# 16 Going on 24: GlobalFiler and the New STR Loci

Becky Hill

National Institute of Standards and Technology

Mid-Atlantic Association of Forensic Scientists

Annual Conference

State College, PA

May 20, 2014

# GlobalFiler STR Kit

Launched Friday, September 14, 2012

**Human Identification**

- GlobalFiler™ Kit
- Go Faster
- Go Further
- Go Global
- Powered by 6-Dye™
- Human Identification Home


**Introducing the world's most powerful STR kit**

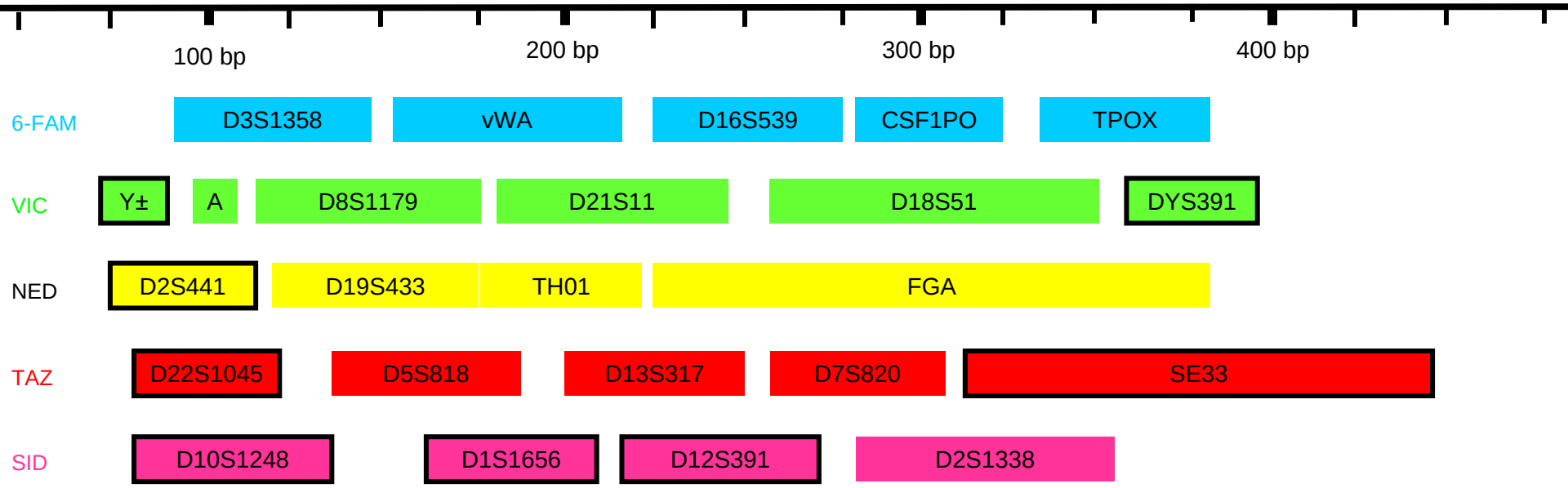
Around the world, forensic labs are being asked to do more with less. That's why the new GlobalFiler™ STR Kit combines reduced amplification time with maximum data recovery power. As part of the only fully integrated and validated forensic workflow, this breakthrough 6-dye, 24-loci technology is designed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

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

**http://www.invitrogen.com/site/us/en/home/Products-and-Services/Applications/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

Two separate kits

- GlobalFiler gives ~12 orders of magnitude improvement using the NIST 1036 data set

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

# 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*

*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 |

**Better for mixtures (more alleles seen)**

# 29 STR Loci present in STR kits

rank ordered by their variability

# 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**

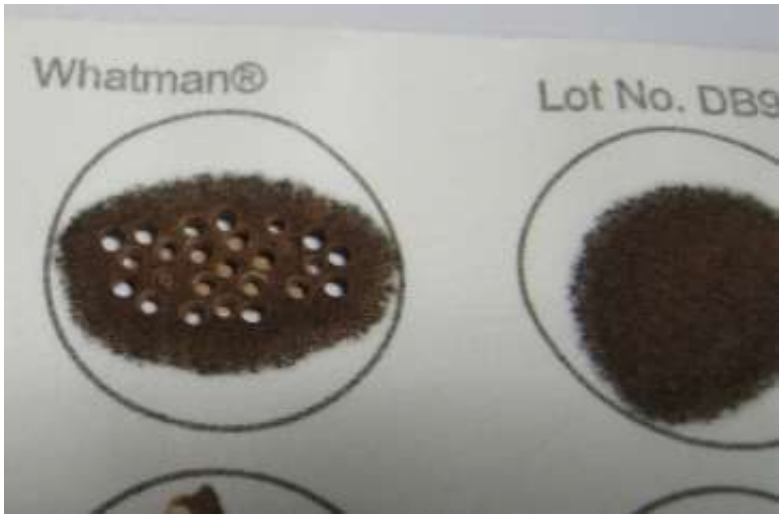
**Better for kinship (low mutation rate)**

# GlobalFiler: What has been done at NIST?

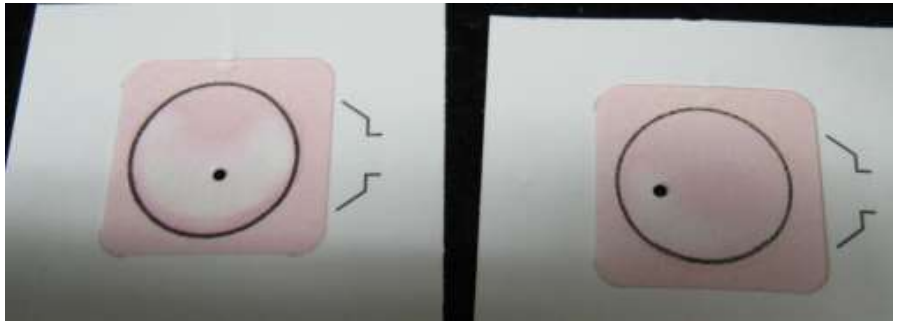
- We received a prototype GlobalFiler Express (direct amp) kit in July 2012
- All requested “VTS” experiments performed
  - Sensitivity, Reproducibility, and Performance
- 50 NIST bloodstains on FTA
- 50 NIST bloodstains on 903 (same DNA samples)
  - Prep-n-Go treatment
- 2 buccal samples collected with Whatman EasiCollect
- Degraded DNA and concordance studies

