Beyond the STRs:
A Comprehensive View of Current Forensic DNA Markers Characterized in the **PCR-Based DNA Profiling Standard** **SRM 2391d**

Carolyn (Becky) Steffen, Kevin Kiesler, Lisa Borsuk, and Peter Vallone
National Institute of Standards and Technology
September 2, 2017

Development of the Next PCR-Based DNA Profiling Standard

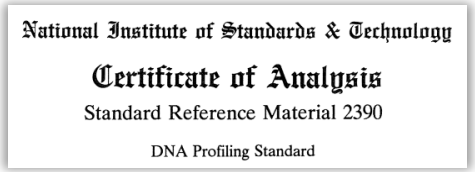
- As a successor to SRM 2391c
  - Inventory may be depleted by late 2018
  - Develop SRM 2391d now to ensure availability when needed

- **Next Generation Sequencing** will be used for certification in addition to **Capillary Electrophoresis** testing
  - Length- and sequence-based genotypes will be provided
  - Include information values for all commercially available forensic markers ***beyond the STRs***

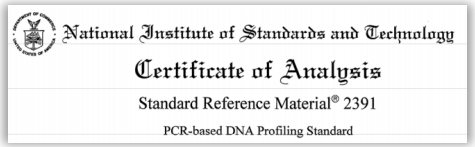
Goal: SRM 2391d will be the most comprehensive NIST forensic SRM to date

NIST Forensic DNA SRMs
Historical Perspective: Past, Present, Future

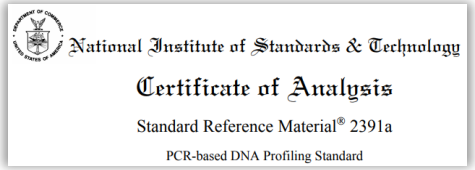
Past


National Institute of Standards & Technology
Certificate of Analysis
Standard Reference Material 2390
DNA Profiling Standard

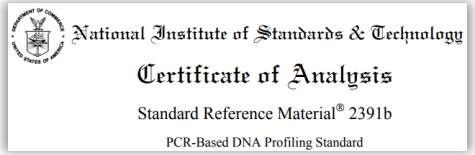
RFLP Testing & DNA Probes (1990)


National Institute of Standards and Technology
Certificate of Analysis
Standard Reference Material® 2391
PCR-based DNA Profiling Standard

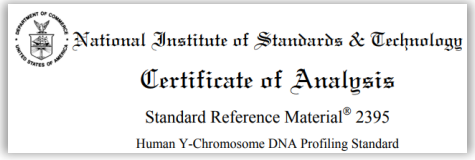
PCR-Based Testing (1995)
- VNTR, Dot Blot
- STR typing (updated 1998)


National Institute of Standards & Technology
Certificate of Analysis
Standard Reference Material® 2391a
PCR-based DNA Profiling Standard

PCR-Based Testing (2000)
- Focus on STR typing
- VNTR, Dot Blot

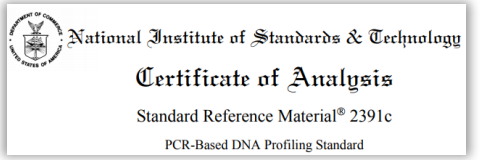

National Institute of Standards & Technology
Certificate of Analysis
Standard Reference Material® 2391b
PCR-Based DNA Profiling Standard

PCR-Based Testing (2003)
- Autosomal STR loci
- More STR loci added (updated 2008)


National Institute of Standards & Technology
Certificate of Analysis
Standard Reference Material® 2395
Human Y-Chromosome DNA Profiling Standard

PCR-Based Y-STR Testing (2003)
- Y-STR loci
- More Y-STR loci added (updated 2008)

Present


National Institute of Standards & Technology
Certificate of Analysis
Standard Reference Material® 2391c
PCR-Based DNA Profiling Standard

PCR-Based STR Testing (2011)
- Autosomal and Y-STR loci
- More autosomal and Y-STR loci, X-STR loci, and Indels added (updated 2015)
- Identity and Ancestry SNPs, and Y-Indel added (updated 2017)

How will SRM 2391d values be assigned?

