samples: 45 min amplification
- 24plex
  - Casework samples: 60 min amplification

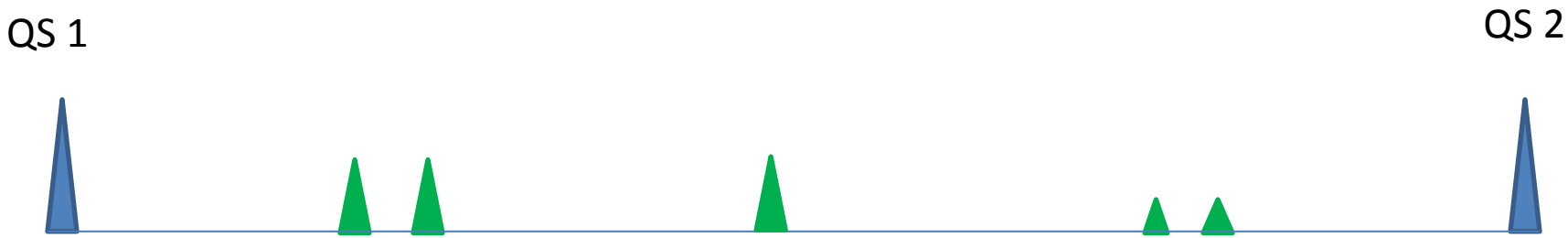
Two separate kits

- Qiagen 24plex gives ~12 orders of magnitude improvement using the NIST 1036 data set

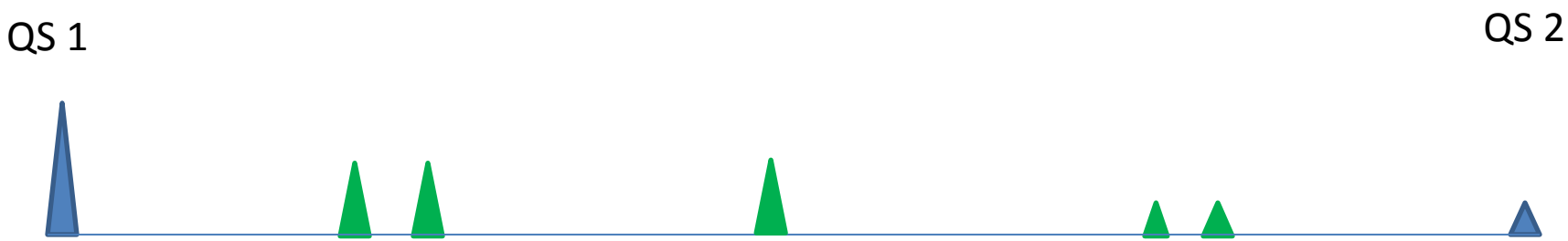
http://www.qiagen.com/resources/e-learning/webinars/human-identity-and-forensics/multiplex-pcr-global-str/

Quality Sensors



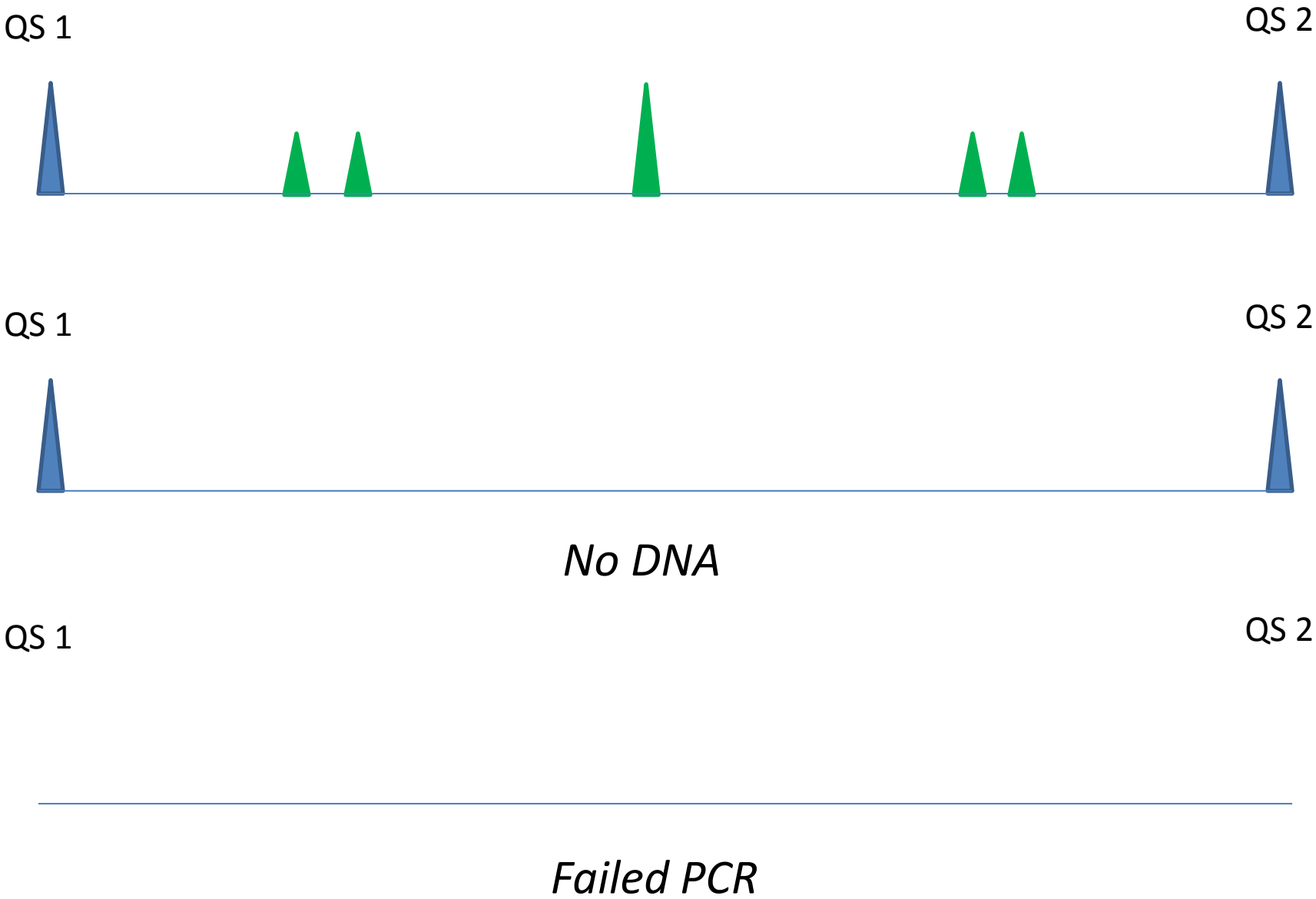


Degradation



Inhibition

Quality Sensors



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

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.

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

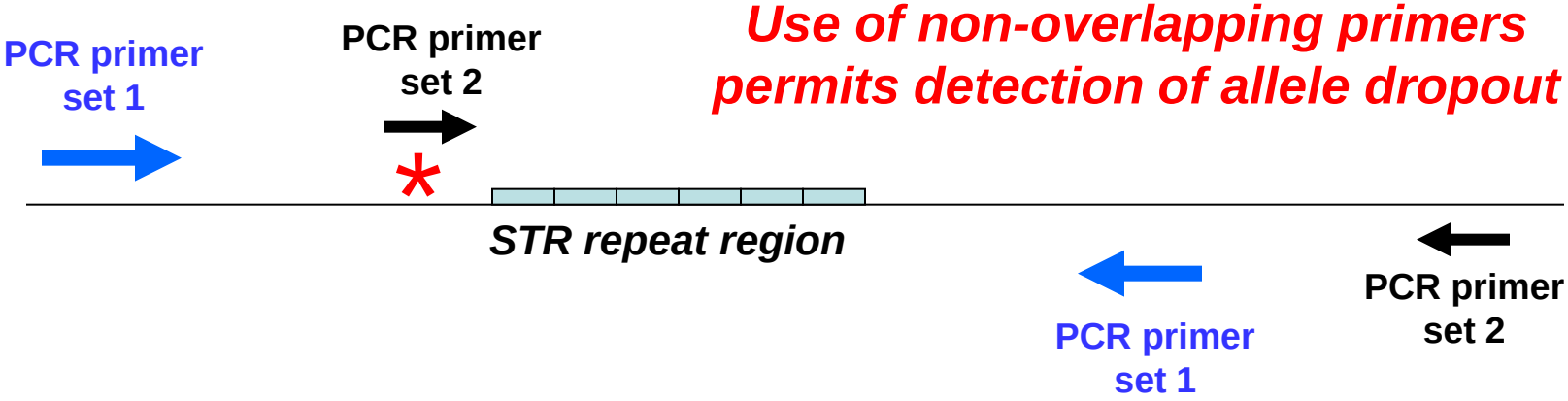
- **Related individuals removed** based on Y-STR and mtDNA results

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

* represents potential mutations impacting primer annealing

Use of non-overlapping primers permits detection of allele dropout



If no primer binding site mutations

Set 1 Amplicons = **Set 2 Amplicons**

Completed Concordance Studies

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/Qiplex	1426	16	22,816	28	99.873
ID/PP16	662	14	9,268	4	99.957
ID/MiniFiler	1137	9	10,233	38	99.746
ID/NGM	1407	11	15,837	3	99.981
ID/NGMx	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/Qiplex	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/NGMx	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+/Qiplex	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+/NGMx	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+/Qiplex	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+/NGMx	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
Qiplex/MiniFiler	1137	9	10,233	48	99.531
Qiplex/NGM	1426	11	15,686	30	99.809
Qiplex/NGMx	657	11	7,227	17	99.765
Qiplex/ESX17	1426	11	15,686	28	99.821
Qiplex/ESI17	1426	11	15,686	27	99.828
Qiplex/ESS	1426	11	15,686	1	99.994
Qiplex/ESSplexSE	657	11	7,227	1	99.986
Qiplex/Hexaplex	653	2	1,306	1	99.923
MiniFiler/NGM	1137	6	6,822	13	99.809
MiniFiler/NGMx	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/NGMx	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
NGMx/ESX17	662	17	11,254	4	99.964
NGMx/ESI17	662	17	11,254	14	99.876
NGMx/ESS	653	16	10,448	17	99.837
NGMx/ESSplexSE	662	17	11,254	34	99.698
NGMx/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/NGMx	663	1	663	4	99.397
SE33/ESSplexSE	662	1	662	21	96.828
ESX17p/ESX17	477	17	8,109	7	99.914
ESX17p/NGMx	477	17	8,109	2	99.975
ESX17p/ESSplexSE	477	17	8,109	42	99.482
ESI17p/SE33	477	1	477	4	99.161
PP16D/D3	50	16	800	2	99.750
PP16D/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*/NGMx	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
2/qiplex/ESX17	1443	3	4329	4	99.908
2/qiplex/ESI17	1443	3	4329	0	100.000
2/qiplex/NGM	1437	3	4311	11	99.745
2/qiplex/NGMx	663	3	1989	0	100.000
2/qiplex/ESS	1433	3	4299	0	100.000
2/qiplex/ESSplexSE	662	3	1986	0	100.000
2/qiplex/Hexaplex	653	3	1959	2	99.898
2/qiplex/ESX17*	663	3	1989	0	100.000
minSTRs/ESX17	663	3	1989	3	99.849
minSTRs/ESI17	663	3	1989	0	100.000
minSTRs/NGM	657	3	1971	3	99.848
minSTRs/NGMx	663	3	1989	0	100.000
minSTRs/ESS	653	3	1959	0	100.000
minSTRs/ESSplexSE	662	3	1986	0	100.000
minSTRs/Hexaplex	653	3	1959	2	99.898
minSTRs/ESX17*	663	3	1989	0	100.000
PP21/IdentifilerPlus	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/Qiplex	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/NGMx	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	2	99.912
PP21/Smoffiler	568	16	9088	1	99.989
PP21/IdentifilerPlus	639	16	10224	4	99.961
PP21/PP16	639	16	10224	1	99.990
PP21/Qiplex	639	16	10224	18	99.824
PP16/IdentifilerPlus	652	16	10432	2	99.981
PP16/PP16	652	16	10432	2	99.981
PP16/SGM+	651	11	7161	1	99.986
PP16/Pro+	647	10	6470	1	99.985
PP16/Qiplex	652	16	10432	16	99.847
PP16/MiniFiler	652	9	5868	14	99.761
PP16/ESX17	652	17	11084	1	99.991
PP16/ESI17	652	17	11084	0	100.000
PP16/NGM	652	16	10432	1	99.990
PP16/NGMx	652	17	11084	1	99.991
PP16/ESS	652	16	10432	16	99.847
PP16/ESSplexSE	652	17	11084	16	99.856
PP16/Hexaplex	652	6	3912	2	99.949
PP16/Smoffiler	568	16	9088	1	99.989
PP16/PP21nk	639	21	13419	0	100.000
PP16/PP23	652	1	652	0	100.000
PP16/YF	652	1	652	0	100.000
PP16/PPY	652	1	652	0	100.000
2/qiplex/IdentifilerPlus	650	16	10400	2	99.981
2/qiplex/PP16	650	16	10400	2	99.981
2/qiplex/SGM+	650	11	7150	1	99.986
2/qiplex/Pro+	647	10	6470	1	99.985
2/qiplex/Qiplex	650	16	10400	16	99.846
2/qiplex/MiniFiler	650	9	5850	14	99.761
2/qiplex/ESX17	650	17	11050	1	99.991
2/qiplex/ESI17	650	17	11050	0	100.000
2/qiplex/NGM	650	16	10400	1	99.990
2/qiplex/NGMx	650	17	11050	1	99.991
2/qiplex/ESS	650	16	10400	16	99.846
2/qiplex/ESSplexSE	650	17	11050	16	99.855
2/qiplex/Hexaplex	650	6	3900	2	99.949
2/qiplex/Smoffiler	568	16	9088	1	99.989
2/qiplex/PP21nk	639	19	12141	0	100.000
2/qiplex/PPY	650	21	13650	0	100.000
2/qiplex/PP23	650	1	650	0	100.000
2/qiplex/YF	650	1	650	0	100.000
Totals	138,031	1712	1,404,800	1373	99.902