# Samples Tested

**Bloodstains on FTA paper**




**Buccal swab on FTA paper (Whatman EasiCollect)**

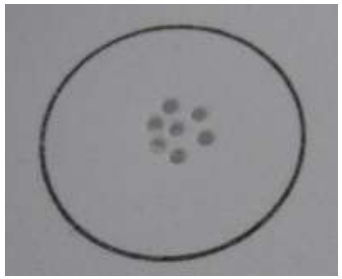


*Worked well*

*No result*
*(not enough cells collected?)*

**AB001**
**Positive control**

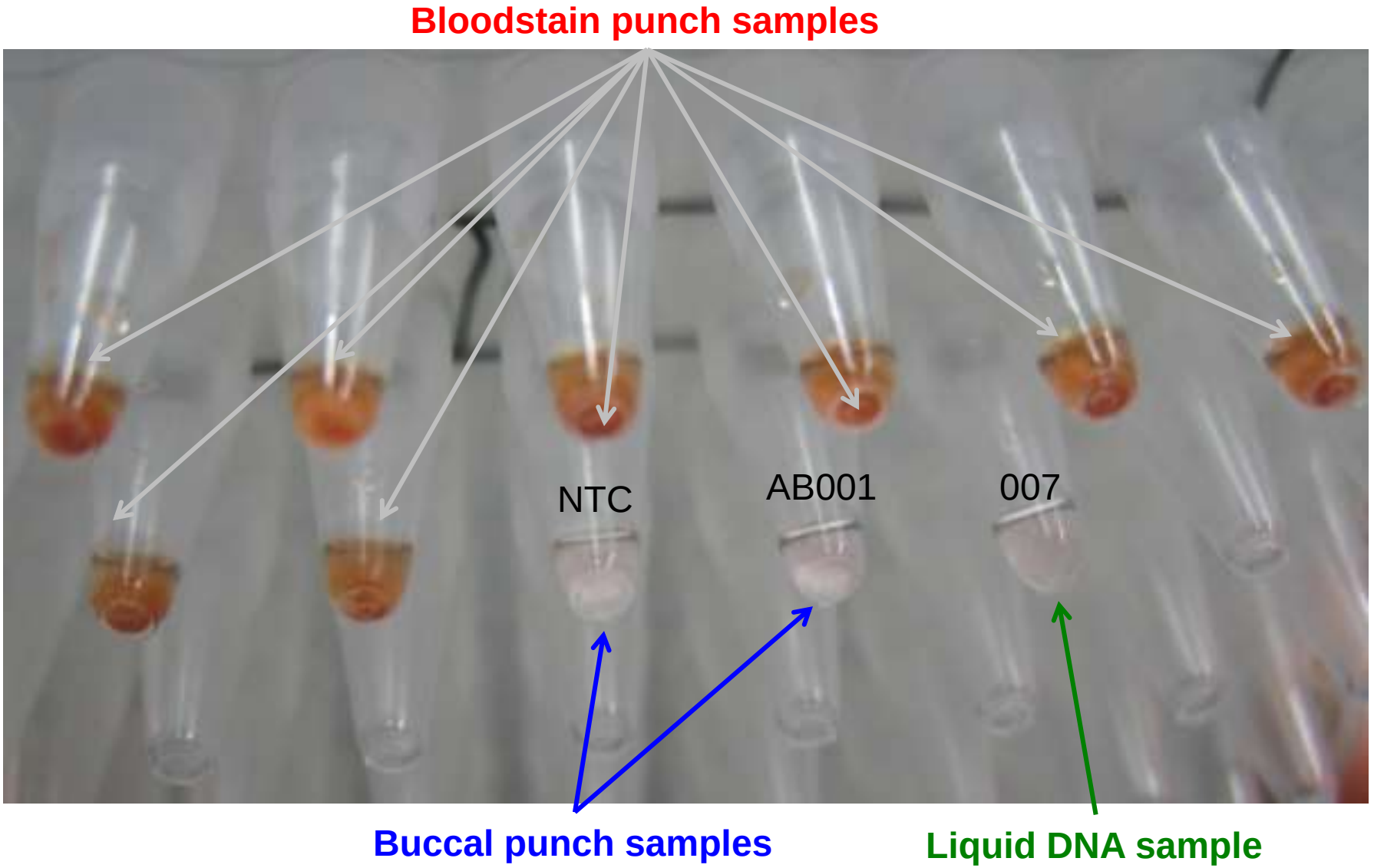

**Negative control**



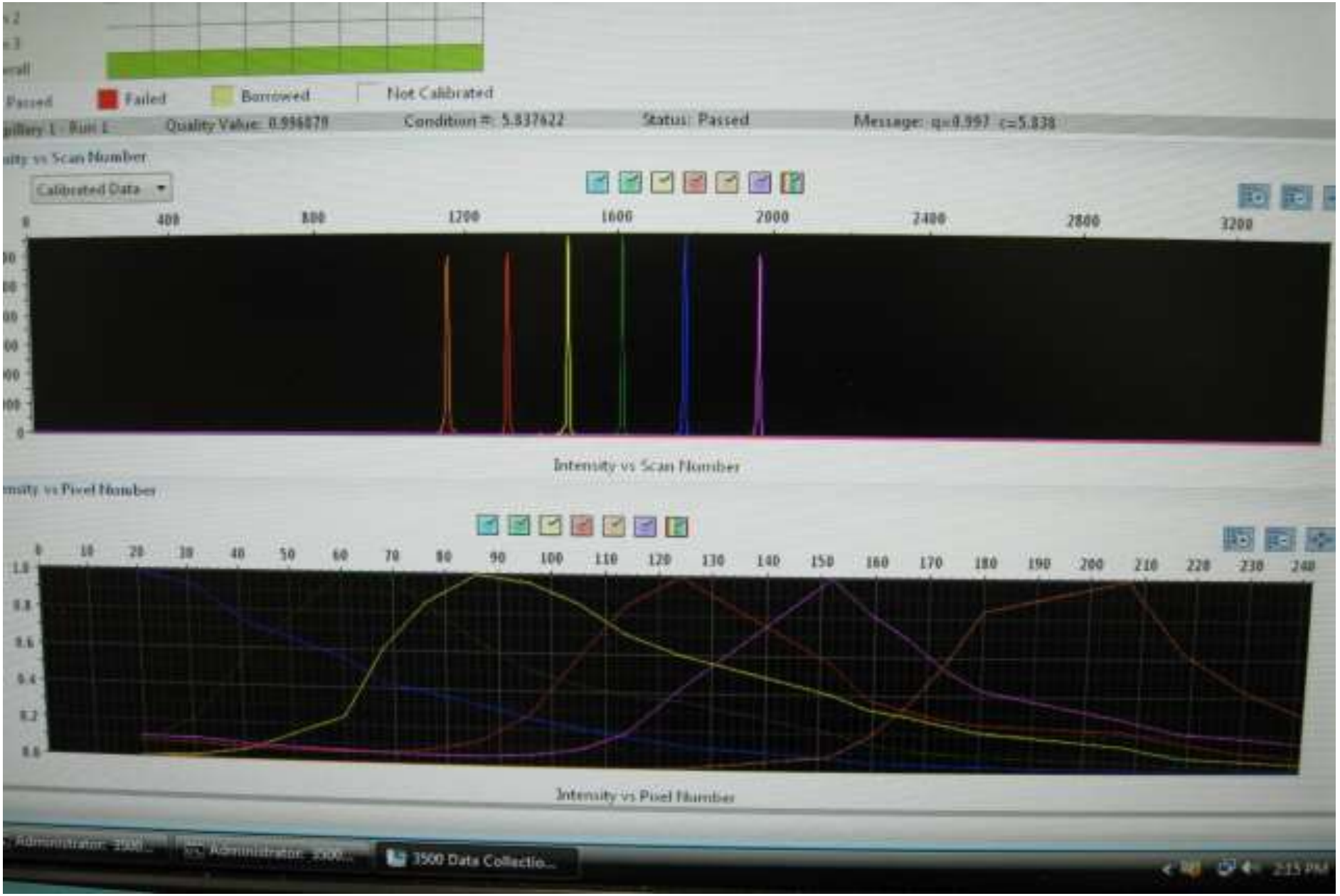
Paper punches with
no sample applied

Automated Harris puncher used

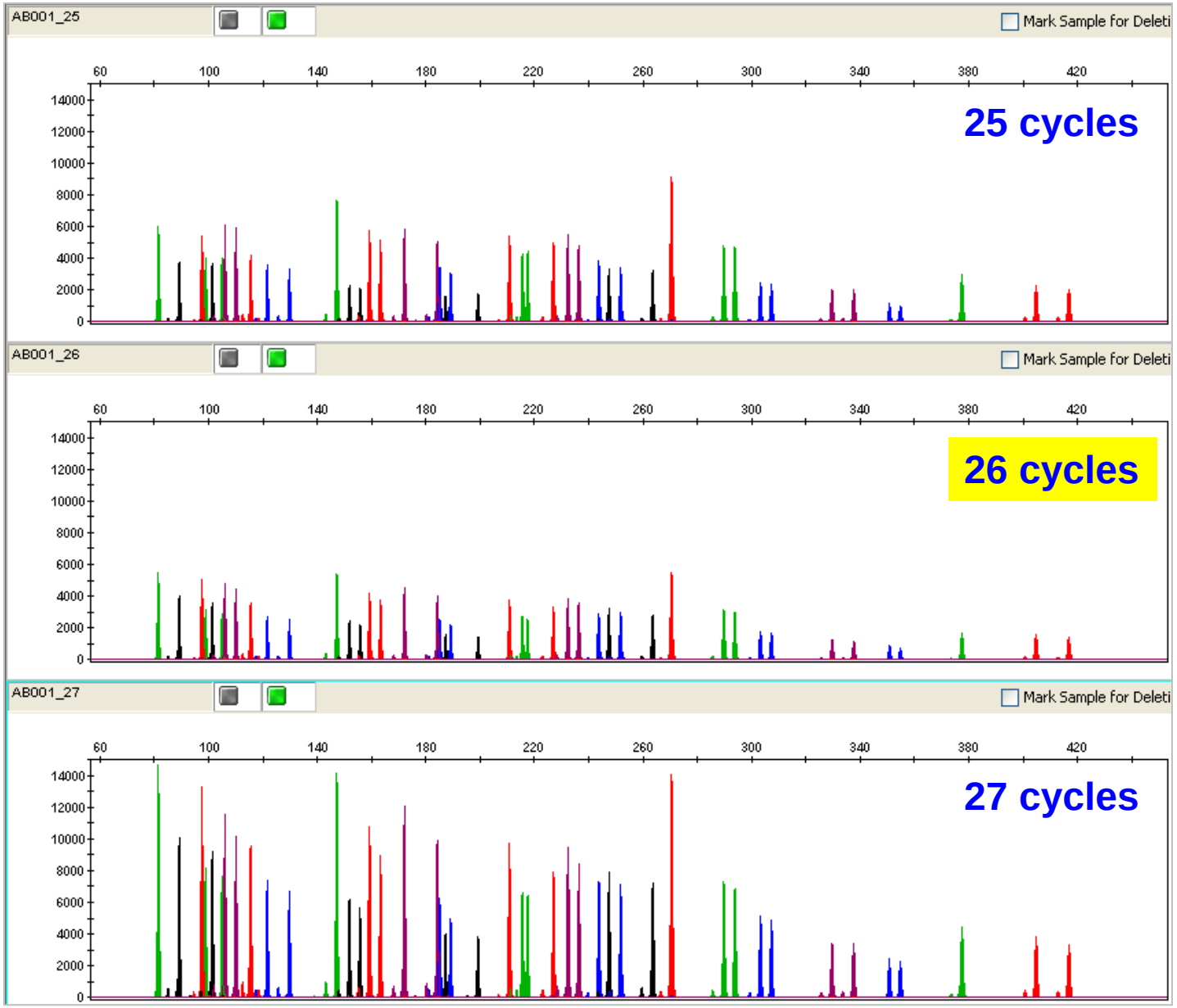
# Samples Following Direct PCR



# The J6 6-Dye Spectral Calibration Worked Well



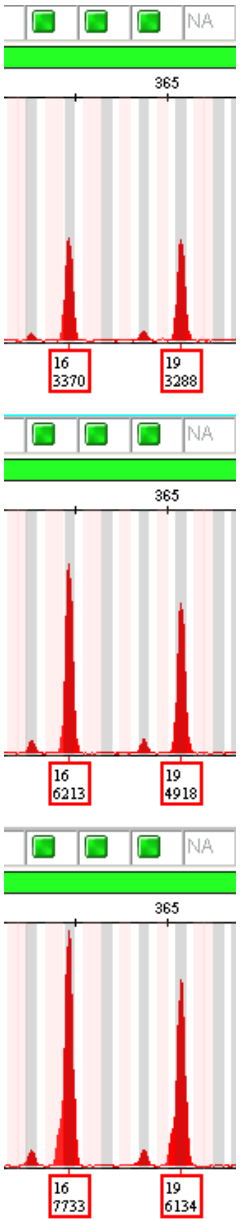
# Cycle Number Experiments




*Selected
26 cycles for all
additional studies

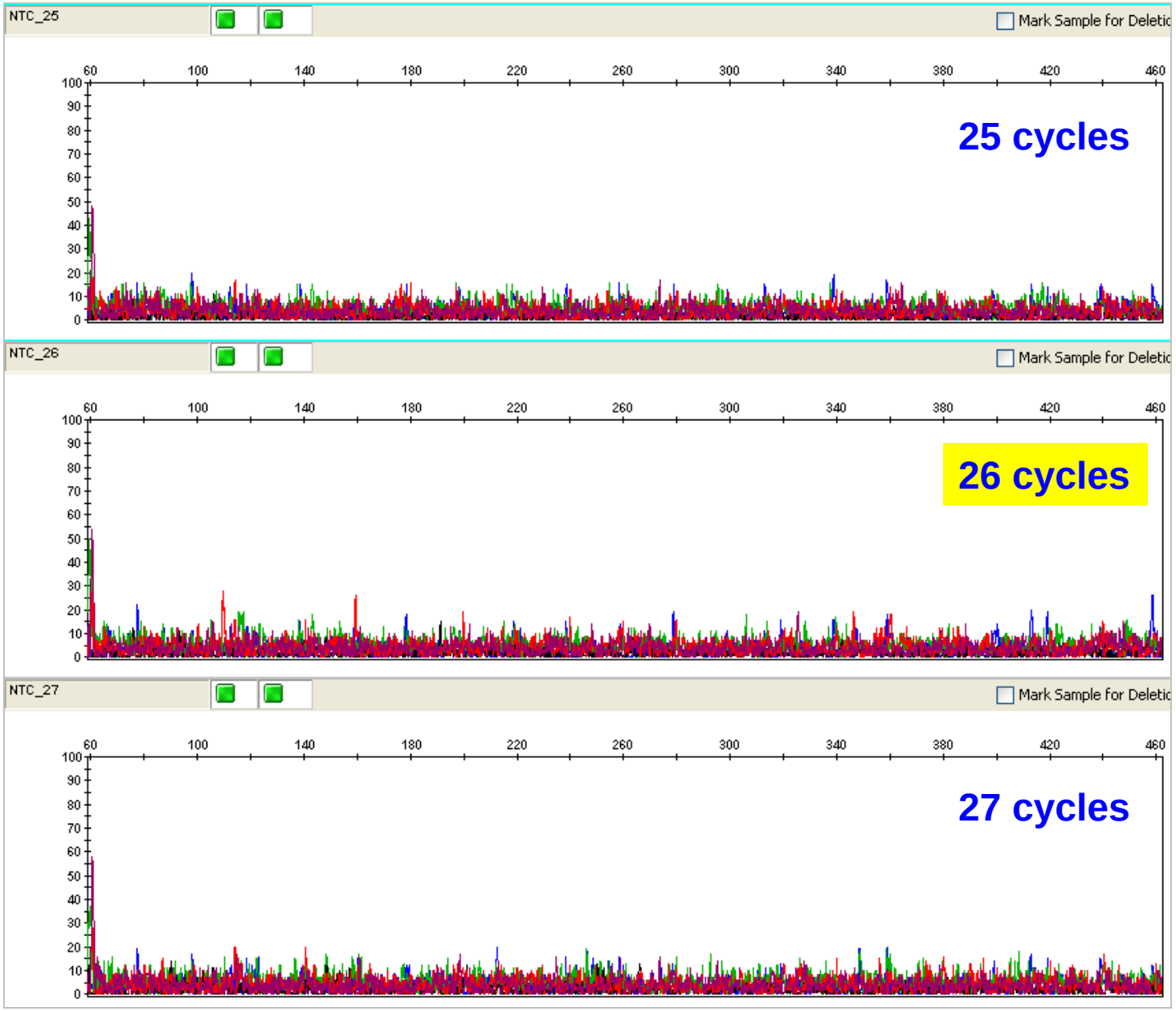