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

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

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

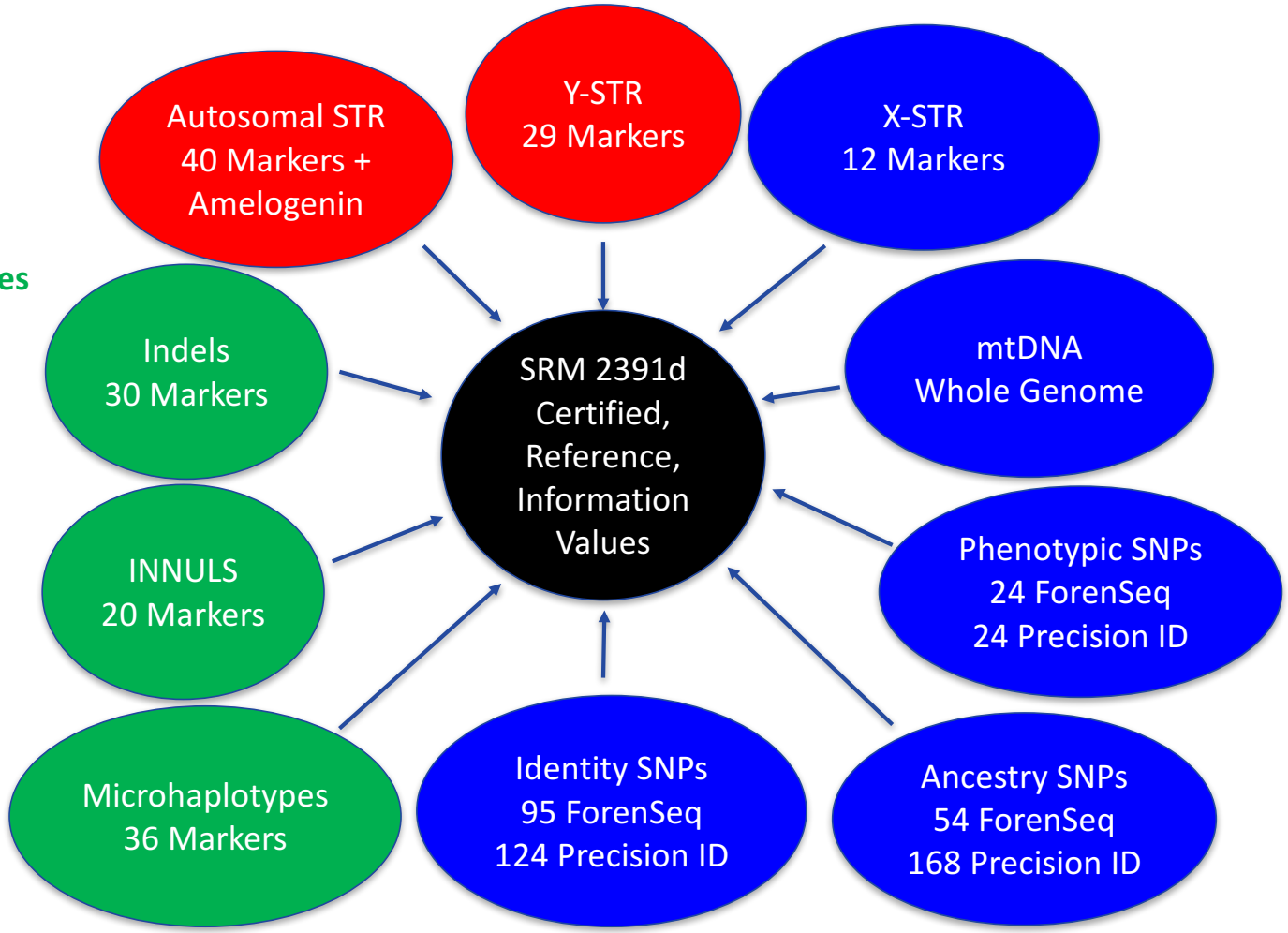
***Fit for purpose*; not all sources of uncertainty have been examined**

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

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

SRM 2391d: Forensic Markers Planned to be Included

- **Certified Values**
- **Reference Values**
- **Information Values**



Which Autosomal STR Markers will have Certified Values?