	Samples Typed	Loci Compared	# Allele Comparisons Made	# Differences	Concordance (%)
Totals	138,031	1712	1,404,800	1373	99.902

1,404,031 allele comparisons

1,373 total differences

99.9% concordance

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

SRM 2391b/2391c PCR-Based Profiling Standard

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

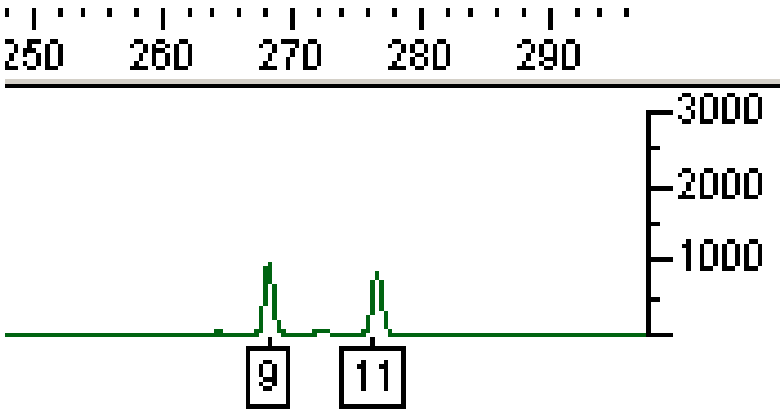
- All new kits tested have been completely concordant with the certified values of all markers for each component for SRM 2391b and 2391c

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

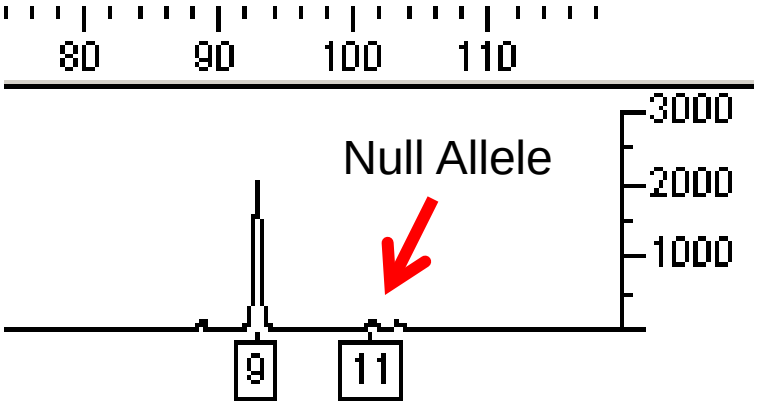
SRM 2391b Genomic 8 with D16S539

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

Identifiler

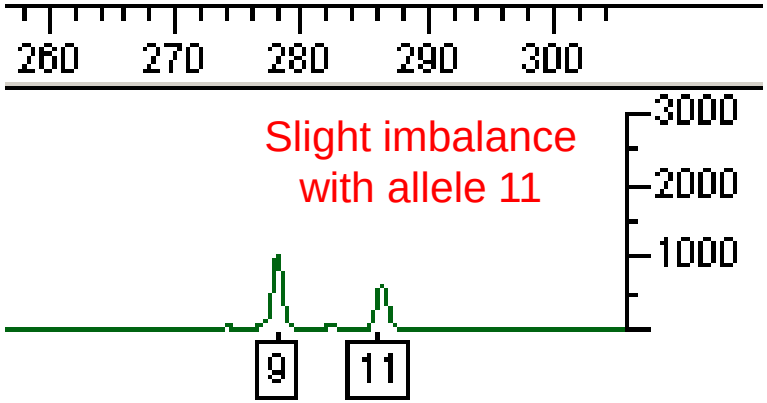


MiniFiler



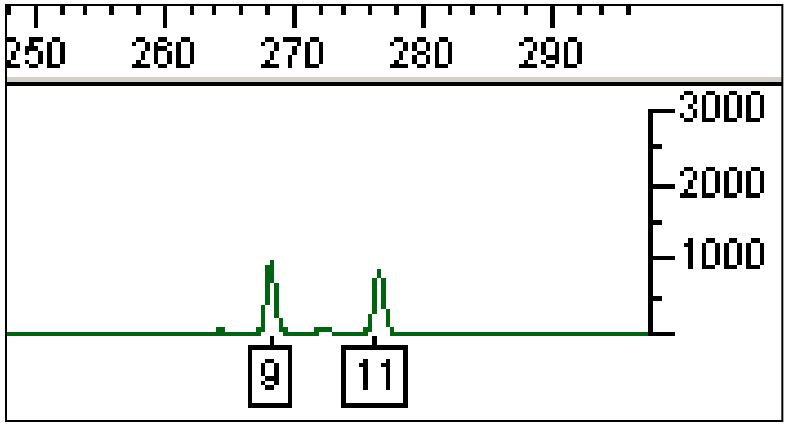
***Due to primer binding site mutation**

PowerPlex 16

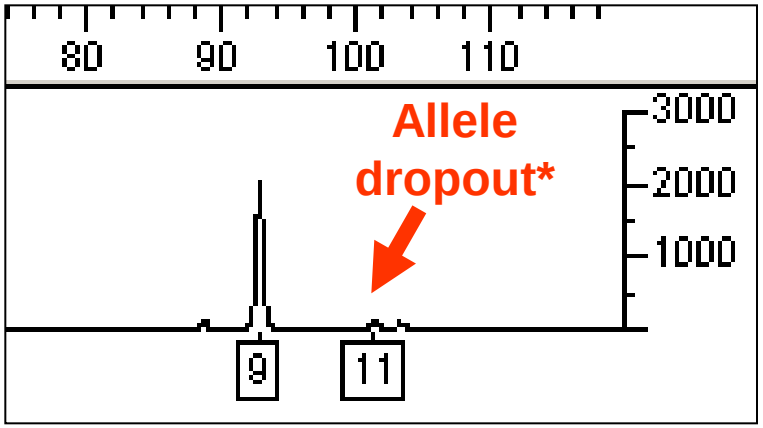


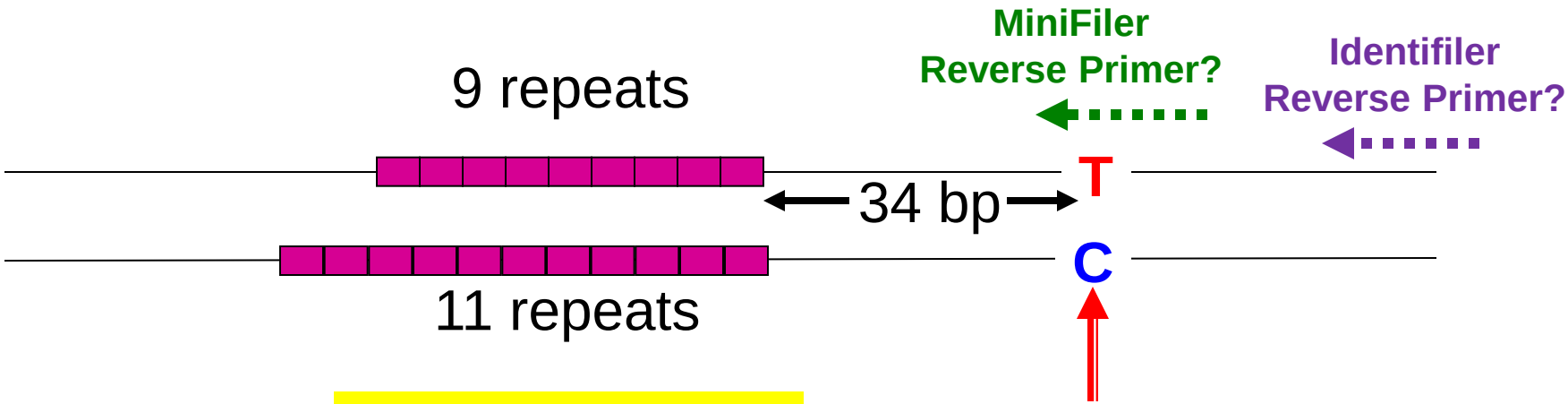
SRM 2391b Genomic 8 with D16S539

Identifiler



MiniFiler



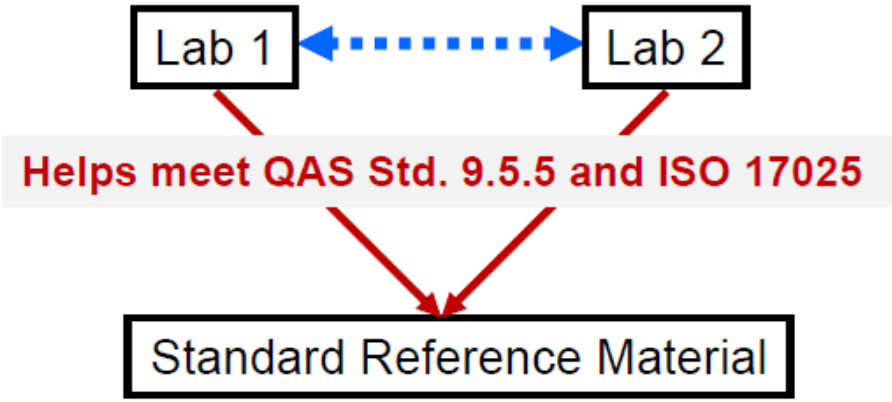


Type 9T,11C

SRM 2391c: PCR-Based DNA Profiling Standard

- Components A through D are DNA extracts in liquid form
- Components E and F are DNA spotted on 903 paper or FTA paper
- Certified values are for STR alleles based on length polymorphisms observed using capillary electrophoresis