FTA_25_2
Mark Sample for Deletio
65 105 145 185 225 265 305 345 385
28000 24000 20000 16000 12000 8000 4000 0
SE33
FTA_26_2
Mark Sample for Deletio
65 105 145 185 225 265 305 345 385
28000 24000 20000 16000 12000 8000 4000 0
SE33
FTA_27_2
Mark Sample for Deletio
65 105 145 185 225 265 305 345 385
28000 24000 20000 16000 12000 8000 4000 0
SE33

## SE33


NA
365
16
3370
19
3288
NA
365
16
6213
19
4918
NA
365
16
7733
19
6134

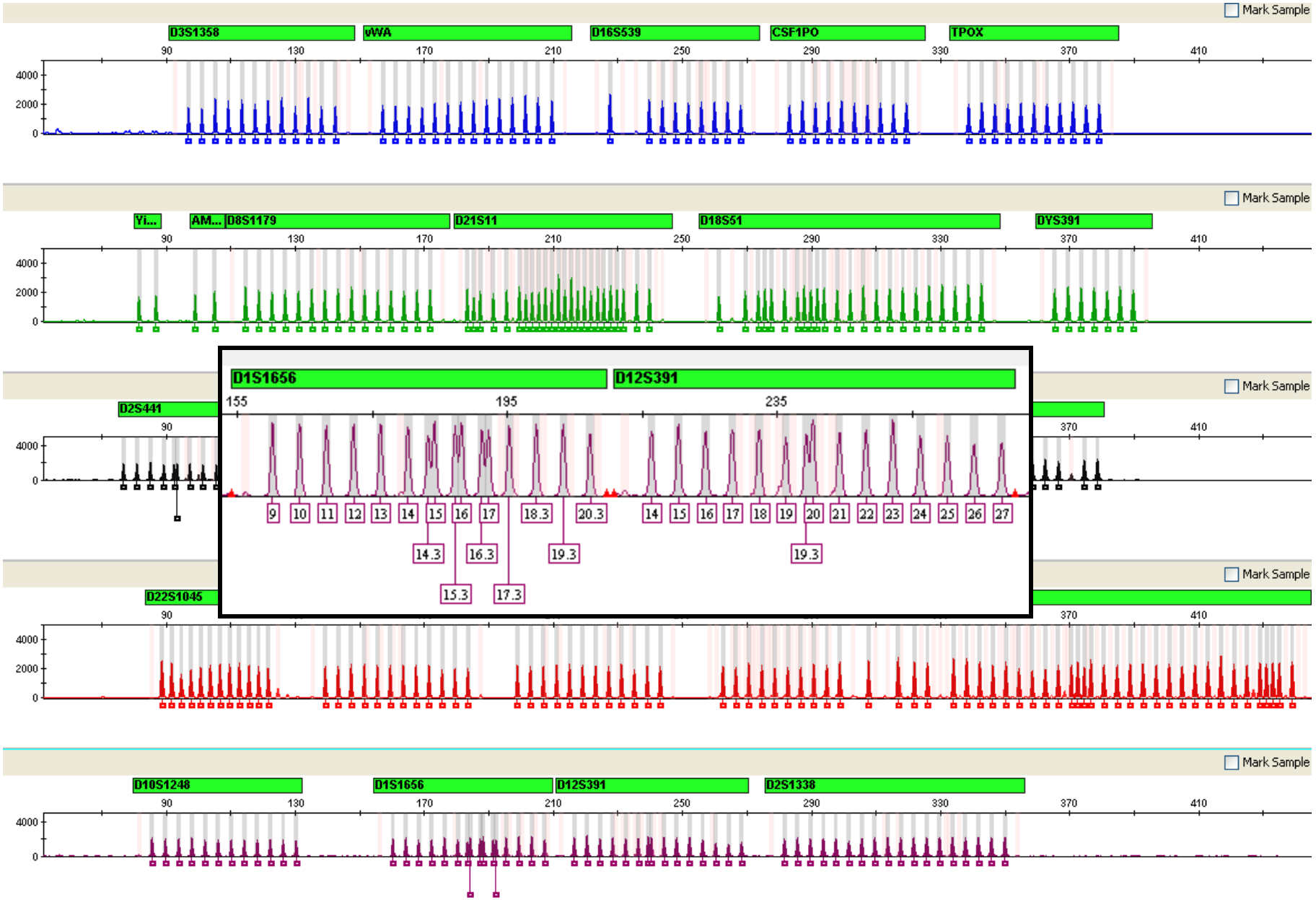

FTA_25_10
Mark Sample for Deleti
D12S391
FTA_26_10
Mark Sample for Deleti
D12S391
FTA_27_10
Mark Sample for Deleti
D12S391
-A peaks
D12S391
240
19
3912
21
3570
NA
240
19
6288
21
5670
NA
240
19
6199
21
5773