Autosomal STR Marker List	MiniFiler	Identifiler	Identifiler Plus	NGM	NGM SElect	NGM Detect	Verifiler Express	GlobalFiler	PP S5	PP CS7	PP 16	PP 16 HS	PP 18D	PP 21	PP ESX 17 Fast	PP ESI 17	PP ESI 17 Pro	PP ESI 17 Fast	PP Fusion	PP Fusion 6C	ESSplex SE	24plex QS	ForenSeq	Precision ID GF	Precision ID Mixture ID GF	PowerSeq 46GY	CODIS 20	European Standard Set	Certified Value	Reference Value	Information Value
D1S1656																													X		
D1S1677																															X
D4S2408																														X	

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

Which Y-STR Markers will have Certified Values?

Y-STR Markers
ThermoFisher CE STR kits
Promega CE STR kits
Qiagen Investigator CE STR kits
Illumina NGS kit
ThermoFisher NGS kits
Promega NGS kits

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

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

Screening and Planning Phase of Development

- **Sample format**:
  - 4 extracted DNA samples
    - 3 single source and 1 mixed sample at a 3:1 ratio (female:male)
  - 2 cell lines spotted onto FTA and 903 paper (intact cells)

Same sample format as SRM 2391c

- **Concentration of the samples will be ~1-2 ng/µL DNA for the extracted DNA and 7.5 x 10^4 cells spotted on paper**
  - The concentrations will NOT be certified values – just information values

---

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


“Beyond the STRs”

How will SRM 2391d be tested "*Beyond the STRs*"?

- **CE** will be performed with kits from two commercial companies:

Kit Provider	
Qiagen Inc. (2)	**Innogenomics (1)**
Investigator DIPplex	InnoTyper 21
Investigator Argus X-12 QS	

- **NGS** will be performed with the HID available sequencing panels/kits

Kit Provider	
Thermo Fisher (4)	**Illumina (2)**
Precision ID Ancestry Panel	ForenSeq DNA Signature Prep Kit
Precision ID Identity Panel	Nextera XT Sample Prep Kit (mtDNA)
Precision ID mtDNA Whole Genome Panel	
Ion Ampliseq DNA Phenotyping Panel	

How will SRM 2391d be tested?

- **Capillary Electrophoresis (CE)** will be performed with three different instruments:
  - 3130xL and 3500xL Genetic Analyzer (ThermoFisher)
  - Spectrum CE System (Promega) – **when available**

- **Next Generation Sequencing (NGS)** will be performed with two different instruments:
  - MiSeq FGx (Illumina)
  - Ion S5 XL (ThermoFisher)

3130xl

3500xl

Spectrum

MiSeq FGx

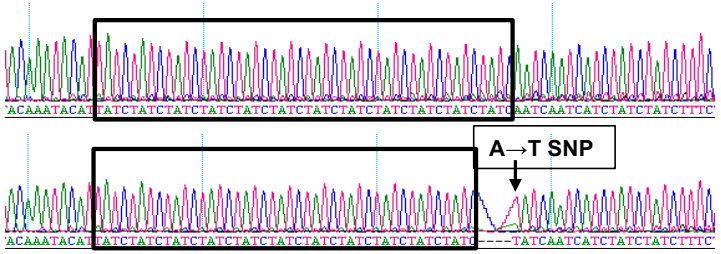
Ion S5 XL

What about Sanger Sequencing?

- There are currently **NOT** plans to Sanger Sequence every certified type for the components of SRM 2391d
  - Sanger and NGS methods were used in parallel to characterize all STR alleles for SRM 2391c
    - All results were fully concordant

- However, if there are any issues, concerns or questionable results:
  - Discordant results between kits
  - Null alleles
  - Any other ambiguities that are observed

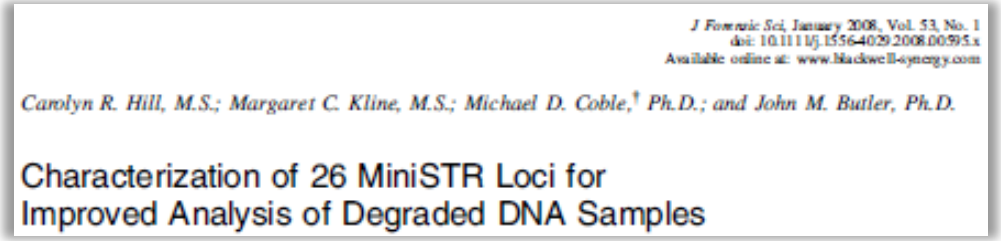
Sanger Sequencing will be used to confirm results



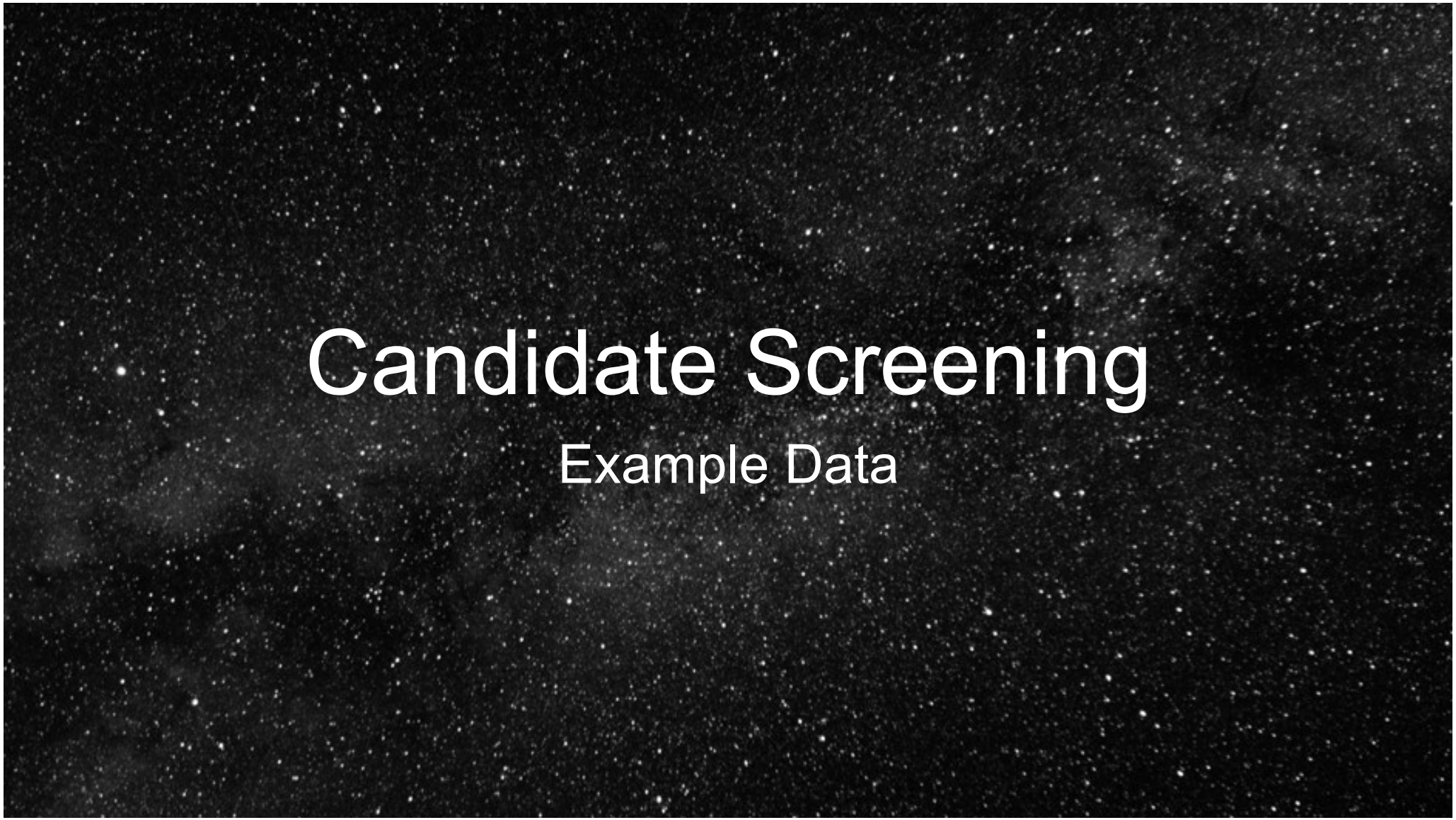
Other Forensic Markers Under Consideration

- Rapidly-Mutating Y-STR markers

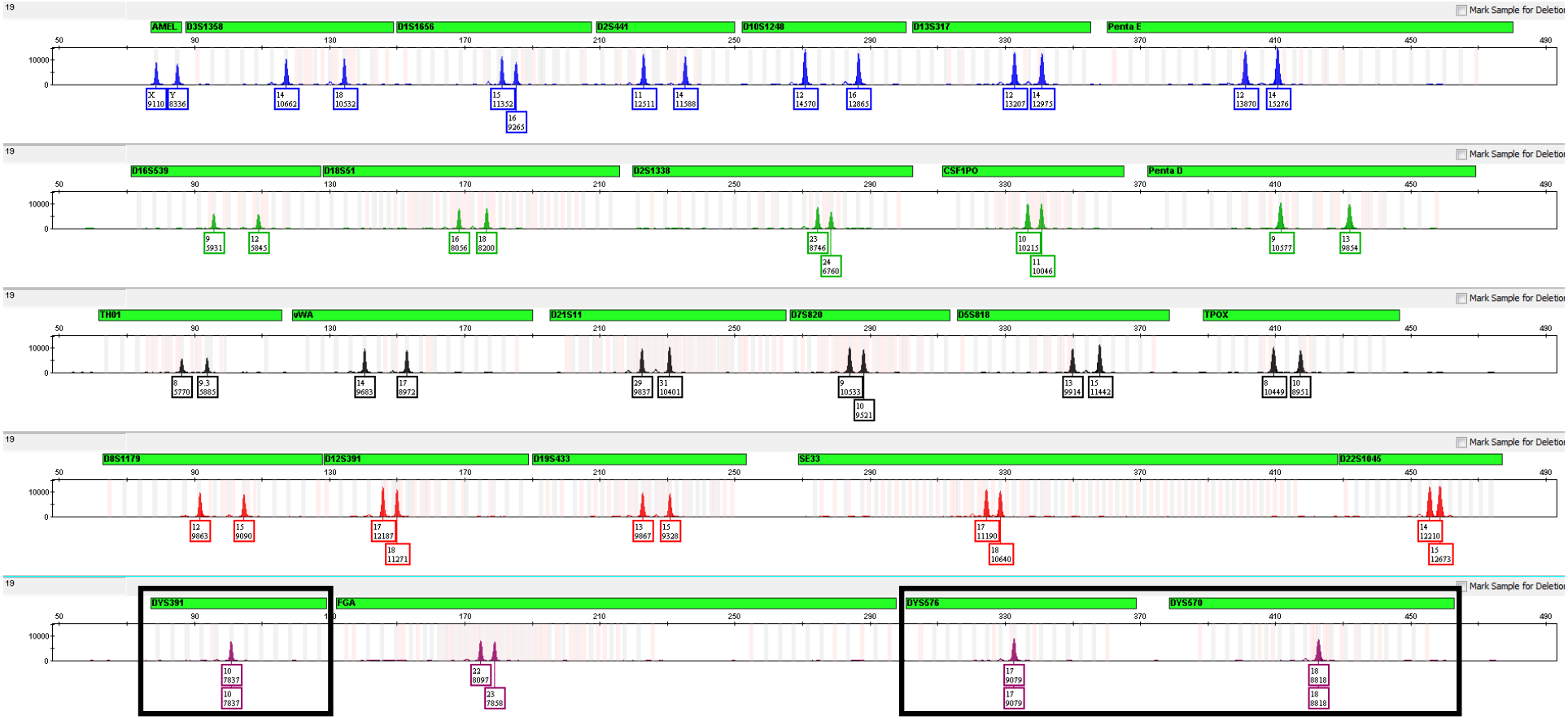
- Non-CODIS STR markers beyond commercial kits



- Others??
  - New kits, instruments and/or software yet to be released


Candidate Screening
Example Data

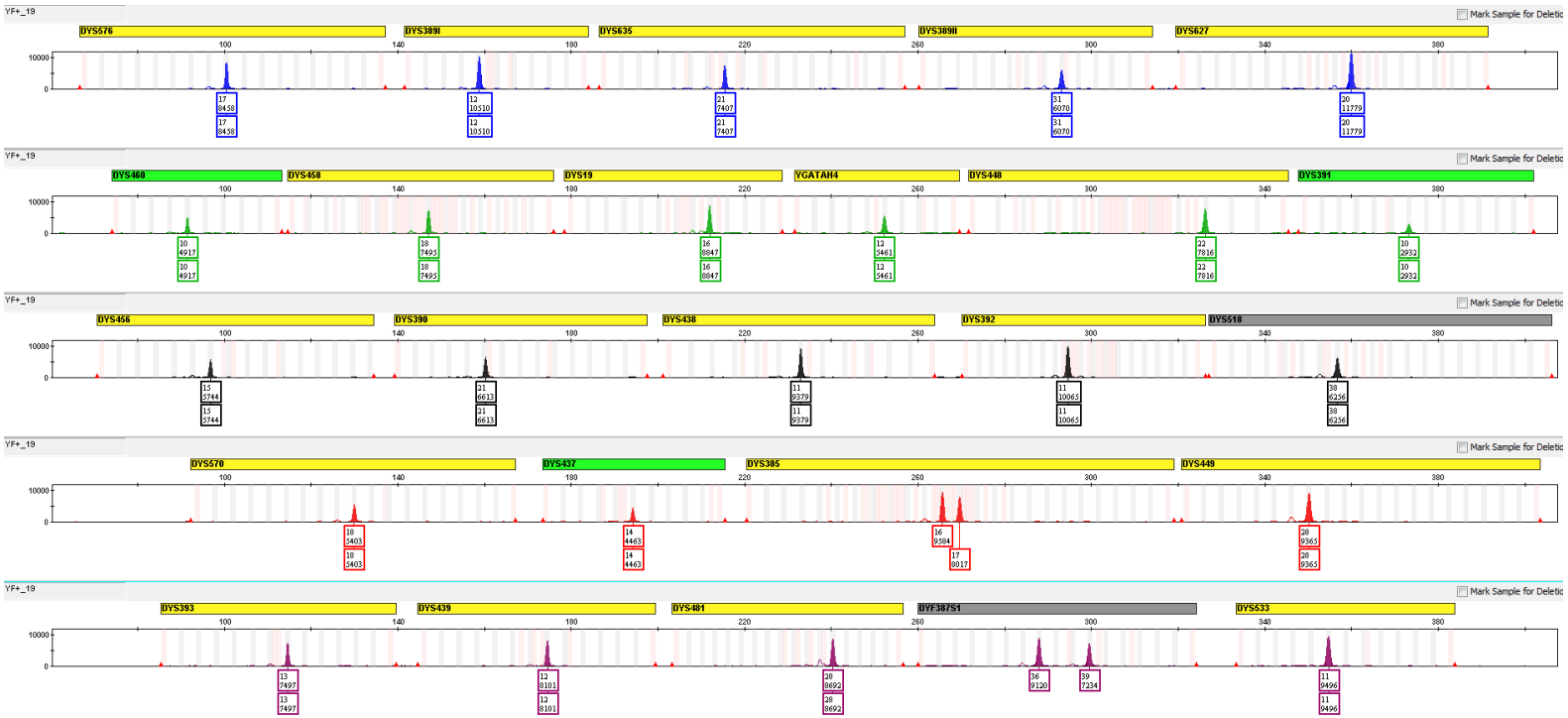
Data Collection for Sample Screening: Autosomal STR Example Candidate Sample



Fully Heterozygous with PowerPlex Fusion 6C

*Y-STR Markers

Data Collection for Sample Screening: Y-STR



YFiler Plus Profile

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

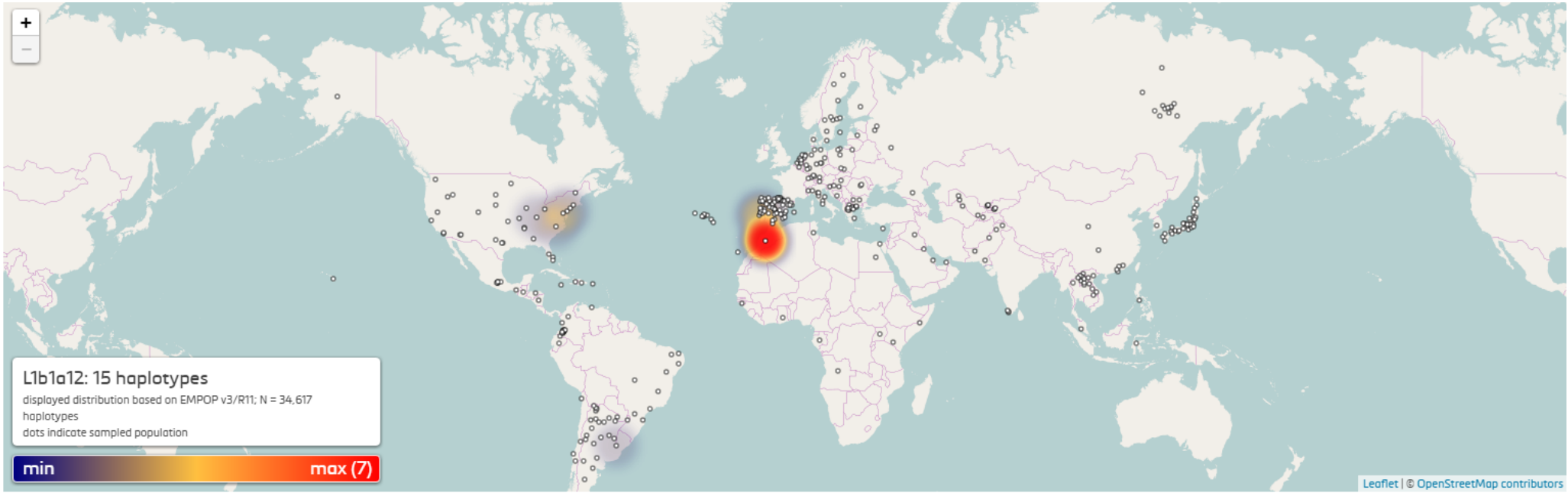
Whit Athey's Haplogroup Predictor: E1b1a
http://www.hprg.com/hapest5/hapest5a/hapest5.htm?order=num

Results Table

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

Data Collection for Sample Screening: mtDNA

Illumina mtDNA Whole Genome Sequencing protocol with Nextera XT Sample Prep Kit



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

Haplogroup	Ancestry	Match
L1b1a12	African	unique

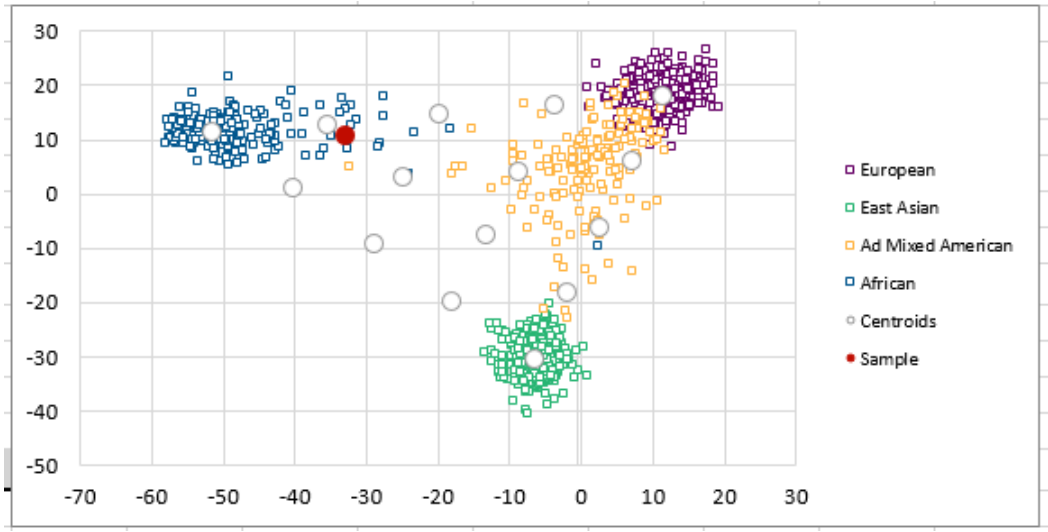
Data Collection for Sample Screening: SNPs

ForenSeq SNP Phenotype and Ancestry Estimation

Hair Color Results	
Brown	0.16
Red	0.00
Black	0.84
Blond	0.00

Eye Color Results	
Intermediate	0.00
Brown	1.00
Blue	0.00

Biogeographical Ancestry Results	
Distance to Nearest Centroid	3.36



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

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

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

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


Applications of SRM 2391d

What can you use SRM 2391d for?

- To meet the FBI Quality Assurance Standards: QAS 9.5.5

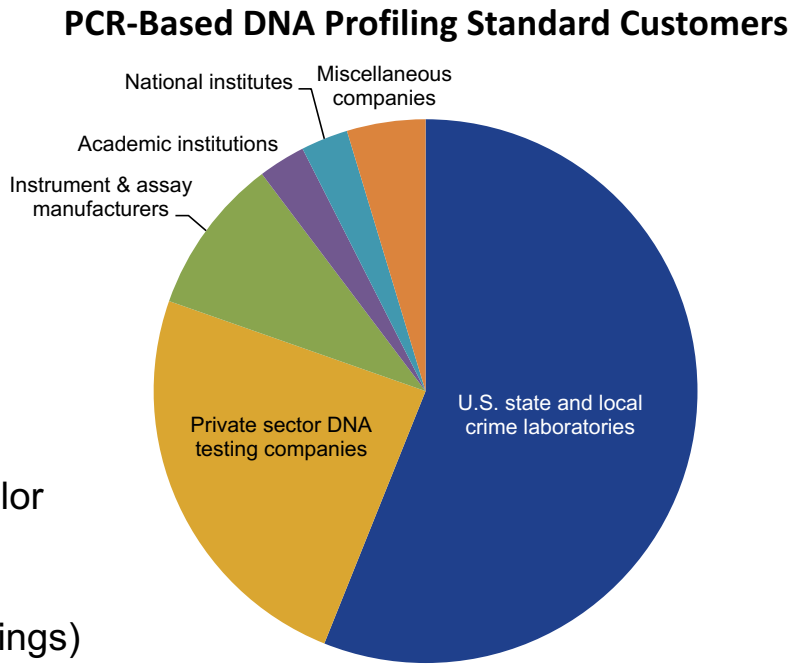
> 9.5.5 The laboratory shall check its DNA procedures annually or whenever substantial changes are made to a procedure against an appropriate and available NIST standard reference material or standard traceable to a NIST standard.

- Validation Studies: instrument, commercial kit, and software
  - Developmental and Internal Validations
  - Known, ***well-characterized*** samples for all systems commercially available

- Make NIST traceable materials (see http://ts.nist.gov/traceability/)

Support to the Forensic Community

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



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

Summary and Final Thoughts

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

Thank you for your attention!

Questions?

becky.steffen@nist.gov
1-301-975-4275

Acknowledgements

Margaret Kline
David Duewer
Michael Coble
Katherine Gettings
Erica Romsos

Disclaimer: Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Commerce. Certain commercial equipment, instruments, and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement by NIST, nor does it imply that any of the materials, instruments, or equipment identified are necessarily the best available for the purpose.

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

Funding: NIST Special Programs Office: *Forensic DNA*

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