Genomic DNAs characterized for the expanded CODIS core loci and Y-STRs



Calibration with SRMs enables confidence in comparisons of results between laboratories

Certified, Reference, & Information Values of SRM 2391c

- **Certified Values**: 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
  - **2 or more methods are used to compare values (i.e. Sanger sequencing, genotyping using multiple sets of primers)**

- **Reference Values**: High-confidence estimate of the true value but where all possible sources of bias have not been fully investigated by NIST
  - **Genotyping with only 2 sets of primers to compare**

- **Information Values**: Data that may be of interest and use to the SRM user, but insufficient information is available to access the confidence of the assignment
  - **Genotyping of only 1 kit is available**

Current 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

Other Markers that were Sequenced and are Certified in the Update

- Additional non-core autosomal STR markers
  - D6S1043 (Sinofiler, PowerPlex 21)

- Newly available Y-STR markers in commercial kits
  - DYS460, DYS481, DYS533, DYS549, DYS643, (PowerPlex Y23, Yfiler Plus)

- Rapidly mutating (RM) Y-STRs in commercial kits
  - DYS449, DYS518, DYS570, DYS576, DYS627, DYF387S1a/b (PowerPlex Y23, Yfiler Plus)

STR Typing Kits and Primer Mixes

Original SRM 2391c certificate

Kit Provider			
Thermo Fisher (12)	Promega Corp. (8)	Qiagen Inc. (2)	Primer Mixes (2)
Identifiler	PowerPlex 16	ESSplex	26plex
Identifiler Plus	PowerPlex 16 HS	IDplex	miniSTRs
NGM	PowerPlex ESX 17		
NGM SElect	PowerPlex ESI 17		
COfiler	PowerPlex ES		
Profiler	PowerPlex S5		
Profiler Plus	PowerPlex Y		
Profiler Plus ID	FFFL		
SGM Plus			
SEfiler			
MiniFiler			
Yfiler			

100% Concordance with all kits

Updated SRM 2391c certificate

Kit Provider			
Thermo Fisher (14)	Promega Corp. (16)	Qiagen Inc. (9)	Primer Mixes (3)
Identifiler	PowerPlex 16	ESSplex	26plex
Identifiler Plus	PowerPlex 16 HS	IDplex	miniSTRs
NGM	PowerPlex ESX 17	ESSplex SE	RM Y STRs
NGM SElect	PowerPlex ESI 17	ESSplex SE Plus	
COfiler	PowerPlex ES	ESSplex SE GO!	
Profiler	PowerPlex S5	IDplex Plus	
Profiler Plus	PowerPlex Y	IDplex GO!	
Profiler Plus ID	FFFL	Argus X-12	
SGM Plus	PowerPlex ESI 17 Pro	DIPplex	
SEfiler	PowerPlex ESX 17 Fast		
MiniFiler	PowerPlex ESI 17 Fast		
Yfiler	PowerPlex 18D		
GlobalFIler	PowerPlex 21		
Yfiler Plus	PowerPlex CS7		
	PowerPlex Fusion		
	PowerPlex Y23		

Updated Values for STR Loci

Certified Values		
Autosomal STR (25)	Y STR (29)	X STR (0)
D1S1656	DYS19	None
D2S1338	DYS385a	
D2S441	DYS385b	
D3S1358	DYS389I	
D5S818	DYS389II	
D6S1043	DYS390	
D7S820	DYS391	
D8S1179	DYS392	
D8S1115	DYS393	
D10S1248	DYS437	
D12S391	DYS438	
D13S317	DYS439	
D16S539	DYS448	
D18S51	DYS456	
D19S433	DYS458	
D21S11	DYS635	
D22S1045	Y GATA H4	
CSF1PO	DYS449	
FGA	DYS460	
Penta D	DYS481	
Penta E	DYS518	
SE33	DYS533	
TH01	DYS549	
TPOX	DYS570	
vWA	DYS576	
	DYS627	
	DYS643	
	DYF387S1a	
	DYF387S1b	

Reference Values		
Autosomal STR (23)	Y STR (0)	X STR (0)
D1GATA113	None	None
D1S1627		
D1S1677		
D2S1776		
D3S3053		
D3S4529		
D4S2364		
D4S2408		
D5S2500		
D6S1017		
D6S474		
D9S1122		
D9S2157		
D10S1435		
D17S1301		
D17S974		
D18S853		
D20S1082		
D20S482		
F13A01		
F13B		
FESFPS		
LPL		

Information Values		
Autosomal STR (1)	Y STR (0)	X STR (12)
Penta C	None	DXS7132
		DXS7423
		DXS8378
		DXS10074
		DXS10079
		DXS10101
		DXS10103
		DXS10134
		DXS10135
		DXS10146
		DXS10148
		HPRTB

Argus X-12 kit

*Other Information Value Considerations:
-DIPplex kit
-PGM and MiSeq IISNPs

New Y-STR loci in commercial kits (Yfiler Plus & PPY23)

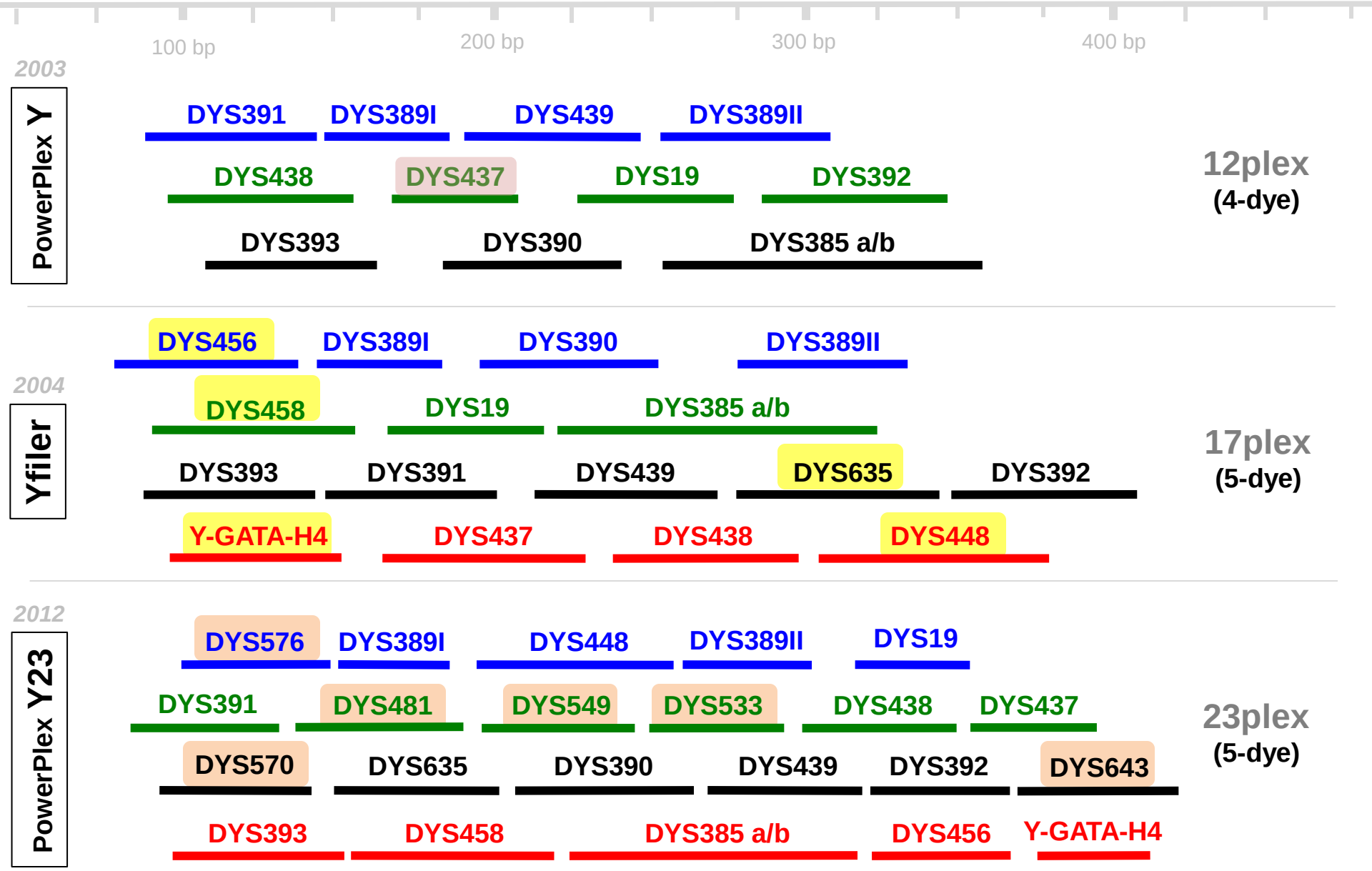
SRM 2391c Certificate is now updated

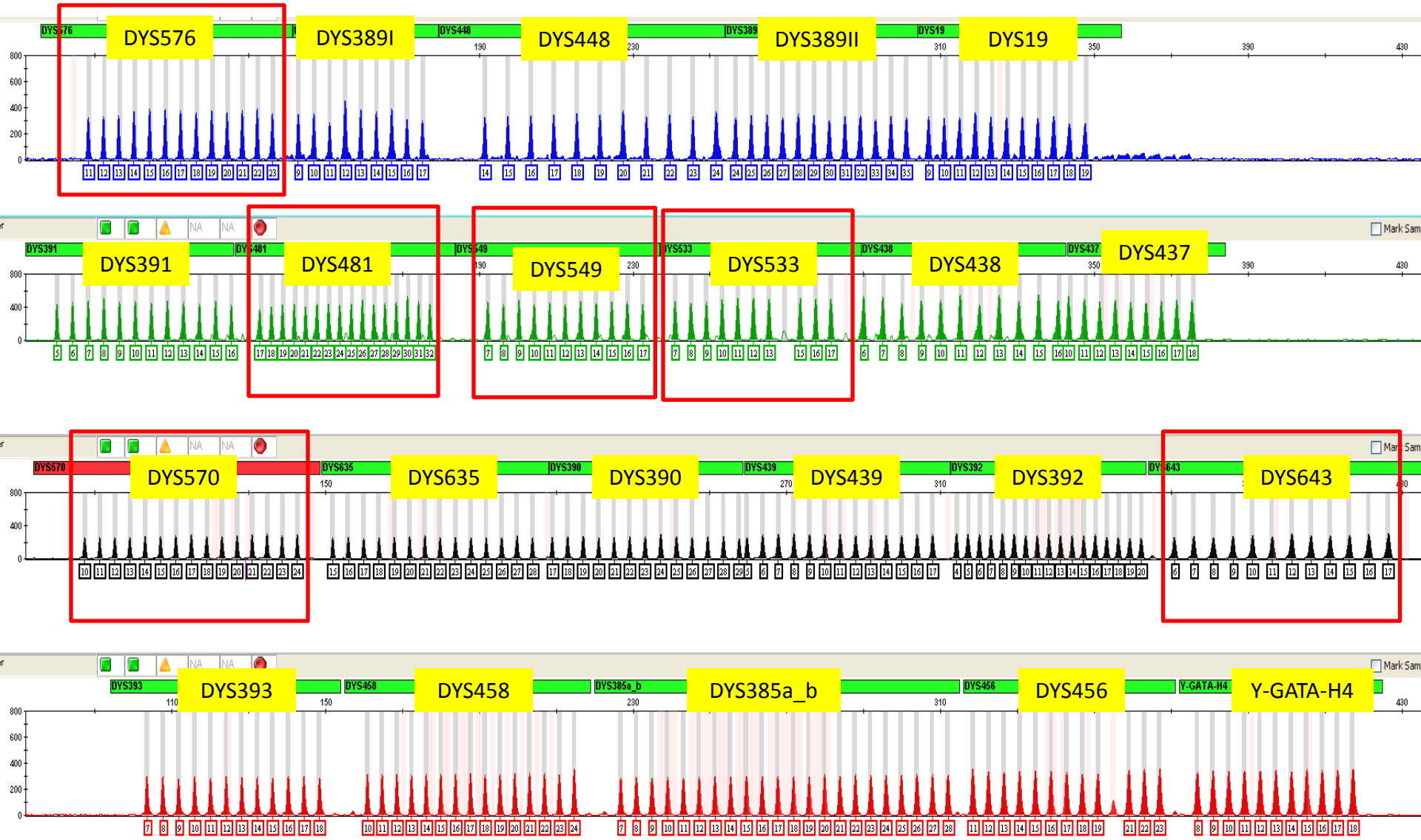
SRM 2391c Update

- Sequencing of Components A-C, E & F has been completed for all remaining autosomal and Y-STR loci
- The SRM 2391c Certificate of Analysis and Report of Analysis has been 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