# Baseline Comparison in Negative Control

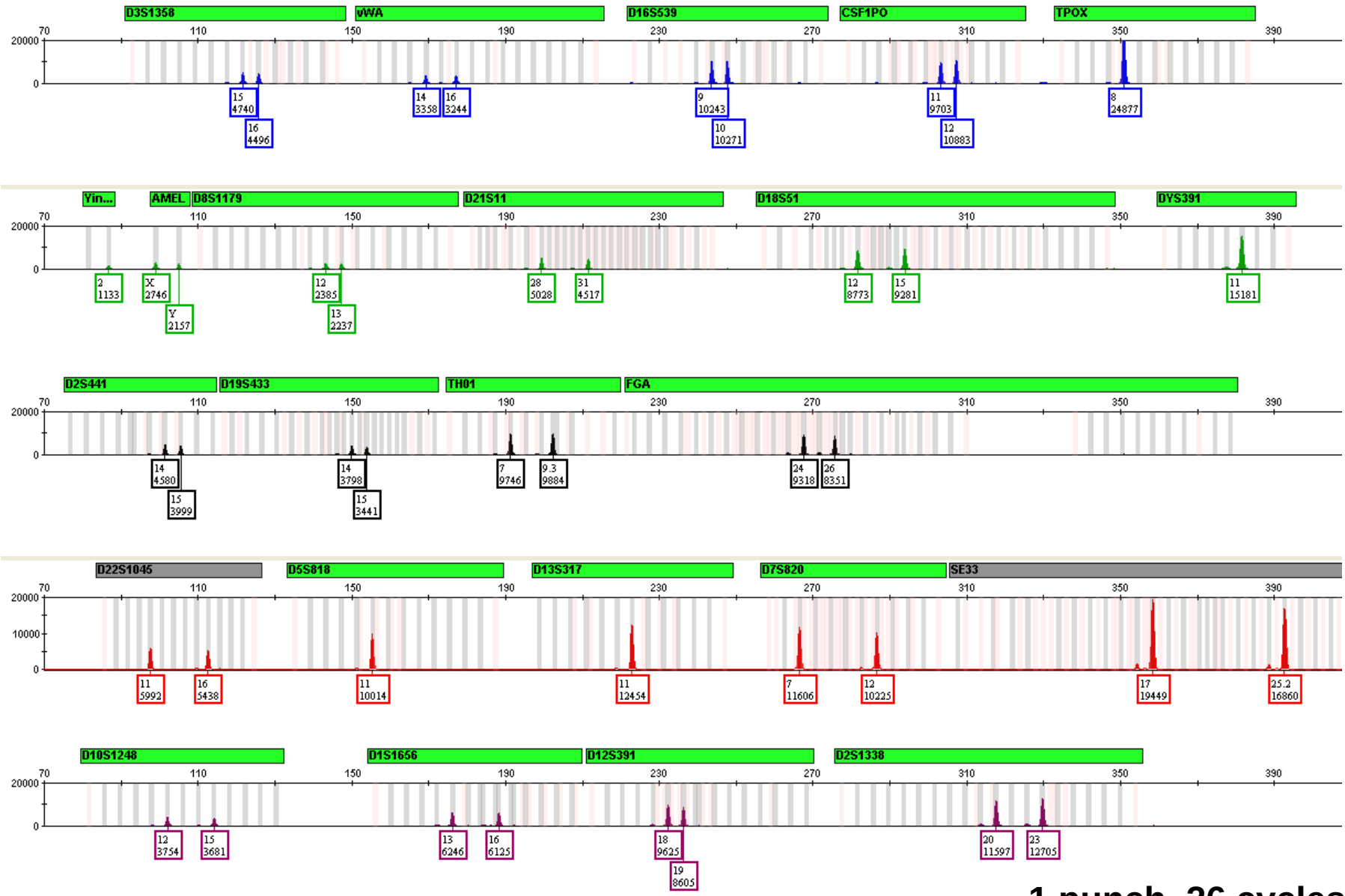


# GlobalFiler Allelic Ladder

343 alleles across these 24 loci

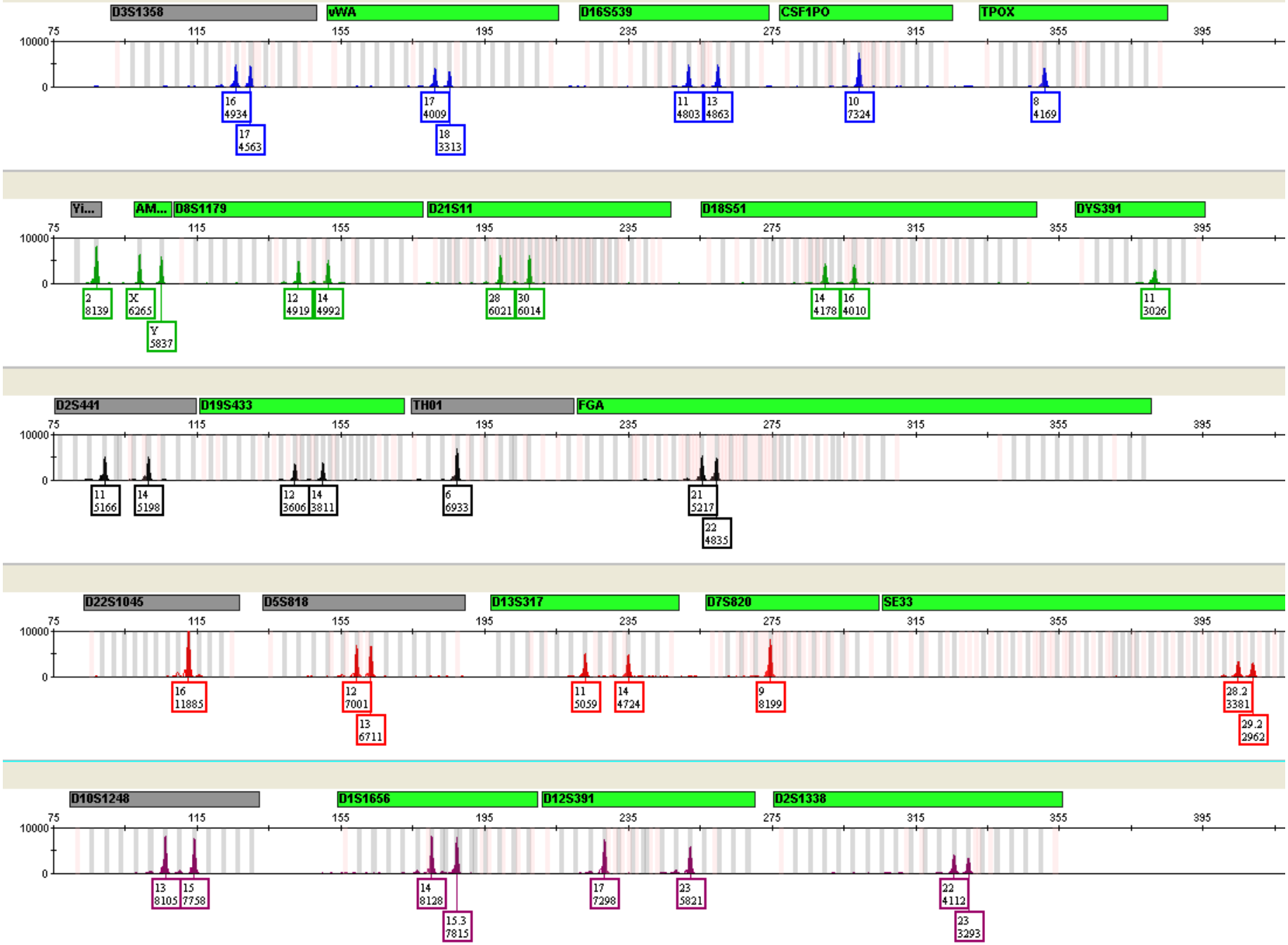


# Positive Control 007

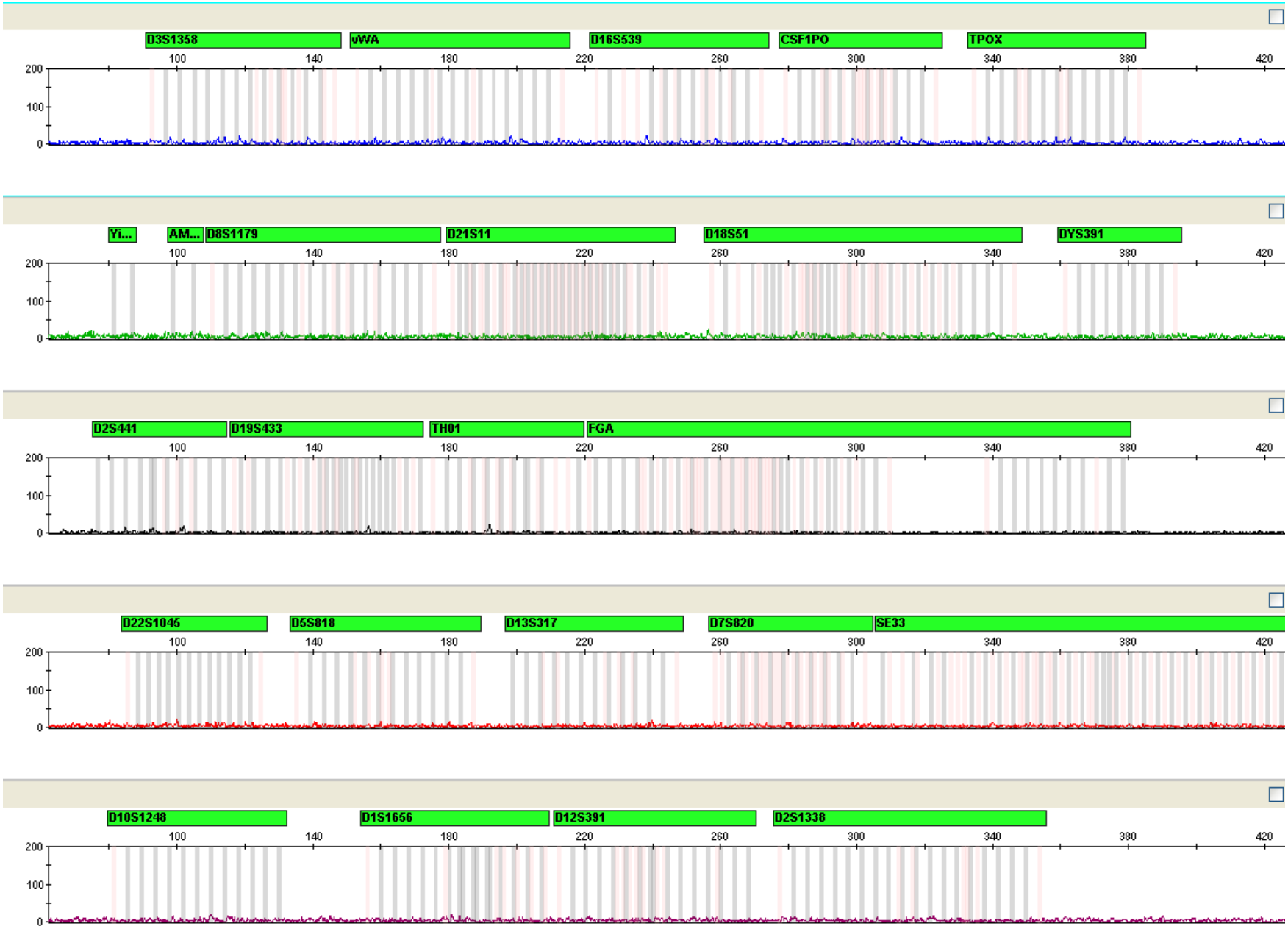



**1 punch, 26 cycles**

# Successful buccal swab (Whatman collector)

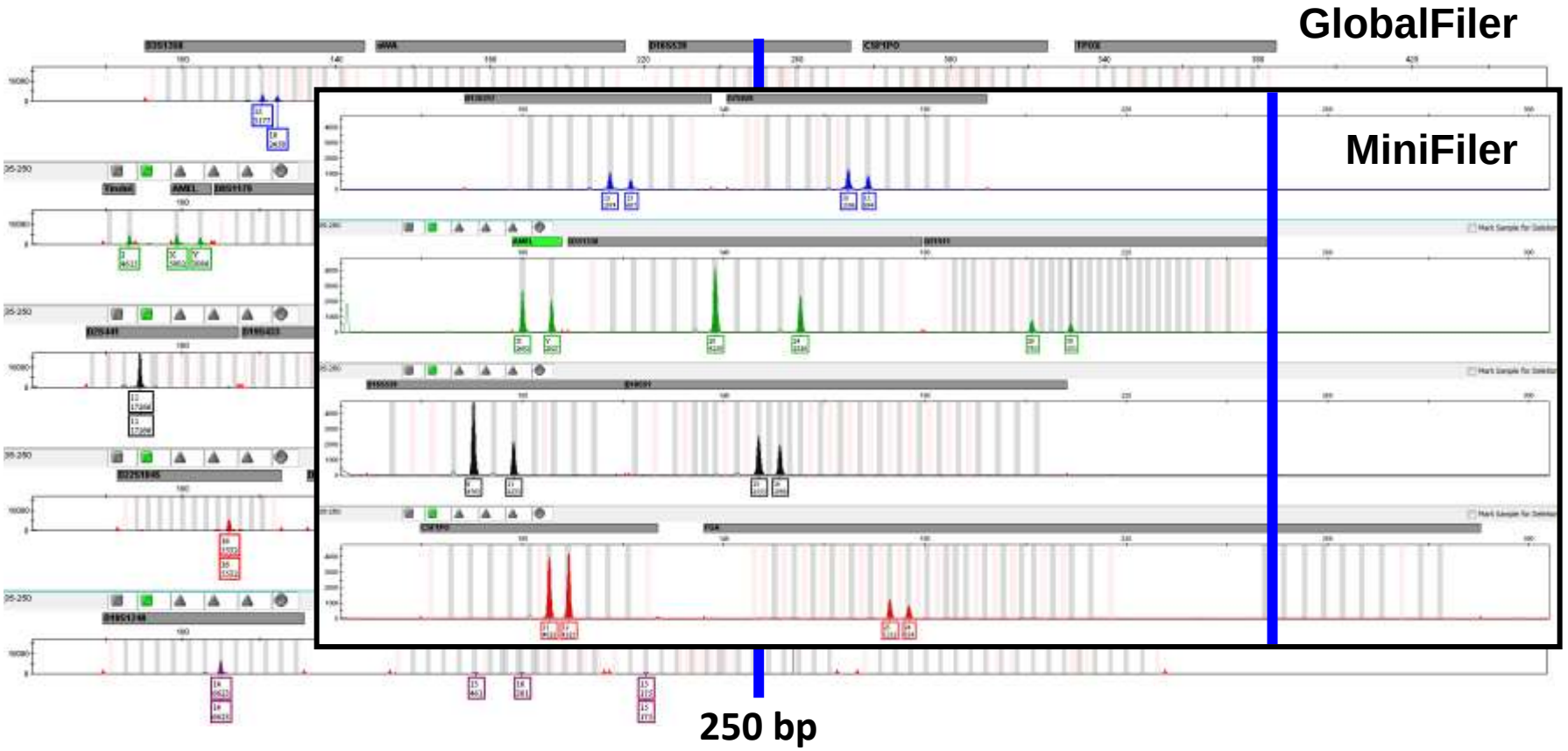