New Y-STRs

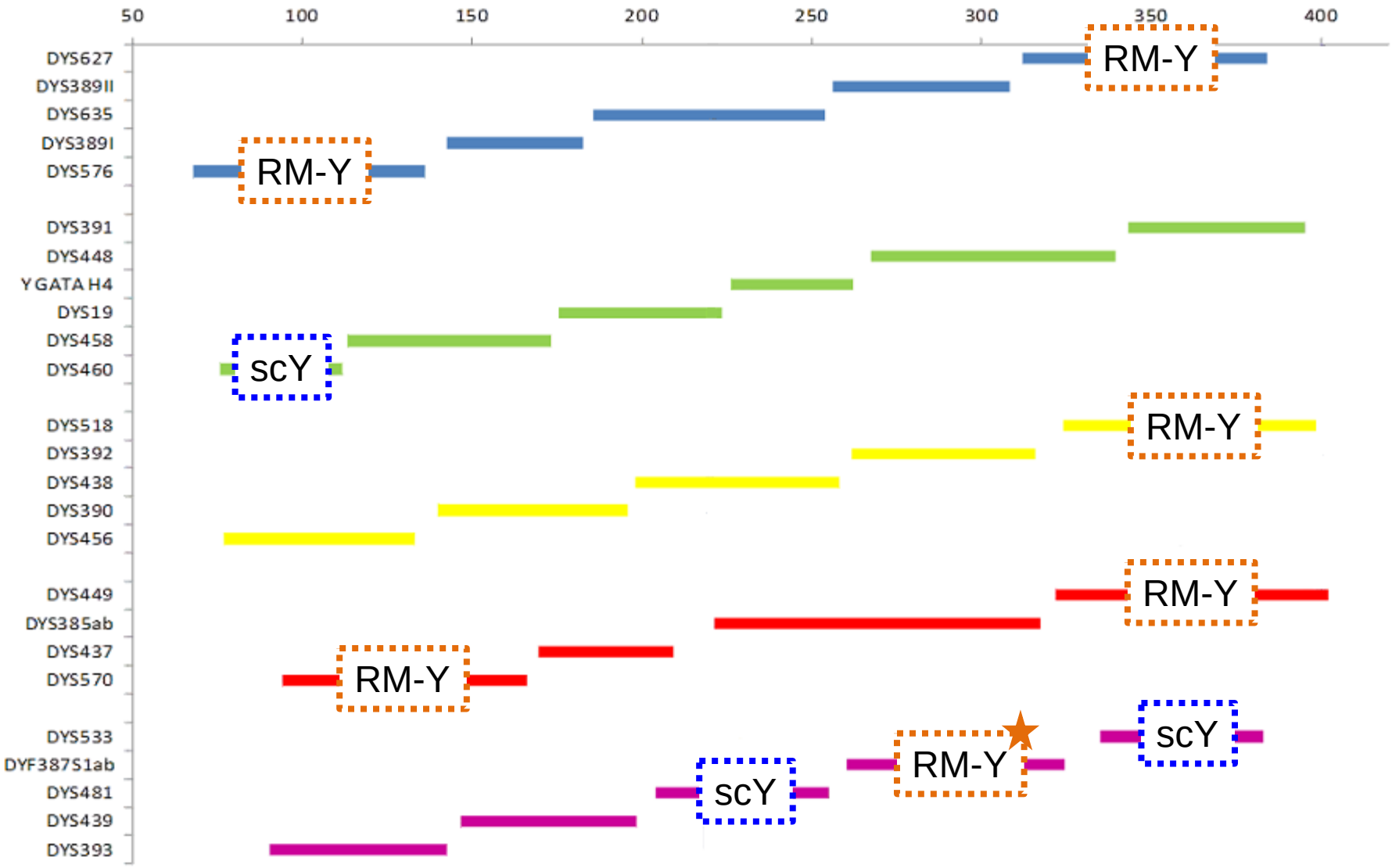
STR Marker Layouts for Y-STR Kits



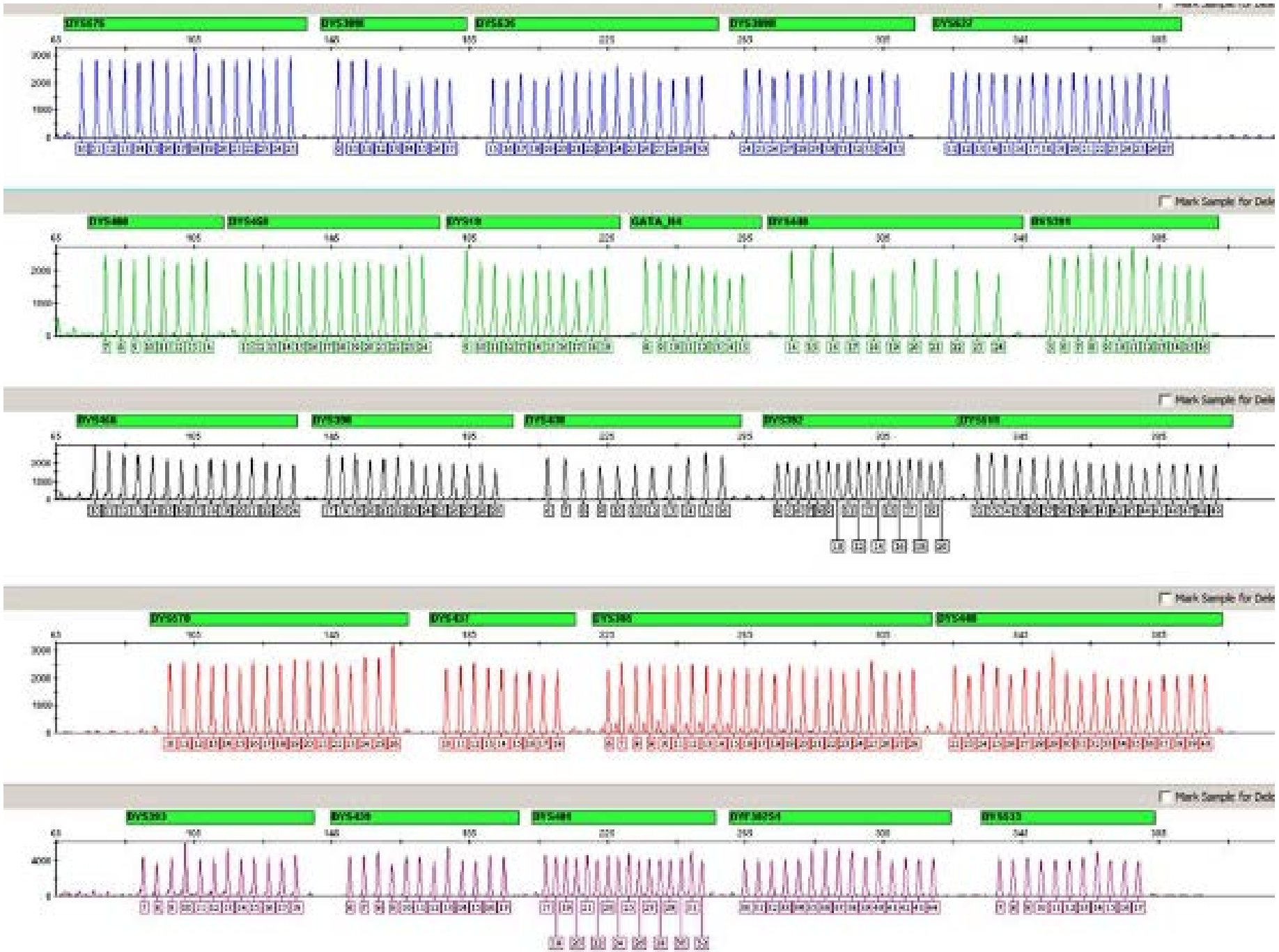


17 Y-Filer + 6 additional loci

Yfiler Plus Kit

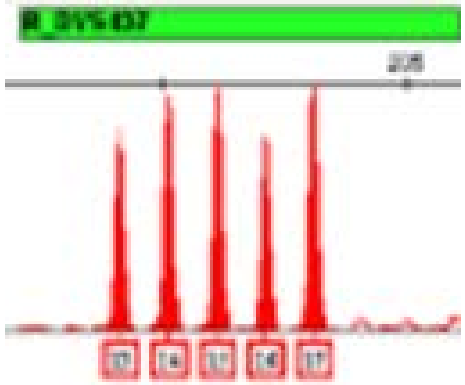


Slide courtesy of Dr. Julio Mulero (Life Technologies)



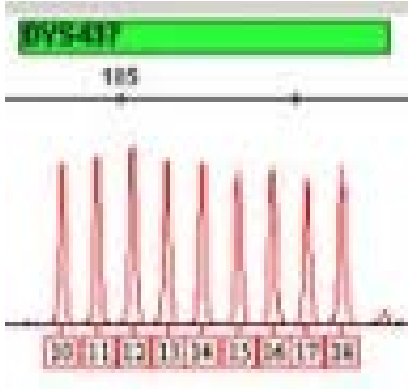
Improved Allelic Ladder (DYS437)

Yfiler



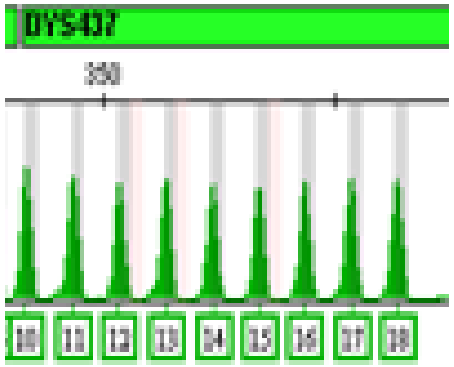
5 Alleles

Yfiler-Plus



9 Alleles

PPY-23



9 Alleles

Performance with related males

Relatives in the NIST Population Samples

Exact matches among the NIST population plates (Y-Filer)

Sample Name	DYS19	DYS389I	DYS389II	DYS390	DYS391	DYS392	DYS393	DYS385a/b	DYS438	DYS439	DYS437	DYS456	DYS458	DYS635	DYS448	Y-GATA-H4
OT07753	14	13	29	24	11	13	13	11,14	12	12	15	15	17	23	19	12
UC10177	14	13	29	24	11	13	13	11,14	12	12	15	15	17	23	19	12
Y16	17	13	30	21	10	11	14	15,20	11	12	14	15	16	22	21	12
Y17	17	13	30	21	10	11	14	15,20	11	12	14	15	16	22	21	12
ZT79304	16	13	30	21	10	11	13	16,17	11	11	14	15	18	21	22	12
PT84253	16	13	30	21	10	11	13	16,17	11	11	14	15	18	21	22	12
WA29612	14	13	29	23	11	13	12	12,14	12	12	15	15	17	23	19	12
WT52482	14	13	29	23	11	13	12	12,14	12	12	15	15	17	23	19	12
WT52470	14	12	28	23	10	11	13	14,15	10	11	16	14	15	21	20	11
MT94827	14	12												21	20	11
Y27	14	13												23	19	13
Y28	14	13												23	19	13
ZT79994	14	13												23	19	12
ZT79995	14	13												23	19	12
PT84213	14	14												22	20	11
TT51696	14	14												22	20	11
MT94839	15	14	31	25	11	11	13	11,14	11	10	14	15	15	23	20	13
MT97158	15	14	31	25	11	11	13	11,14	11	10	14	15	15	23	20	13
JT51484	17	13	30	21	10	11	14	17,19	11	13	14	16	16	21	21	11
ZT79620	17	13	30	21	10	11	14	17,19	11	13	14	16	16	21	21	11
MT97176	14	13	30	24	11	13	13	11,14	12	12	15	17	16	23	19	12
Y4	14	13	30	24	11	13	13	11,14	12	12	15	17	16	23	19	12
JT51476	14	13	29	25	11	14	13	11,13	12	12	15	17	17	23	18	12
UT58337	14	13	29	25	11	14	13	11,13	12	12	15	17	17	23	18	12
TT51702	14	13	29	24	11	13	13	11,14	12	12	15	16	17	23	19	12
MT97185	14	13	29	24	11	13	13	11,14	12	12	15	16	17	23	19	12
ZT79333	14	13	29	24	11	13	13	11,14	12	12	15	16	17	23	19	12
PT83904	13	15	31	24	9	11	13	13,14	10	10	14	16	18	21	20	12
PT84348	13	15	31	24	9	11	13	13,14	10	10	14	16	18	21	20	12
ZT80369	13	15	31	24	9	11	13	13,14	10	10	14	16	18	21	20	12

12 sets of 2 individuals

2 sets of 3 individuals

Adding 6 markers from PP-Y23

Sample Name	DYS576	DYS570	DYS643	DYS533	DYS481	DYS549
OT07753	18	**17**	10	**12**	**23**	13
UC10177	18	**18**	10	**13**	**22**	13
Y16	14	19	14	11	29	11
Y17	14	19	14	11	29	11
ZT79304	15	19	14	11	28	11
PT84253	15	19	14	11	28	11
WA29612	16	17	10	12	22	**14**
WT52482	16	17	10	12	22	**13**
WT52470	**16**	**23**	12	11	**26**	12
MT94827	**17**	**22**	12	11	**25**	12
Y27	19	17	10	12	22	13
Y28	19	17	10	12	22	13
ZT79994	20	16	10	12	23	12
ZT79995	20	16	10	12	23	12

6 pairs still unresolved

Two Samples from the Population Plates ZT79994 and ZT79995

Match exactly over 23 Y-STRs

ZT79994	Pos	Seq	ZT79995	Pos	Seq
	16172	T - C		16172	T - C
	16219	A - G		16219	A - G
	16311	T - C		16311	T - C
	16344	C -T		16344	C -T
	73	A -G		73	A -G
	263	A- G		263	A- G
	315.1	C		315.1	C
	469	C - T		469	C - T
	523	A-del		523	A-del
	524	C-del		524	C-del

Same mtDNA haplotype (U6a)

Autosomal STR Analysis

Sample Info	CSF1PO 1	CSF1PO 2		D13S317 1	D13S317 2		D16S539 1	D16S539 2		D18S51 1	D18S51 2
ZT79994	9	11		12	14		13	13		13	14
ZT79995	11	12		11	14		11	13		13	14
	D19S433 1	D19S433 2		D21S11 1	D21S11 2		D2S1338 1	D2S1338 2		D3S1358 1	D3S1358 2
ZT79994	13	14		29	30		17	25		16	17
ZT79995	14	16		30	30		17	25		17	18
	D5S818 1	D5S818 2		D7S820 1	D7S820 2		D8S1179 1	D8S1179 2		FGA 1	FGA 2
ZT79994	12	12		7	12		13	13		20	21
ZT79995	13	13		8	12		13	13		20	23
	TH01 1	TH01 2		TPOX 1	TPOX 2		vWA 1	vWA 2		D10S1248	D10S1248
ZT79994	7	9.3		8	8		15	18		14	14
ZT79995	7	9.3		8	11		15	17		14	14
	D12S391	D12S391		D1S1656	D1S1656		D22S1045	D22S1045		D2S441	D2S441
ZT79994	21	21		13	13		15	16		10	11
ZT79995	18.3	20		13	16.3		16	17		10	11
	SE33	SE33		Penta_D	Penta_D		Penta_E	Penta_E		D6S1043	D6S1043
ZT79994	16	17		9	12		13	14		12	18
ZT79995	17	23		11	14		13	14		12	12

24 STR Markers

Kinship Analysis

	Combined
Parent-Child	0
Full Sib	**56327**
Half-Sib*	5692
1st Cousin	264

*Is also the same stat for Aunt/Uncle/Niece/Nephew and Grandparent/Grandchild relationships

Performance with related males

Father-Son Mutation Rates

Father-Son Mutations - Summary

	Meioses	Mutations	
AfAm	85	10	
Asian	101	16	
Caucasian	100	6	
Hispanic	100	12	
Sum	386	44	11.40%

	Father to Son		Father to Son	
	+1 repeat	-1 repeat	+2 repeat	-2 repeat
DYS439	2	4		
DYS389II	3	2		
DYS481	3	2		
DYS570	3	2		
DYS576	2	2		
DYS389I	2	2		
Y-GATA-H4	1	1	1	1
DYS458	2	1		
DYS549	1	1		
DYS635	1	1		
DYS19	1	0		
DYS390	0	1		
DYS643	1	0		
DYS385a	1	0		
DYS448	0	0		
DYS391	0	0		
DYS533	0	0		
DYS438	0	0		
DYS437	0	0		
DYS392	0	0		
DYS393	0	0		
DYS385b	0	0		
DYS456	0	0		
(sum)	23	23	1	1

	Mutations
DYS439	6
DYS389II	5
DYS481	5
→ **DYS570**	5
→ **DYS576**	4
DYS389I	4
Y-GATA-H4	4
DYS458	3
DYS549	2
DYS635	2
DYS19	1
DYS390	1
DYS643	1
DYS385a	1
DYS448	0
DYS391	0
DYS533	0
DYS438	0
DYS437	0
DYS392	0
DYS393	0
DYS385b	0
DYS456	0

G Model
FSIGEN-744; No. of Pages 11

ARTICLE IN PRESS

Forensic Science International: Genetics xxx (2011) xxx–xxx

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

A new future of forensic Y-chromosome analysis: Rapidly mutating Y-STRs for differentiating male relatives and paternal lineages

Kaye N. Ballantyne[a,1,2], Victoria Keerl[a,1,3], Andreas Wollstein[a,b], Ying Choi[a], Sofia B. Zuniga[c], Arwin Ralf[a], Mark Vermeulen[a], Peter de Knijff[c], Manfred Kayser[a,*]

[a] Department of Forensic Molecular Biology, Erasmus MC University Medical Center Rotterdam, 3000 CA Rotterdam, The Netherlands
[b] Cologne Center for Genomics, University of Cologne, D-50674 Cologne, Germany
[c] Department of Human Genetics, Leiden University Medical Center, 2300 RC Leiden, The Netherlands

The American Journal of Human Genetics 87, 341–353, September 10, 2010 **ARTICLE**

Mutability of Y-Chromosomal Microsatellites: Rates, Characteristics, Molecular Bases, and Forensic Implications

Kaye N. Ballantyne,[1] Miriam Goedbloed,[1] Rixun Fang,[2] Onno Schaap,[1] Oscar Lao,[1] Andreas Wollstein,[1,3] Ying Choi,[1] Kate van Duijn,[1] Mark Vermeulen,[1] Silke Brauer,[1,4] Ronny Decorte,[5] Micaela Poetsch,[6] Nicole von Wurmb-Schwark,[7] Peter de Knijff,[8] Damian Labuda,[9] Hélène Vézina,[10] Hans Knoblauch,[11] Rüdiger Lessig,[12] Lutz Roewer,[13] Rafal Ploski,[14] Tadeusz Dobosz,[15] Lotte Henke,[16] Jürgen Henke,[16] Manohar R. Furtado,[2] and Manfred Kayser[1,*]

Yfiler Plus Kit

Released August 2014

Yfiler-Plus Kit Performance

Sensitivity

Sensitivity Experiment

Full reaction volumes (25 ul):

- 0.5 ng single amp
- 0.25 ng single amp
- 0.125 ng duplicate amp
- 0.0625 ng duplication amp
- 0.03125 ng duplicate amp

Half reaction volumes (12.5 ul):

- 0.5 ng duplicate amp
- 0.25 ng duplicate amp
- 0.125 ng duplicate amp
- 0.0625 ng duplication amp
- 0.03125 ng duplicate amp

Sensitivity Experiment

FULL REACTION VOLUME (25 ul)								
Marker	0.03125 ng		0.0625 ng		0.125 ng		0.25 ng	0.5 ng
DYS576		184	222	259	491	574	491	2308
DYS389I	96	151	344	318	1083	895	1603	2581
DYS635	195	153	198	152	379	473	1341	2481
DYS389II		115	279	176	725	648	1036	1847
DYS627	142	373	165	140	1257	591	1122	3639
DYS460		463	304	129	263	328	827	2786
DYS458		404	51	261	466	581	1091	3107
DYS19		173	220	143	335	525	1250	2697
GATA_H4	97		159	226	482	153	1896	3212
DYS448	91		154	204	500	196	2255	2489
DYS391	342		231	51	1169	721	2470	3789
DYS456	330	186	269	115	732	241	1524	2143
DYS390		232	123	87	225	331	478	1664
DYS438	144	166	207	205	557	622	999	3563
DYS392		209	53	129	700	156	840	2293
DYS518	232	118	145	185	638	411	619	2322
DYS570		92	163	300	987	327	1673	3398
DYS437	83	248	364	152	749	485	1429	2494
DYS385	109	364	303	135	525	545	1261	2608
DYS385			130		383		1539	1912
DYS449	57	91	321	218	713	482	846	3598
DYS393			485	126	172	389	2184	2353
DYS439		278	130	108	916	313	668	3488
DYS481	254		375	218	1022	588	1707	3605
DYF387S1		203	208	169	413	487	1530	2173
DYF387S1			197	192	696	428	1050	2846
DYS533	209		280	93	942	456	1987	2772