# Negative Control



# Degraded DNA Study

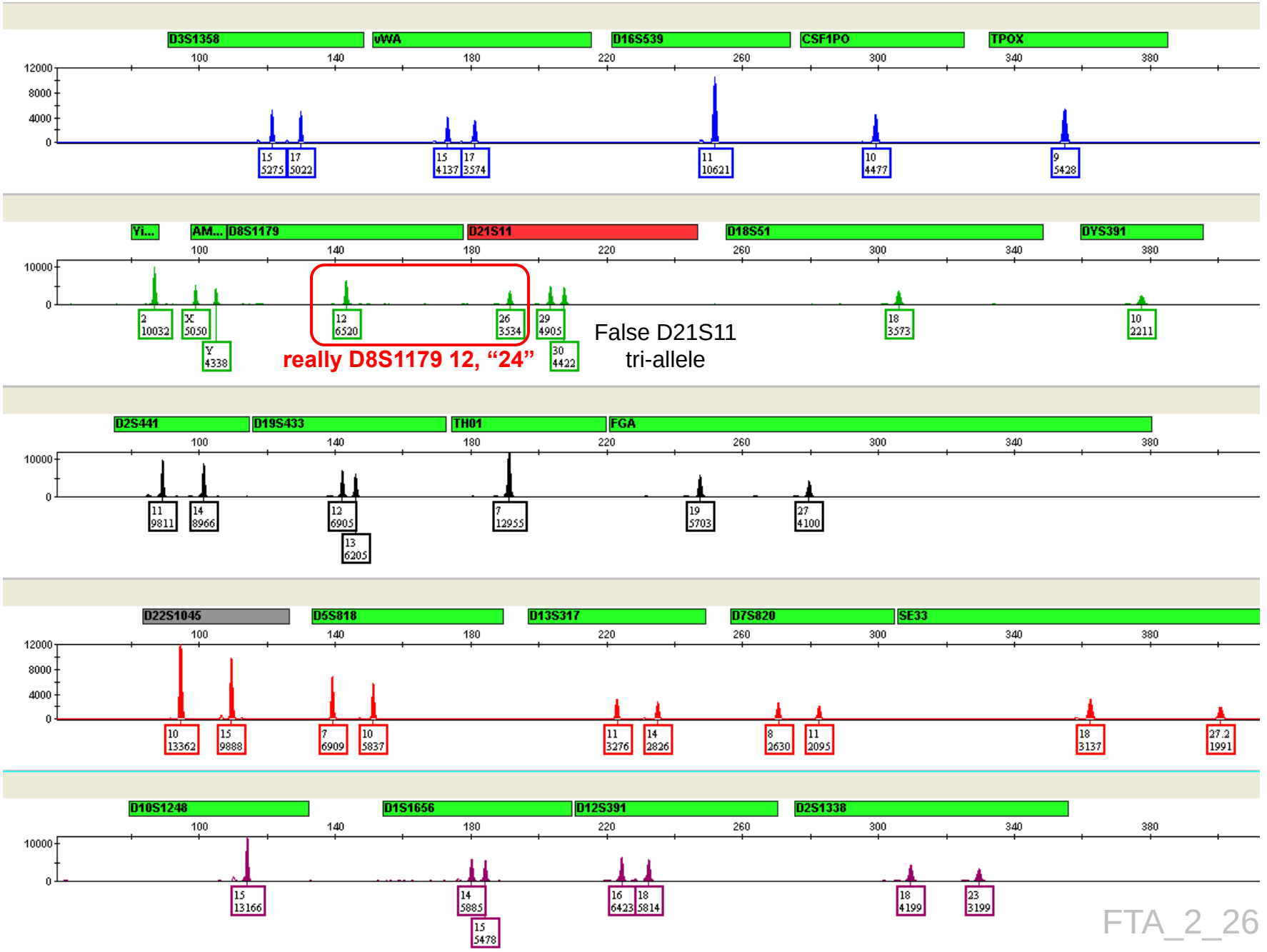
- Tested DNA fragments < 250 bp with GlobalFiler Casework
- GlobalFiler: 10 loci dropped out (>250 bp)
- MiniFiler: Full profile, 7 loci gained (D21S11 and Amel overlap)
- Total of 21 loci included in the profile with combined kits (GF & MF)
- Only 3 loci did not give a signal: TPOX, DYS391 and SE33



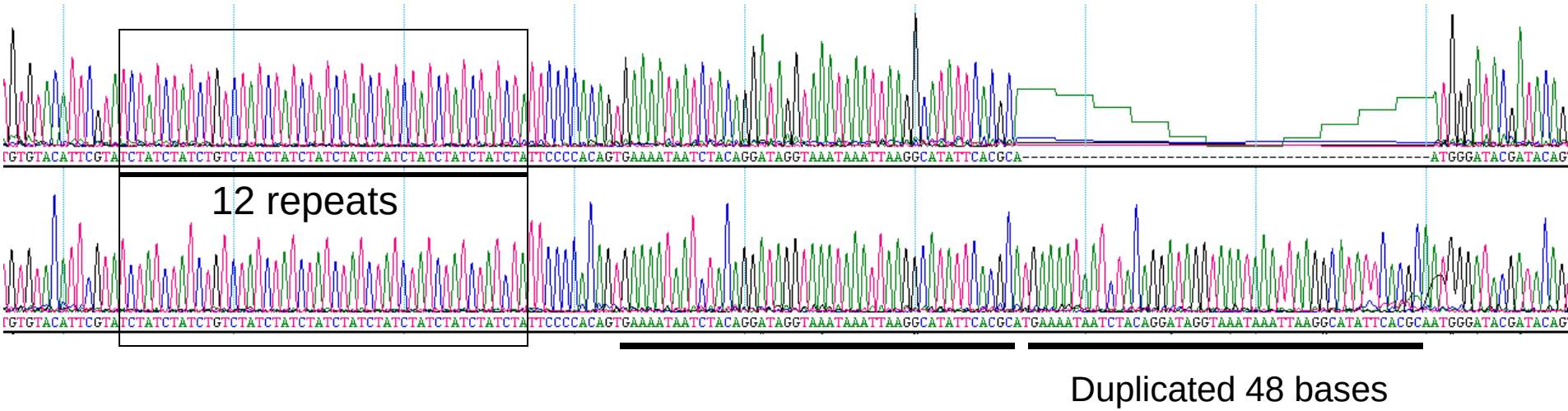
# Concordance Evaluation

- **50 NIST bloodstain sample results were 99.8% concordant** with previous Identifiler Plus, PowerPlex 16 HS, PowerPlex 18D, and PowerPlex ESI 17 Pro results
- **Two loci out of 1100 comparisons** (22 loci x 50 samples) were discordant
  - A large **D8S1179** “24” allele produced a false D21S11 tri-allele in bloodstain #26 (ID/NGM: 25, 29, 30 and GF: 26, 29, 30)
  - **D22S1045** allele 15 dropout/severe imbalance in one sample (bloodstain #29)

# A large D8S1179 allele runs into the adjacent D21S11 locus



# D8S1179 12,“24”



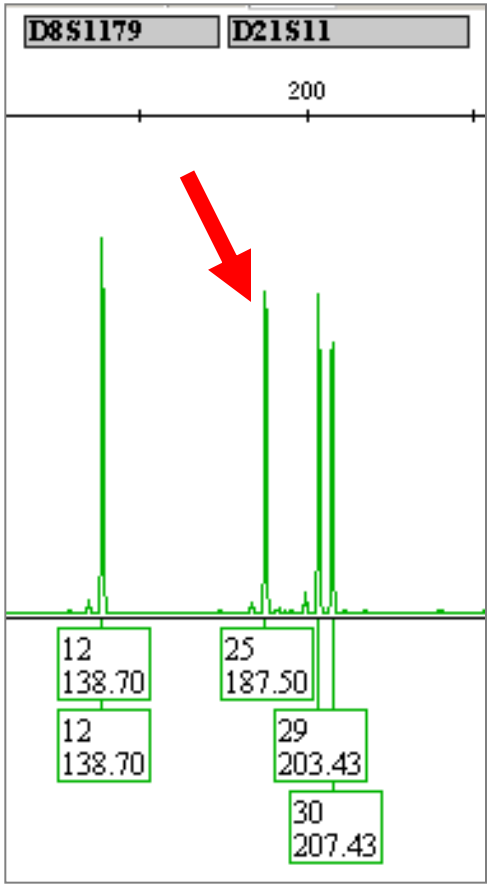

Allele 12 : $[TCTA]_2$ TCTG $[TCTA]_9$

**Allele “24”** : $[TCTA]_2$ TCTG $[TCTA]_9$ **duplication of the 48 bases 10 bases downstream of the repeat**

# Result with This Large D8S1179 Allele

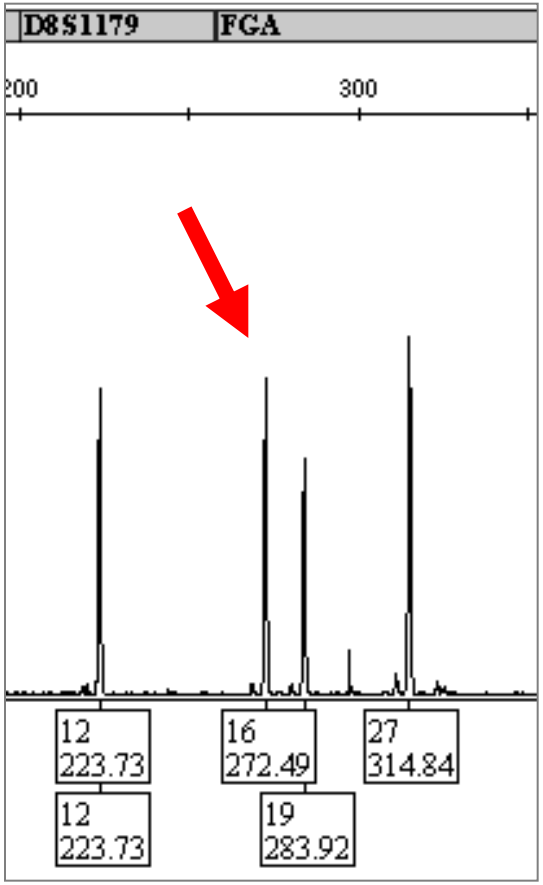
## Using European STR Kits

**NGM SElect**



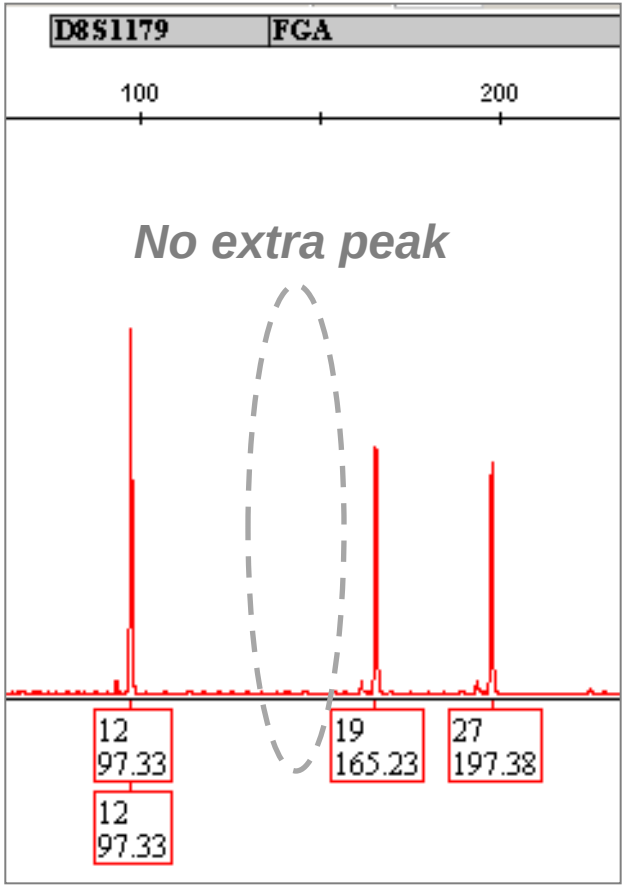

***False D21S11 tri-allele***

**PP ESX 17**




***False FGA tri-allele***

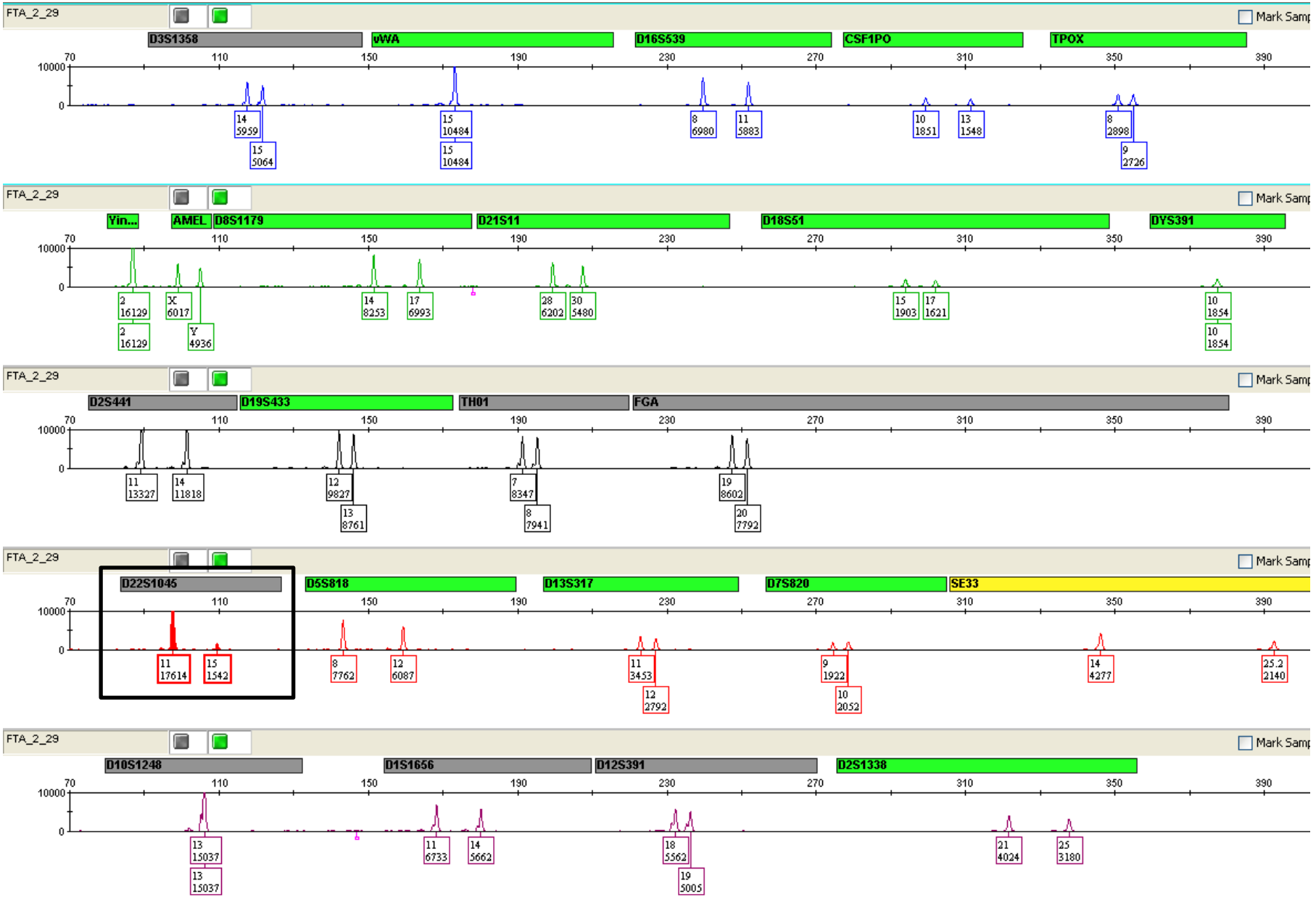
**PP ESI 17**



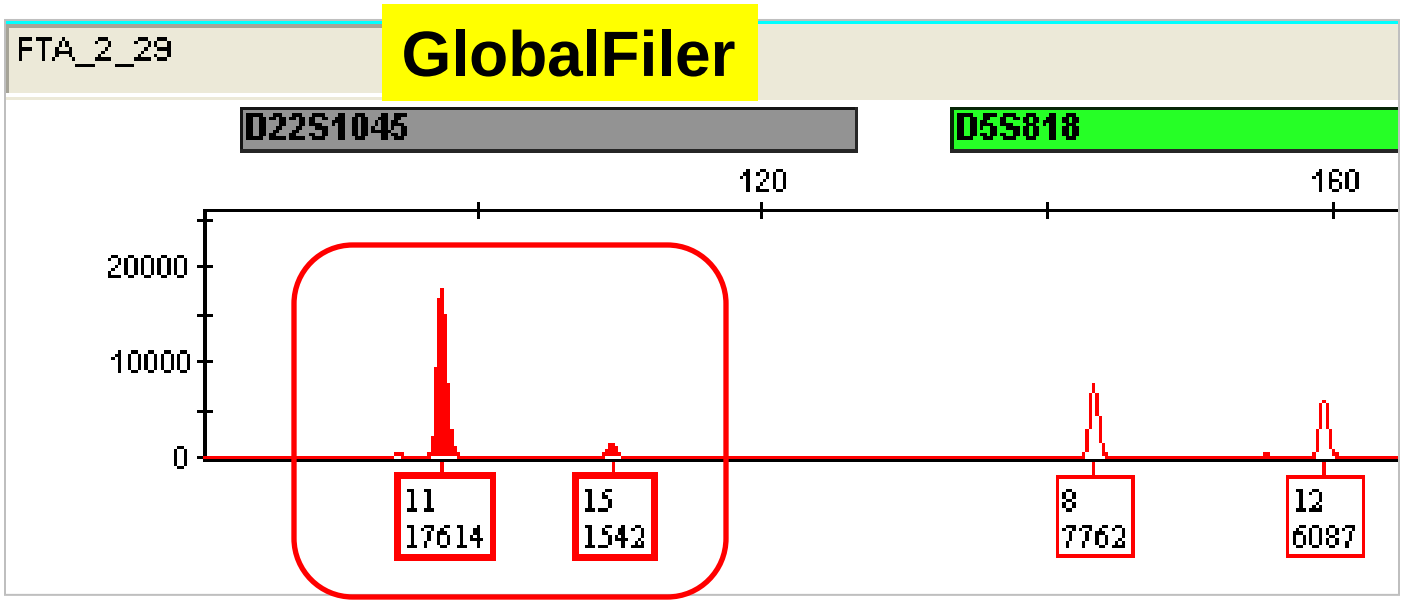

*Reverse primer internal to duplicated flanking region*

Previously presented at CODIS Conference (November 2011)

# D22S1045 Severe Imbalance/Allele Dropout

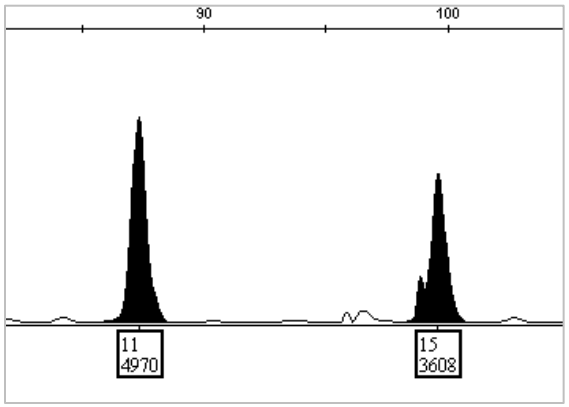


# Likely Primer Binding Site Mutation at D22S1045

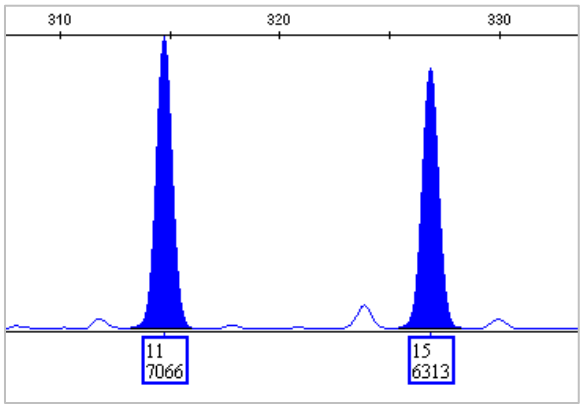



*This sample has not yet been sequenced to discover the underlying sequence around the flanking region*

**NGM SElect**

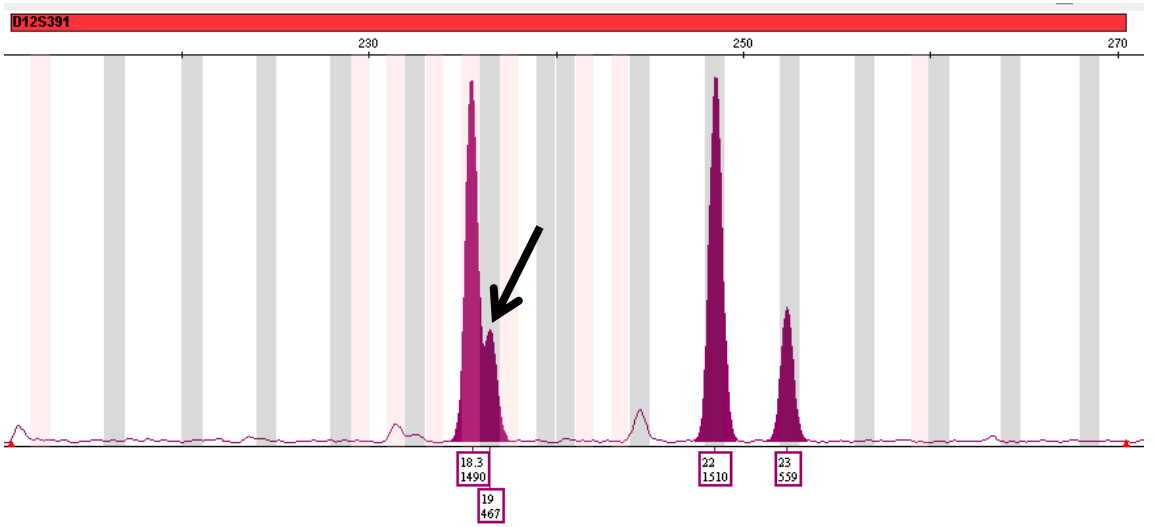



**PowerPlex ESI 17 Pro**



# SRM 2391c Concordance

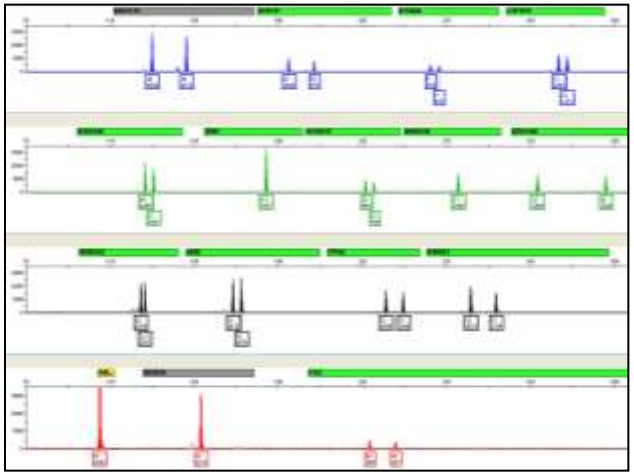
- All SRM 2391c components run with GlobalFiler Casework were **concordant** at all loci
  - Exception: Y indel was not included in the comparison because no other kits use this marker
- Component D at D12S391 shows 1 bp resolution:



# GlobalFiler Express Results

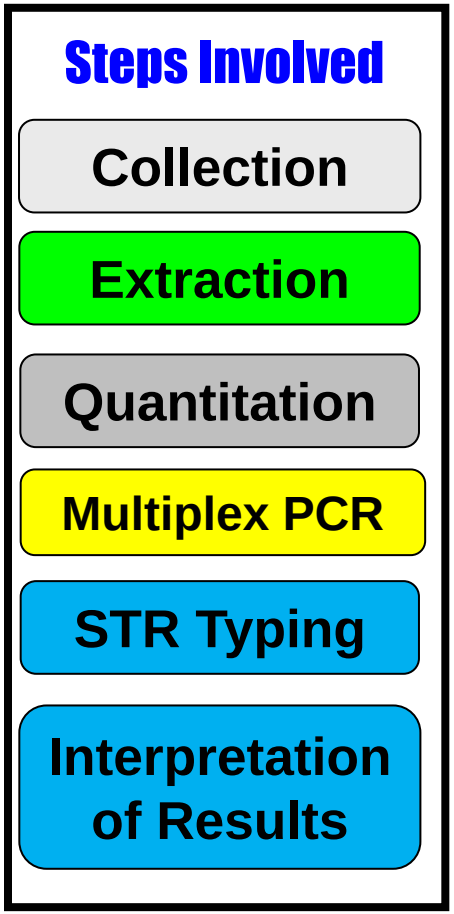
- **Sensitivity Study** (to optimize cycle number)
  - 26 cycles is optimal number in our lab
- **Reproducibility Study**
  - All replicates at 26 cycles gave similar peak height values and ratios
- **Performance Study**
  - All 50 blood samples produced full, well-balanced profiles
- **Concordance Study**
  - Only 2 discordant results observed with 1100 allele comparisons (99.8% concordance)
- **Degraded DNA Study (Casework kit)**
  - Expected results observed with each fragment size

# Rapid DNA Typing: How Fast is Erica Butts?



http://www.cstl.nist.gov/biotech/strbase/pub_pres/Butts-ISHI2012-Rapid-DNA.pdf

# Experimental Design




## GlobalFiler Express

Single Source Reference Samples

**None**
(1.2 mm Blood Punch)

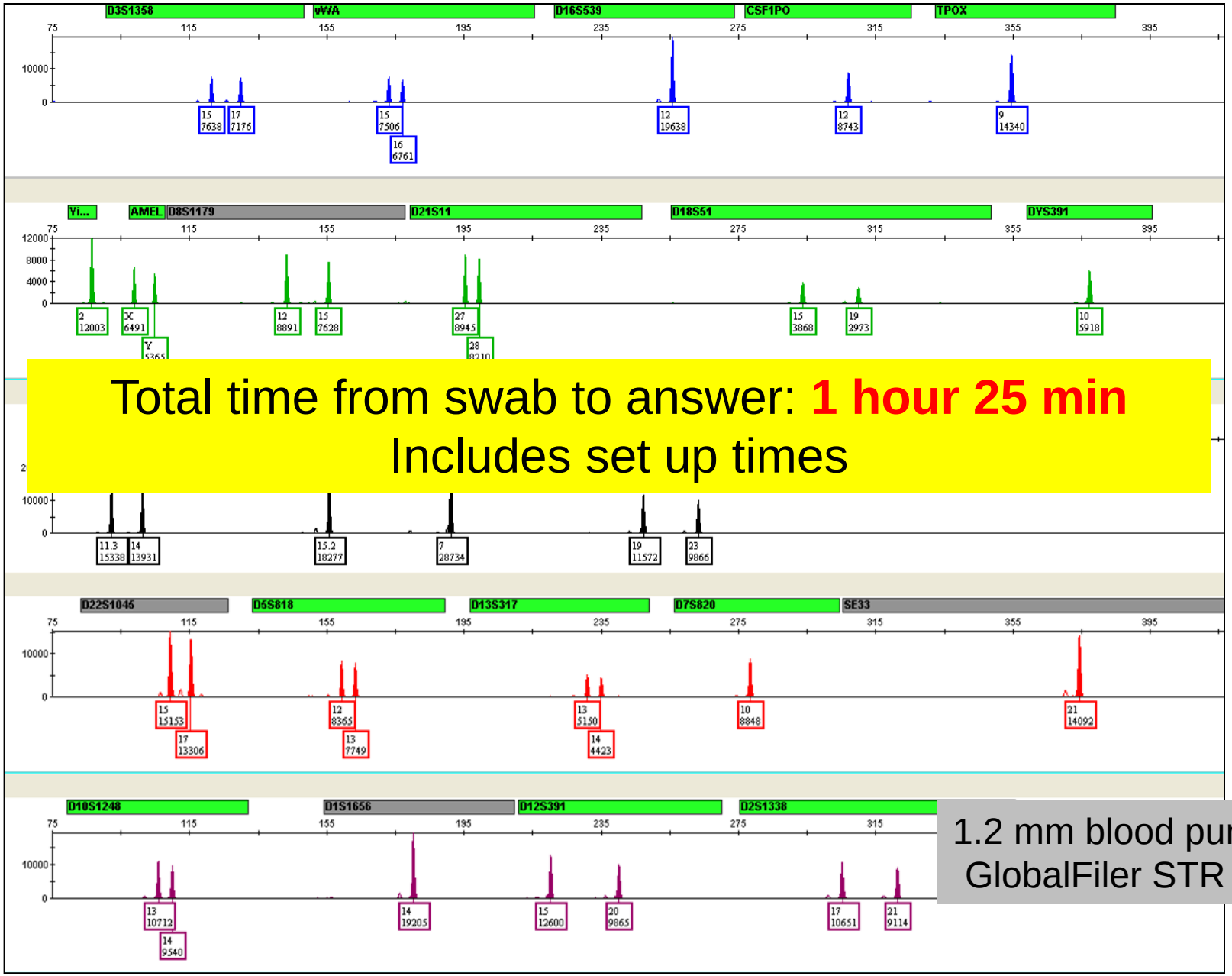
**GlobalFiler Express**
(9700)

**ABI 3500 Genetic Analyzer**
**GeneMapper IDX v1.2**

8 unique samples were typed in parallel

**Testing was timed from collection through data interpretation of results, to include all sample transfer steps**

# FTA → 9700 → 3500



# FBI Consortium Validation Project

- The CODIS database will potentially increase from 13 to 20 core loci
- Two larger STR multiplex kits were developed to meet these requirements
  - GlobalFiler and PowerPlex Fusion
- 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

- Selection of laboratories to participate in validation studies
- **Validation of proposed new CODIS core loci**
- Selection of CODIS core loci
- Implementation of new CODIS core loci into NDIS operations

# 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
  - Data is currently being collected and evaluated from both GlobalFiler and PowerPlex Fusion kits

# Summary

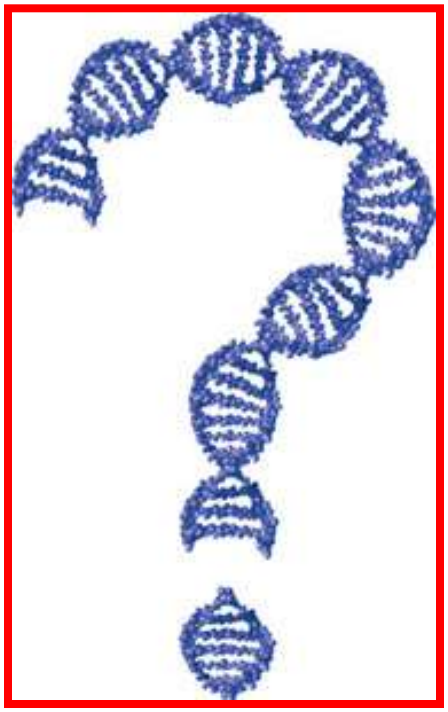
- GlobalFiler is a 6 dye STR multiplex kit that includes the required new loci for upload to CODIS
- GlobalFiler performed well in our initial studies, including concordance, degraded DNA, sensitivity, reproducibility, and performance studies
- GlobalFiler results can be generated in less than 2 hours
  - Overall time includes: collection, sample handling, and liquid transfer steps
- GlobalFiler is currently being validated as part of the FBI Consortium Validation Project and results from these studies will eventually be published and presented

# Thank you for your attention!

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

http://www.cstl.nist.gov/biotech/strbase

**Questions?**



becky.hill@nist.gov
301-975-4275

Acknowledgements:
Erica Butts for GF rapid work

Funding from the **National Institute of Justice (NIJ)** through NIST Office of Law Enforcement Standards