≥ 175 RFU

< 175 RFU

Sensitivity Experiment

HALF REACTION VOLUME (12.5 ul)										
Marker	0.03125 ng		0.0625 ng		0.125 ng		0.25 ng		0.5 ng	
DYS576	84	404	740	484	2271	551	3580	1773	4986	8103
DYS389I		1080	1285	2442	1064	2389	3992	2964	8646	12160
DYS635			331	911	1540	1950	3991	965	7439	5402
DYS389II		1055	717	2561	1026	1985	3080	2211	4781	8069
DYS627			1175	876	1340	2054	2944	2168	12000	4860
DYS460		785	605	1469	2961	2146	4028	2445	3443	5921
DYS458		394	262	2099	2616	1013	4972	2265	5029	6217
DYS19		835	351	1283	1886	1464	2492	2473	4532	4871
GATA_H4		245	475	1690	1766	1274	3980	1493	2813	5701
DYS448			594	1842	1131	1611	4011	2042	3175	5857
DYS391	199	398	1140	398	2861	1632	3359	3067	6691	10923
DYS456	122		121	920	1280	1391	2803	851	3190	6202
DYS390			214	673	771	1428	2059	1212	3429	4472
DYS438	321	155	656	717	2163	1269	2201	1646	5038	6472
DYS392	153		1306	446	1853	858	2242	1960	5151	5722
DYS518	367	262	844	280	771	1661	1657	997	3357	4083
DYS570	196	592	376	1165	1754	1551	3068	2200	4346	7802
DYS437	145	1152	404	325	1424	1582	6373	1199	6266	7528
DYS385	93	613	630	1672	2218	1764	3592	1671	3160	5944
DYS385		197	726	585	1434	894	3548	2302	3706	5323
DYS449	79	112	420	1727	2309	2752	4809	2858	5308	5506
DYS393	77	824	719	621	2240	1716	3786	2111	4809	6484
DYS439		286	841	1266	1449	2781	2165	2468	4563	5620
DYS481	83	618	689	1368	3258	1737	3684	3348	7554	7676
DYF387S1	51	296	842	999	2175	1016	3089	1306	3782	8096
DYF387S1		448	499	1447	1655	2940	2757	2400	4758	4924
DYS533		527	494	4033	2288	952	2599	1869	5163	6919

≥ 175 RFU

< 175 RFU

Yfiler-Plus Kit Performance

F:M Mixtures

High Female [DNA] Experiment

Looking for:

Artifacts

Inhibition

Three male samples (B, C, F) were each combined with same female

Input female DNA was constant at approximately 200 ng

Input male DNA at four levels in the range of approx. 1 ng – 0.05 ng

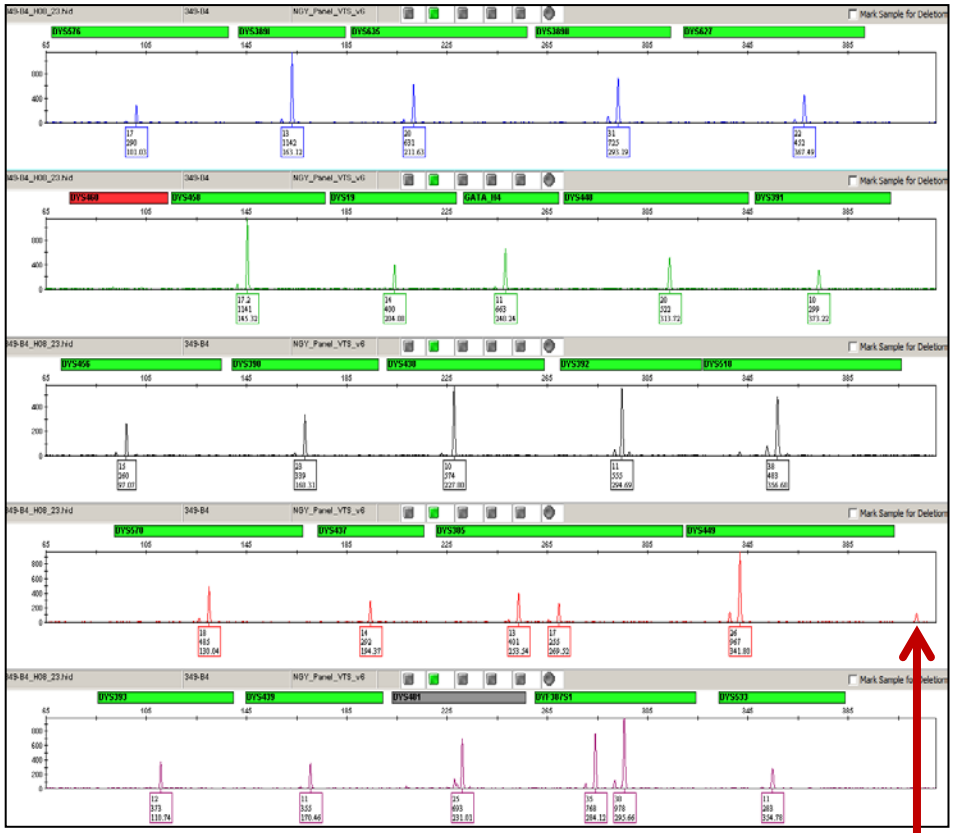
Male only samples for comparison

24 samples total

200 : 1 200 ng + 1 ng	1 ng
700 : 1 200 ng + 0.3 ng	0.3 ng
2000 : 1 200 ng + 0.1 ng	0.1 ng
4000 : 1 200 ng + 0.05 ng	0.05 ng

High Level Female Experiment

Male Sample B



Artifact?

4000 : 1

(approx. 200 ng female + 0.05 ng male)

(approx. 0.05 ng)

349-B1_E08_14.hid
349-B1
NGY_Panel_VTS_v6
Mark Sample for Deletion
DYS449
200 : 1
26
4291
341.81
349-B2_F08_17.hid
349-B2
NGY_Panel_VTS_v6
Mark Sample for Deletion
700 : 1
26
5173
341.80
349-B3_G08_20.hid
349-B3
NGY_Panel_VTS_v6
Mark Sample for Deletion
2000 : 1
26
763
341.80
349-B4_H08_23.hid
349-B4
NGY_Panel_VTS_v6
Mark Sample for Deletion
4000 : 1
26
967
341.80

DYS449
B1_A07_01.hid
B1
NGY_Panel_VTS_v6
Mark Sample for Deletion
26
12420
341.79
B2_B07_04.hid
B2
NGY_Panel_VTS_v6
Mark Sample for Deletion
26
8951
341.80
B3_C07_07.hid
B3
NGY_Panel_VTS_v6
Mark Sample for Deletion
26
2009
341.80
B4_D07_10.hid
B4
NGY_Panel_VTS_v6
Mark Sample for Deletion
26
1159
341.79

349-B1_E08_14.hid
349-B1
NGY_Panel_VTS_v6
Mark Sample for Deletion
DYS449
200 : 1
26
4291
341.81
349-B2_F08_17.hid
349-B2
NGY_Panel_VTS_v6
Mark Sample for Deletion
700 : 1
26
5173
341.80
349-B3_G08_20.hid
349-B3
NGY_Panel_VTS_v6
Mark Sample for Deletion
2000 : 1
26
763
341.80
349-B4_H08_23.hid
349-B4
NGY_Panel_VTS_v6
Mark Sample for Deletion
4000 : 1
“TAZ 413” artifact
26
967
341.80

Performance with unrelated males

NIST U.S. Population Samples

Discrimination Capacity

- is a measure of the number of unique haplotypes in a given population

$$DC = \frac{\#H}{N}$$

# of Haplotypes

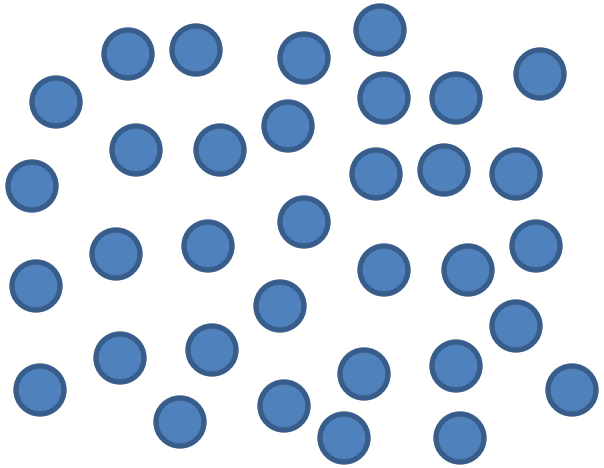
Population size

Results so far (582 individuals)

582 Males

Yfiler Plus

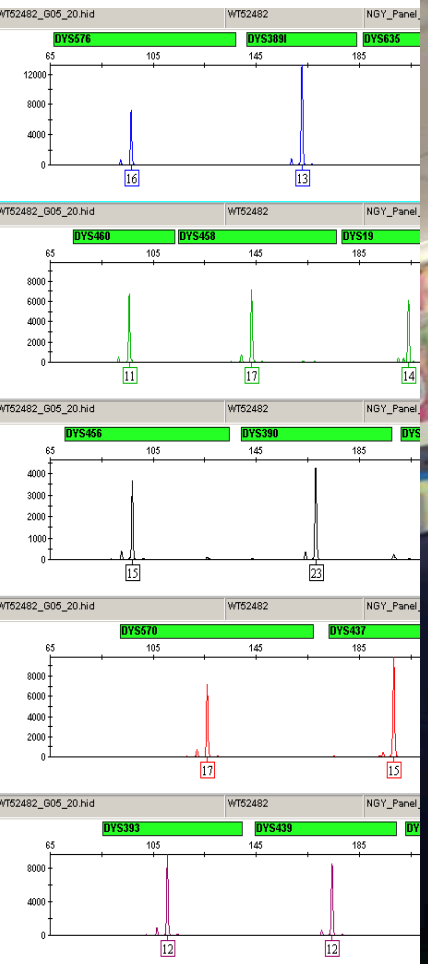
27 Y-STRs

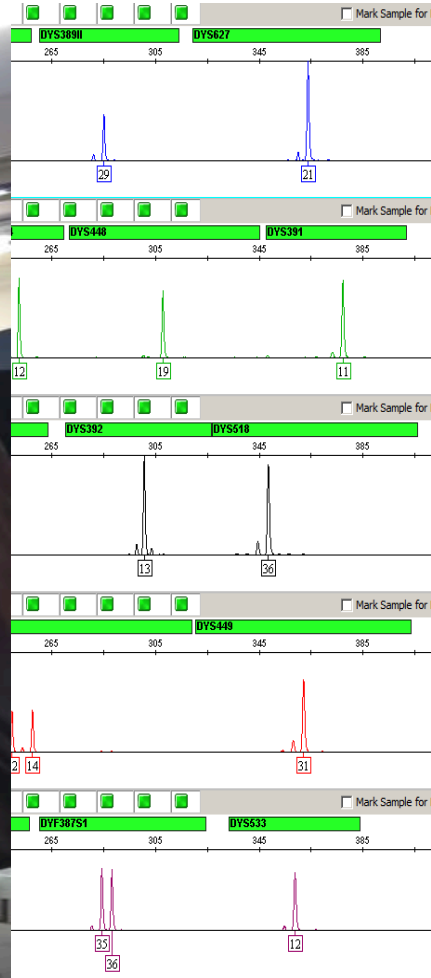


580 Unique Haplotypes

1 Shared Haplotype

DC = 0.998


WT52482_G05_20.hid
WT52482
NGY_Panel_
DYS576
DYS389I
DYS635
16
13
DYS460
DYS458
DYS19
11
17
14
DYS456
DYS390
15
23
DYS570
DYS437
17
15
DYS393
DYS439
12
12


Mark Sample for
DYS389II
DYS627
29
21
DYS448
DYS391
12
19
11
DYS392
DYS518
13
36
DYS449
14
31
DYF387S1
DYS533
35
36
12

Kinship stats over 24 aSTRs

	Combined
Parent-Child	0.000
Full Sib	0.000
Half-Sib*	0.121
1st Cousin	0.814

Two Caucasians

Same Haplotype at

Yfiler Plus Loci

Rapidly Mutating (RM) Y-STRs

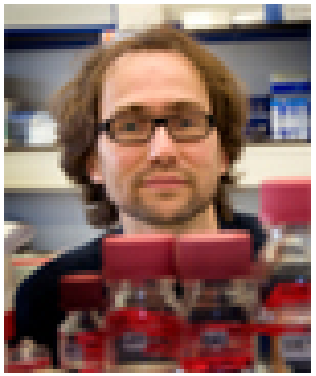
Trying to separate
close male relatives

Rapidly Mutating Y-STRs

The American Journal of Human Genetics 87, 341–353, September 10, 2010 **ARTICLE**

Mutability of Y-Chromosomal Microsatellites: Rates, Characteristics, Molecular Bases, and Forensic Implications

Kaye N. Ballantyne,[1] Miriam Goedbloed,[1] Rixun Fang,[2] Onno Schaap,[1] Oscar Lao,[1] Andreas Wollstein,[1,3] Ying Choi,[1] Kate van Duijn,[1] Mark Vermeulen,[1] Silke Brauer,[1,4] Ronny Decorte,[5] Micaela Poetsch,[6] Nicole von Wurmb-Schwark,[7] Peter de Knijff,[8] Damian Labuda,[9] Hélène Vézina,[10] Hans Knoblauch,[11] Rüdiger Lessig,[12] Lutz Roewer,[13] Rafal Ploski,[14] Tadeusz Dobosz,[15] Lotte Henke,[16] Jürgen Henke,[16] Manohar R. Furtado,[2] and Manfred Kayser[1,*]



Manfred Kayser

13 markers evaluated

Forensic Science International: Genetics 6 (2012) 208–218

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

A new future of forensic Y-chromosome analysis: Rapidly mutating Y-STRs for differentiating male relatives and paternal lineages

Kaye N. Ballantyne [a,1,2], Victoria Keerl [a,1,3], Andreas Wollstein [a,b], Ying Choi [a], Sofia B. Zuniga [c], Arwin Ralf [a], Mark Vermeulen [a], Peter de Knijff [c], Manfred Kayser [a,*]

[a] *Department of Forensic Molecular Biology, Erasmus MC University Medical Center Rotterdam, 3000 CA Rotterdam, The Netherlands*
[b] *Cologne Center for Genomics, University of Cologne, D-50674 Cologne, Germany*
[c] *Department of Human Genetics, Leiden University Medical Center, 2300 RC Leiden, The Netherlands*

Using Y-STRs with a higher mutation rate, father-son and brother pairs can sometimes be distinguished

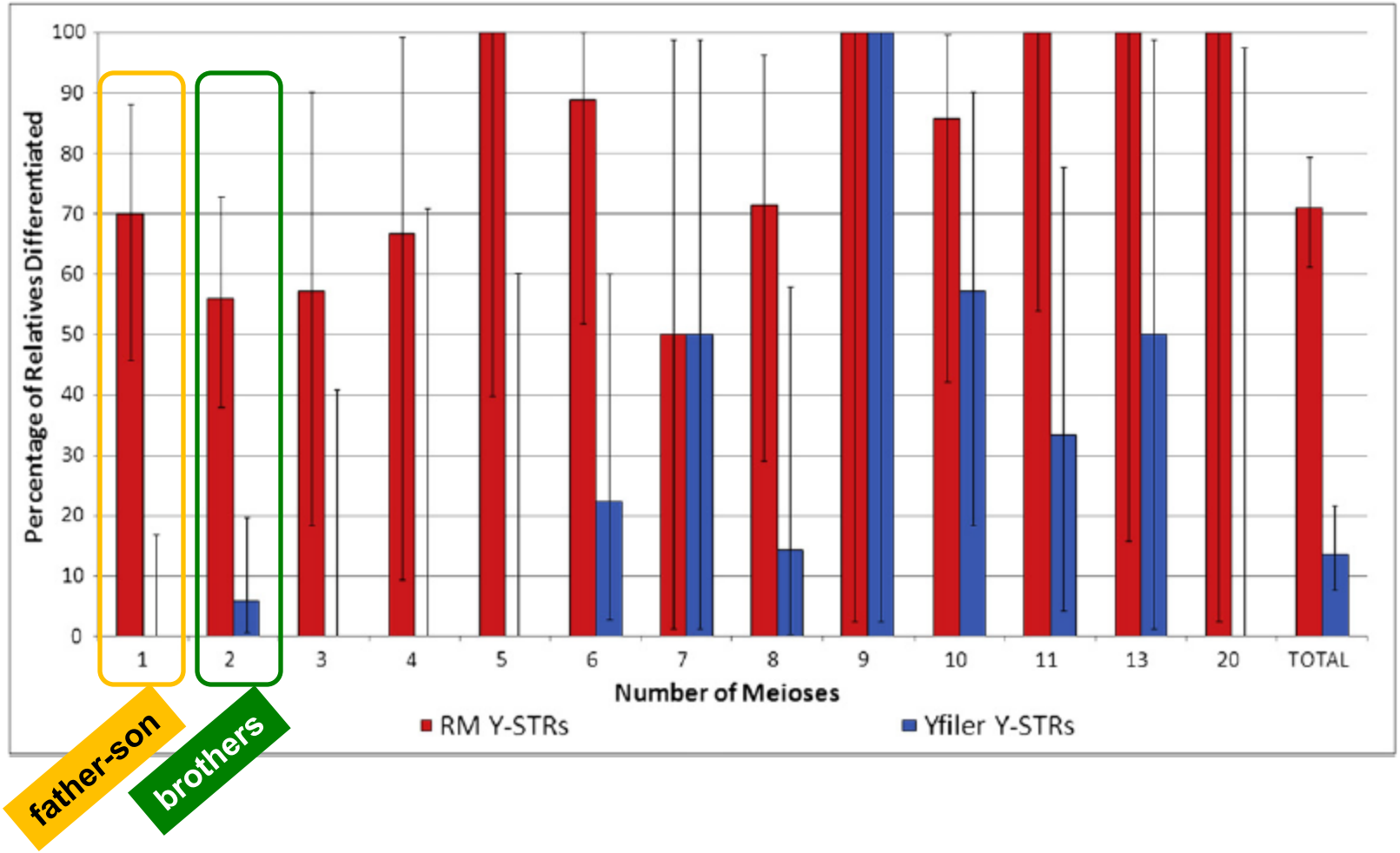


Figure 4. Male Relative Differentiation with Newly Identified 13 RM Y-STRs and Commonly Used 17 Yfiler Y-STRs
Results from differentiating between male relatives from analyzing 103 pairs from 80 male pedigrees, sorted according to the number of generations separating pedigree members, based on 13 RM Y-STRs (in red) and 17 Yfiler Y-STRs (in blue). Error bars represent 95% binomial confidence intervals. Note that these samples are independent from the father-son pairs initially used to establish the Y-STR mutation rates.

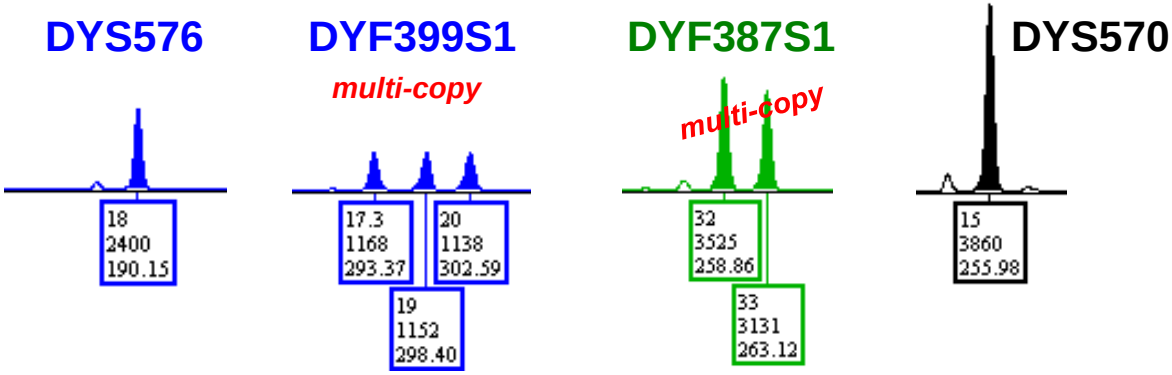
The American Journal of Human Genetics 87, 341–353, September 10, 2010

K. Ballantyne et al. 2010; K. Ballantyne et al. 2012

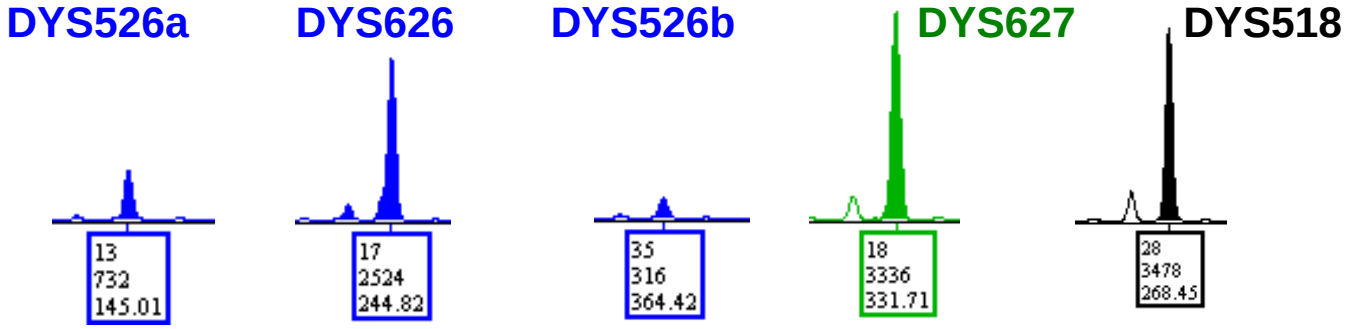
Rapidly Mutating (RM) Y-STRs

NIST supplied data from 1,296 U.S. samples (634 population + 331 father/son pairs)
to RM Y-STR Study Group led by Manfred Kayser

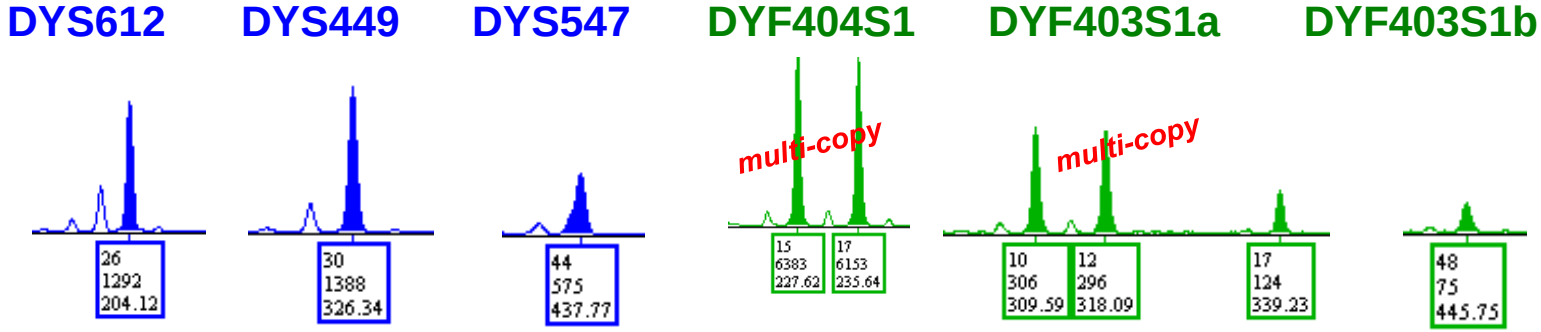
RM Y-STR
Multiplex 1



RM Y-STR
Multiplex 2



RM Y-STR
Multiplex 3

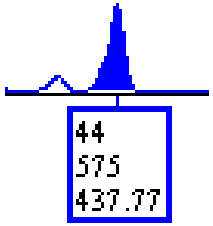


Why do these markers mutate “rapidly”?

Markers in Yfiler Plus

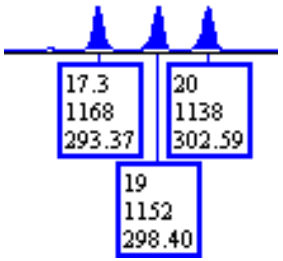
Locus	(average mutation rate)
DYS449	(1.2%)
DYS518	(1.8%)
DYS547	(2.4%)
DYS570	(1.2%)
DYS576	(1.4%)
DYS612	(1.4%)
DYS626	(1.2%)
DYS627	(1.2%)
DYF387S1	(1.6%)
DYF399S1	(7.7%)
DYF403S1 a/b	(3.1/1.2%)
DYF404S1	(1.3%)
DYS526 a/b	(1.3%)

“Large” number of repeats



DYS547

Multi-copy Markers



DYF399S1

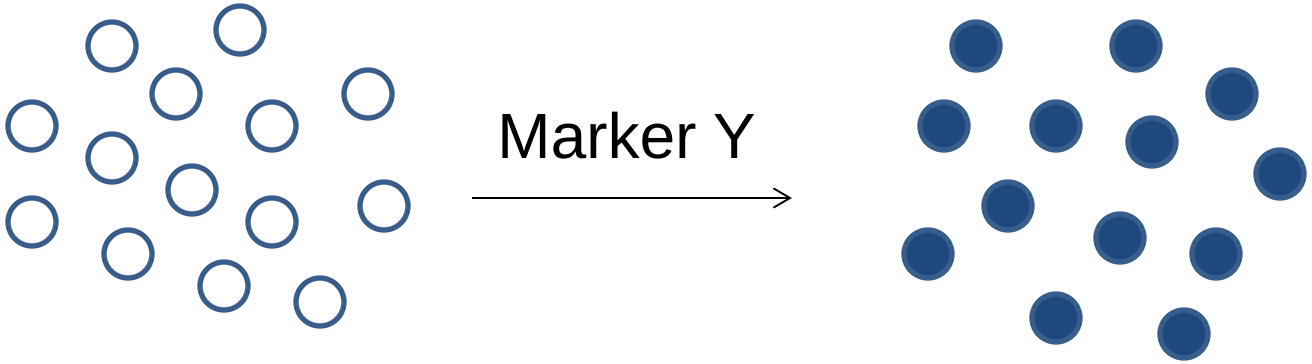
DYS458 (0.64%) is highest in Yfiler loci where average is ~0.2%

Gene Diversity

- is a measure of the uniqueness of a particular marker in a given population

$$GD = (1 - \sum_i x_i^2)$$

Relative frequency of each allele

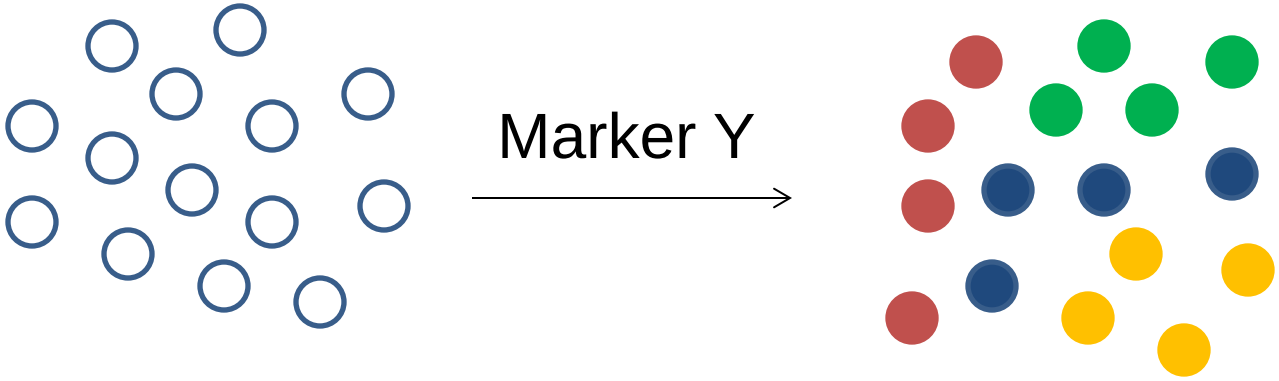


1 type = 100%

N = 100

$$GD = \underbrace{\left(1 - \sum_i x_i^2\right)}_{0}$$

$$DC = 1/100 = 0.01$$

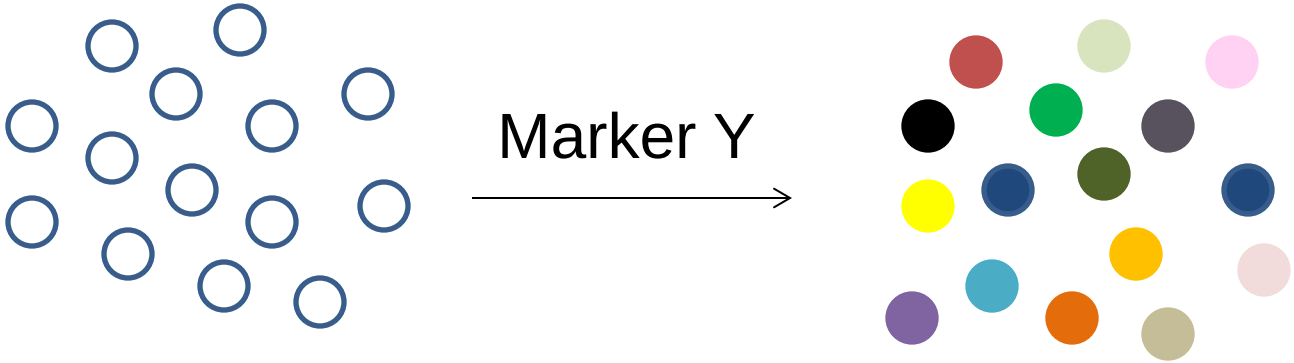


4 types = 25%

N = 100

$$GD = \underbrace{\left(1 - \sum_i x_i^2\right)}_{0.75}$$

DC = 4/100 = 0.04



100 types = 0%

N = 100

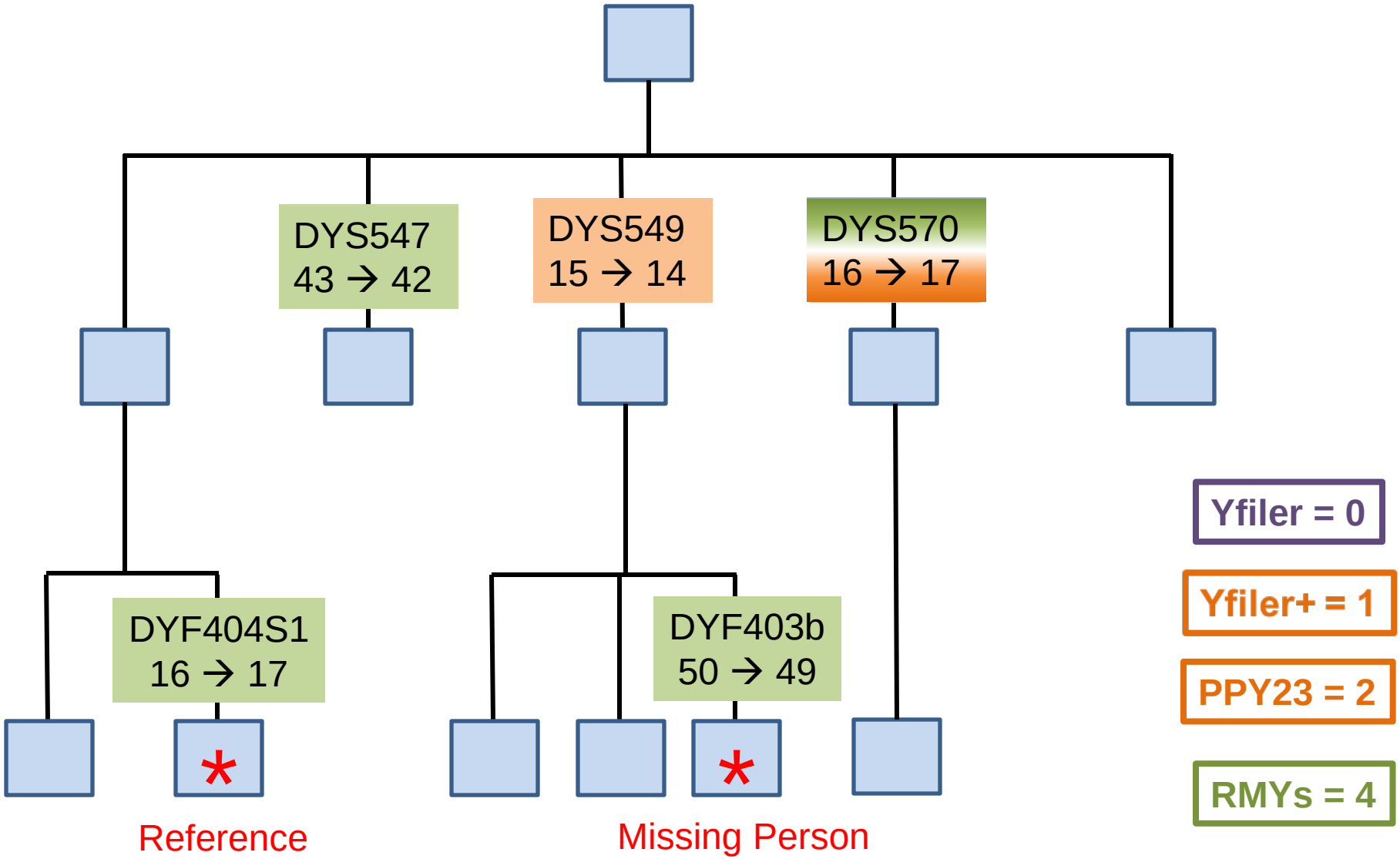
$$GD = \underbrace{\left(1 - \sum_i x_i^2\right)}_{0.99}$$

DC = 100/100 = 1.0

Gene Diversity of the YFP Markers

Locus	Gene Diversity
DYF387S1a/b	0.919
DYS385a/b	0.919
DYS627	0.8584
DYS449	0.8315
DYS481	0.82
DYS518	0.8196
DYS576	0.7954
DYS570	0.7852
DYS458	0.7671
DYS390	0.7645
DYS635	0.7457
DYS389II	0.7375
DYS448	0.7202
DYS456	0.7015
DYS438	0.693
DYS19	0.6681
DYS439	0.6533
DYS533	0.6372
DYS437	0.6305
GATA_H4	0.6026
DYS392	0.6001
DYS460	0.5736
DYS389I	0.548
DYS391	0.5352
DYS393	0.4749

Y-STR mutations in a Paternal Lineage



Interpretational Issues

- We will need to move away from simply "excluding" based upon a number of discordant markers.
- A Likelihood Ratio can provide weight to the evidence based upon competing propositions.
- This will require information on the haplotype frequency and mutation rate data.

Relating two deep-rooted pedigrees from Central Germany by high-resolution Y-STR haplotyping

Manfred Kayser [a,*], Mark Vermeulen [a,b], Hans Knoblauch [c], Herbert Schuster [d], Michael Krawczak [e], Lutz Roewer [f]

Forensic Science International: Genetics 1 (2007) 125–128.

Thank you for your attention

Acknowledgments: NIJ & FBI Funding

Contact Information

Becky Hill
Research Biologist
becky.hill@nist.gov
301-975-4275

Mike Coble
Research Biologist
michael.coble@nist.gov
301-975-4330

Our team publications and presentations are available at:
http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm