Emerging DNA Technologies

Applications and Information to
Resolve Common Mitotypes

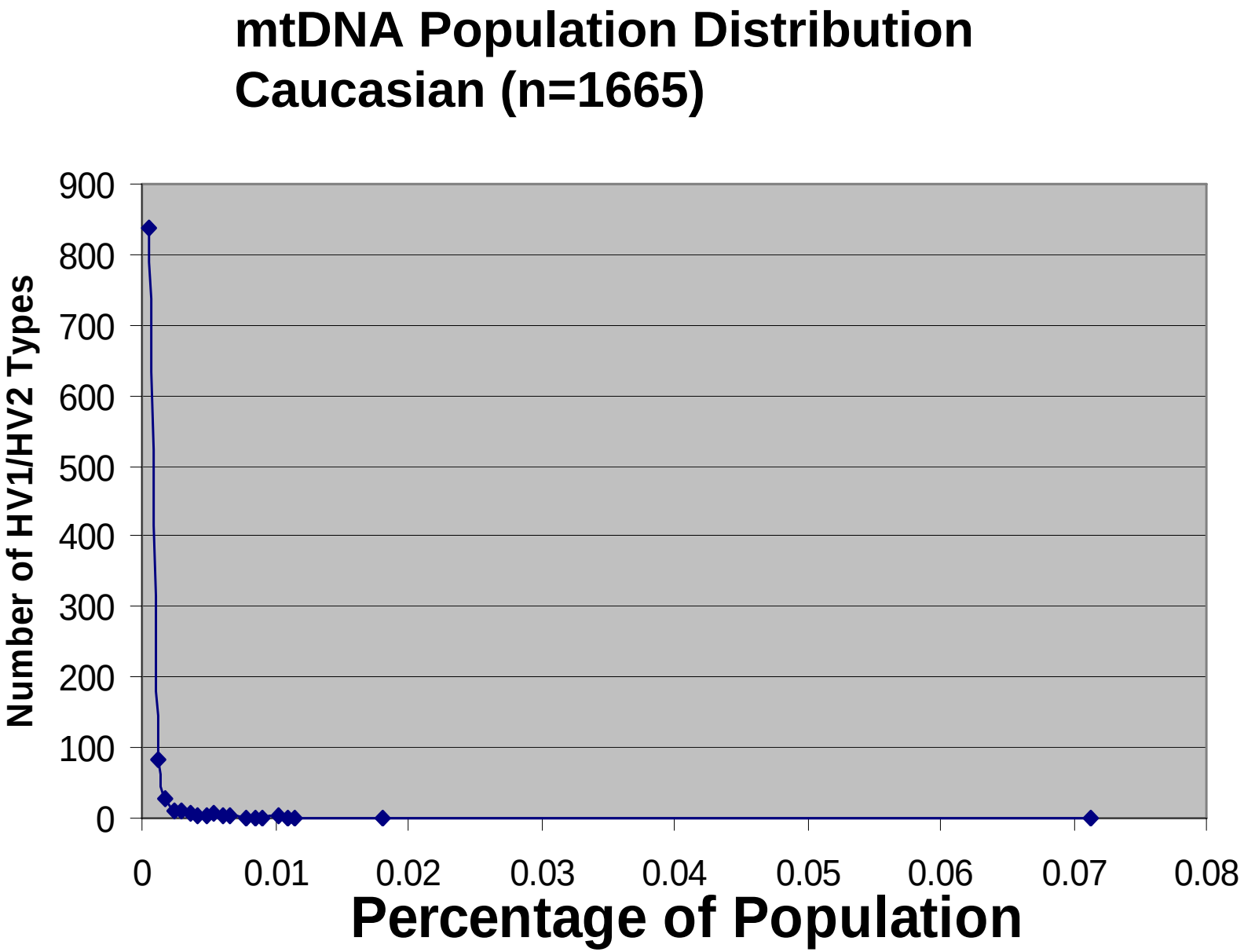
Suni M. Edson
AFDIL

Goals

- Discuss population resolution of mitochondrial DNA
- Examine some new applications that could assist in resolving common mitotypes
  - Cytochrome b
  - Low-copy number (LCN) STR analysis

Why go to mtDNA?

- Disadvantages
  - mtDNA is not a positive form of identification
    - 7% of US Caucasians have the most common type
    - Most common type in US African Americans is shared by ~2.25%.
    - Additional information is required for identification.

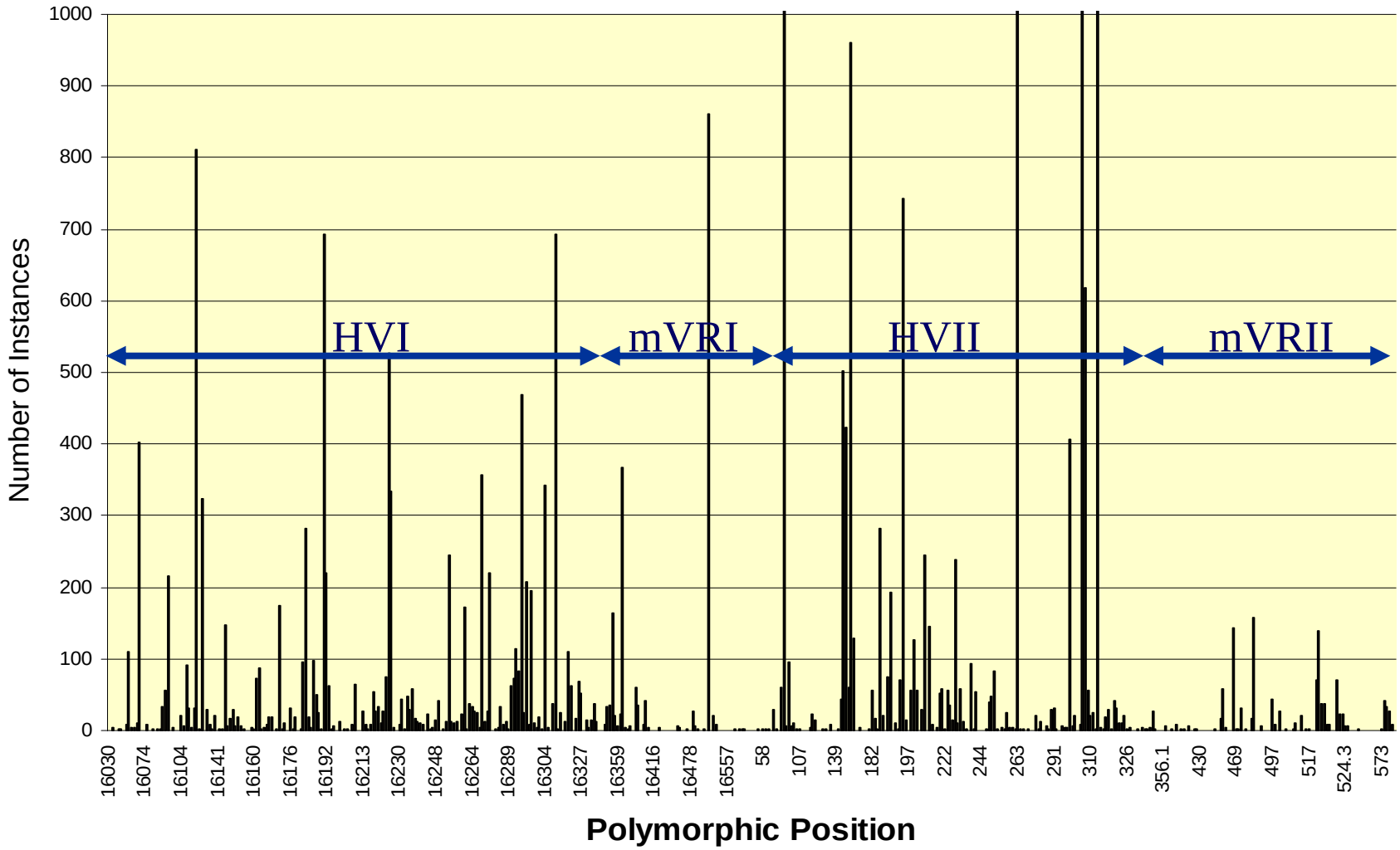

mtDNA Population Distribution
Caucasian (n=1665)
Number of HV1/HV2 Types
900
800
700
600
500
400
300
200
100
0
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
Percentage of Population

Population Resolution

- So, why do we target HV1/HV2 rather than the entire control region (CR) for skeletal remains?

- DiRenzo and Wilson (1991) examined Middle Eastern and Sardinian populations and found that the of the variation in the CR is in the first 400 bases, roughly HV1.

CR variation



Identification

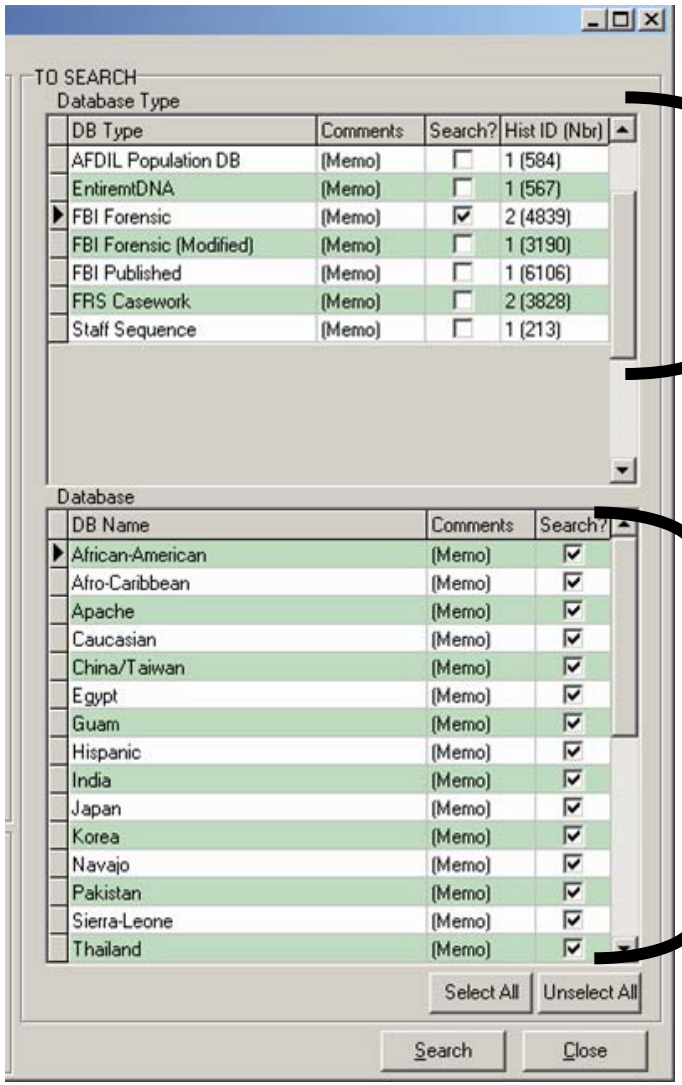
“Multi-hit Match”

Requires additional information

“Single Hit Match”

Less Dependent on additional information

Databases



Available databases

Populations represented in the selected database

SWGDAM database

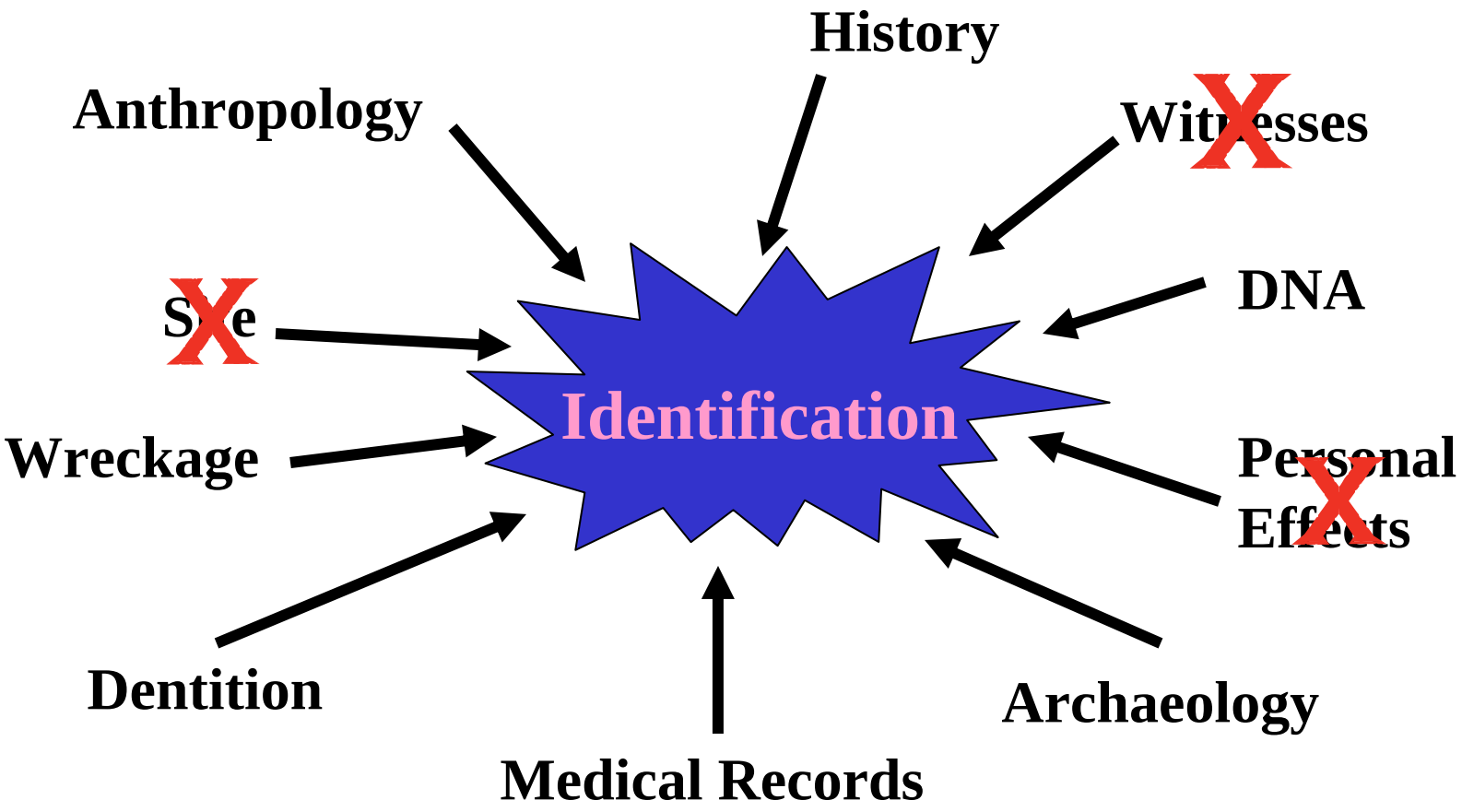
- Caucasians – n=1655
- African Americans – n=1148
- Hispanics – n=686
- African Egyptians – n=75
- African Sierra Leone – n=109
- Japanese – n=163
- Koreans – n=182
- Thai – n=52
- Chinese/Taiwanese – n=329
- Navajo – n=146
- Apache – n=180
- Guam – n=87
- India – n=19
- Pakistan – n=8

Total = 4839

Publications

- In the last two years discussing the need for more data and better quality control for mtDNA databases:
  - Budowle and Polanskey (2005). Science 307(5711): 845-7.
  - Allard, et al. (2005). Fors. Sci. Intl. 148(2-3): 169-79.
  - Just, et al. (2004). Fors. Sci. Intl. 146(Suppl.): S147-9.
  - Budowle, et al. (2004). J. Fors. Sci. 49(6): 1256-61.
  - Bandelt, et al. (2004). Intl. J. Legal Med. 118(5): 267-73.
  - Brandstatter, et al. (2004). Intl. J. Legal Med. 118(5): 294-306.
  - You, et al. (2004). Fors. Sci. Intl. 141(1): 1-6.
  - Parson, et al. (2004). Fors. Sci. Intl. 139(2-3): 215-26.
  - Pereira, et al. (2004). Intl. J. Legal Med. 118(3): 132-6.

Identification with a Loss of Context



Mitochondrial DNA Phylogenetics and Forensics

- Many populations are underrepresented in the available databases.
- We found this to be particularly true in the case of the events surrounding September 11th, 2001.

September 11th, 2001

- AFDIL worked on the remains from the crash of American Airlines 77 into the Pentagon and of United Airlines 93 at Shanksville, Pennsylvania.
- This was a finite set of individuals in both instances.

September 11th, 2001

- Nuclear DNA in conjunction with dental records and fingerprinting were able to identify
  - 178 of the 183 individuals at the Pentagon site
  - All 40 victims from UA93
- We were left with
  - 5 non-referenced profiles from the Pentagon
  - 4 non-referenced profiles from UA93

DNA Analysis

- DNA analysis on these 9 non-referenced profiles
  - STR analysis showed that two were siblings
  - mtDNA analysis confirmed this.
- When searched against a global database, only one sequence generated produced any matches
  - Showed a frequency of 0.14% across all populations.

DNA Analysis

- A fairly recent study by Richards, et. al (2000) examined the HV1 sequence information of 1088 individuals of Near East origin.
- This data can be found at http://www.stats.ox.ac.uk/~macaulay/founder2000/index.html
- The un-referenced sequences were searched against this database as well.

		Database	
Event	Sequence	Global (*n* = 4142)	Near Eastern (*n* = 1088)
AA Flight 77	Sequence 01/02	6	25
	Sequence 03	0	2
	Sequence 04	0	1
	Sequence 05	0	1
UA Flight 93	Sequence 01	0	0
	Sequence 02	0	0
	Sequence 03	0	0
	Sequence 04	0	0

Sequence	Haplogroup	Euro. Caucasian (%) (*n* = 2804)	Near Eastern (%) (*n* = 1088)
Sequence 01	N1b	8 (0.28%)	19 (1.7%)
Sequence 02	U1a	12 (0.43%)	29 (2.7%)
Sequence 03	J1	75 (2.70%)	52 (4.8%)
Sequence 04	U4	84 (3.0%)	21 (1.9%)

Phylogenetics and Larger Databases

- While this exercise was mostly academic, it does illustrate the need for larger databases of underrepresented populations.

- The Research Section of AFDIL is working on several databases for these populations as are numerous laboratories around the world.

Other methods

- For resolving common mitotypes:
  - SNPs
  - LCN-STRs
- For resolving Species:
  - Cytochrome b

Cyt b

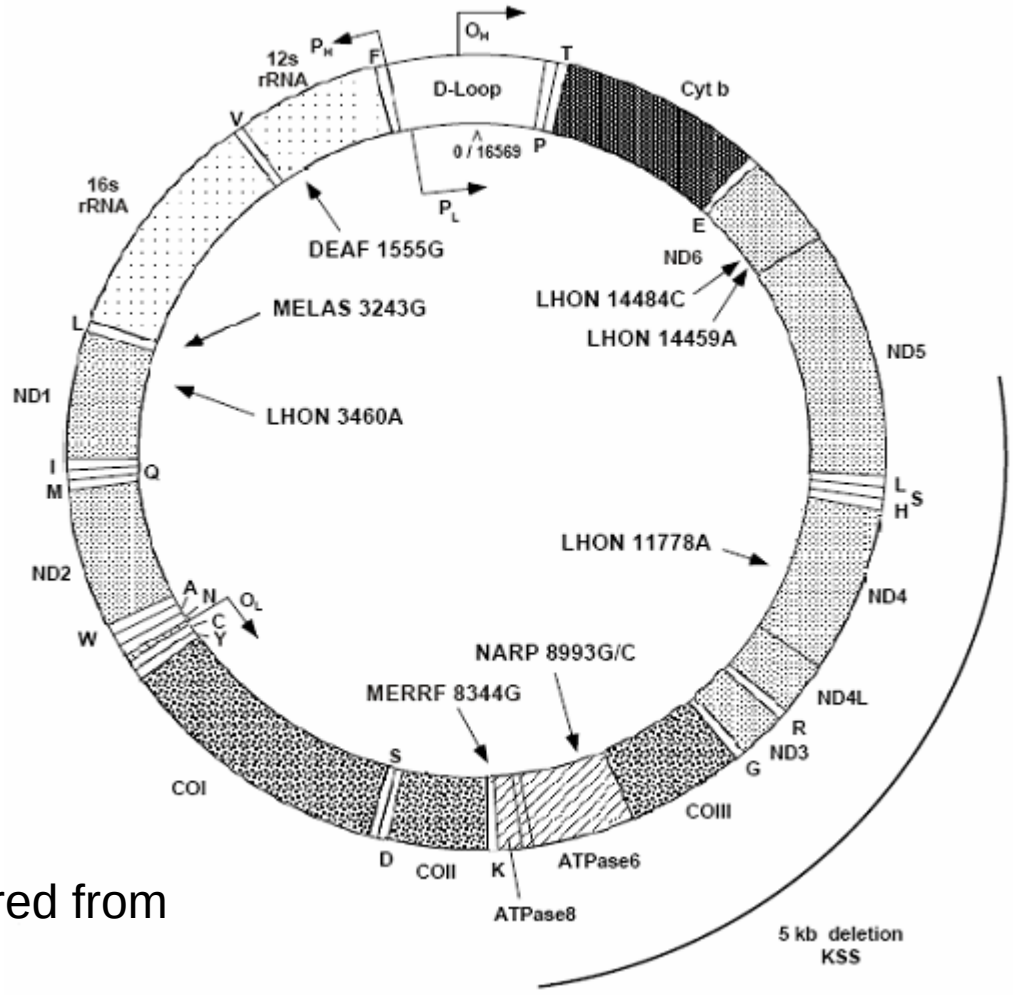


Image acquired from mitomap.org

Cytochrome b

- Shows some promise at individuating species:
  - Work is ongoing at AFDIL to identify primers and positions that could be diagnostic.
- Also commonly used for taxonomic purposes.
  - Much work has been done to use this region for species identification.

Cytochrome b

- Species specific primers have been developed for
  - Dog
  - Cow
  - Sheep
  - Horse
  - Pig
  - Mouse
  - Various other common species.

- While this may not seem particularly pertinent to human identification, it is very useful when looking at sets of remains with bone or tissue fragments that are unidentifiable as human or otherwise.

More Information

- Cytochrome b for species identification is currently ongoing at Strathclyde University and was presented at AAFS 2006.
  - S Tobe and A Linacre (2006). A Single Step Multiplex PCR to Identify Mammalian Species in the United Kingdom.
- Recent publications in Cytochrome b analysis:
  - Hsieh, et al. (2006). Species Identification of Kachuga tecta Using the Cytochrome b Gene. J. Fors. Sci. 57(1): 52-6.
  - An, et al. (2006). A molecular genetic approach for species identification of mammals and sex determination in birds in a forensic case of poaching from South Korea. Fors. Sci. Int.: *Epub ahead of print.*

LCN-STR analysis

- Low-copy number (LCN) short tandem repeat analysis is fairly recent addition to the toolkit of sorting common mitotypes.
- Used more often for ‘touch DNA’ samples.
- Peter Gill at the FSS has done a lot of the foundation work for this technique.
  - Calling guides
  - Reaction conditions
  - Gill, P (2001). Application of low copy number DNA profiling. Croat. Med. J. 42(3): 229-32.

LCN is Not Mini-STR's!

- Mini-STR's refer to the moving of the primers closer to the target repeat region.
  - Decreases size of amplicon overall
  - Designed for degraded or low-copy samples.

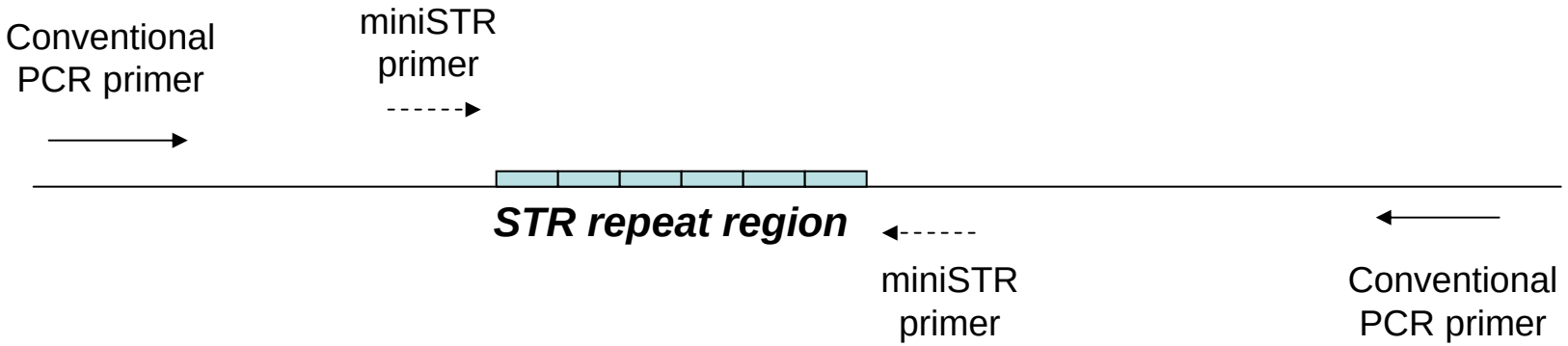


Figure 7.2, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

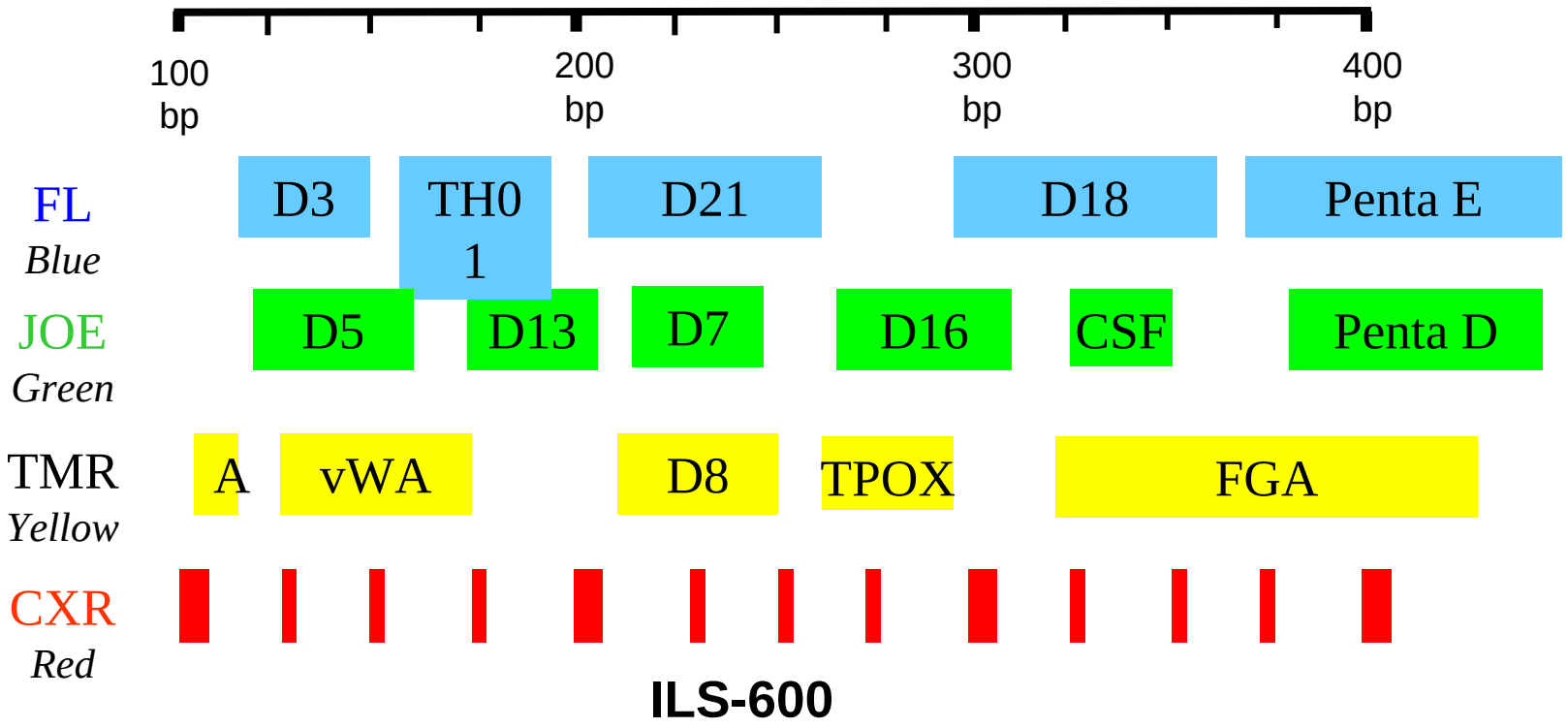
LCN-STR

- Most commonly refers to an input of less than 100pg of DNA.
- Although this varies by lab and the kit used
  - AmpFlSTR Identifiler
  - PowerPlex16
- AFDIL's preferred input of template DNA for STR analysis is 1ng – 250pg.

LCN-STR Analysis

- Modifications to commercially available kits are used.
  - Increased cycle number
  - Increased Taq
  - Addition of other components that could increase primer fidelity.

PowerPlex® 16



Low Copy Number (LCN) nucDNA Modifications

Powerplex 16	LCN – PP16
0.8 ul Taq Gold	1.6 ul Taq Gold
30 cycles	36 cycles
Input DNA – Targeted between 250pg and 1ng	Input DNA – below 250pg

Still under development at AFDIL. These are our initial parameters that we have not completed validation on.

LCN Concerns

- Increased artifacts due to the increased sensitivity of the reaction and the assumedly degraded nature of the sample.
  - Increased stutter
  - Allelic drop-in
  - Allelic drop-out
  - Incomplete adenylation

- Of course these are all issues with regular STR analysis, but could be more prevalent in LCN.
- Inefficient Quantitation?

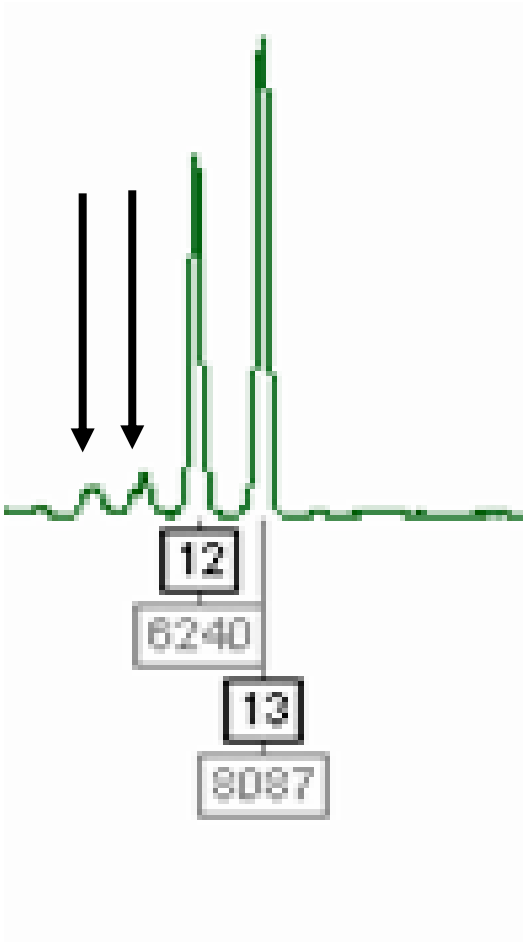
Stutter Products

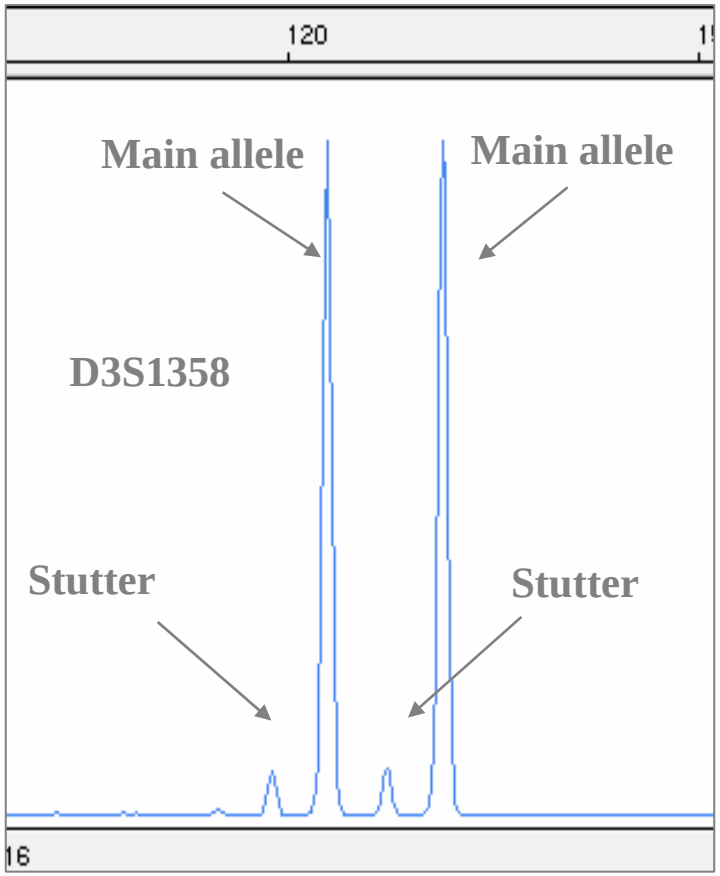
- Peaks that show up primarily one repeat less than the true allele as a result of strand slippage during DNA synthesis
- Stutter is less pronounced with larger repeat unit sizes

  (dinucleotides > tri- > tetra- > penta-)
- Longer repeat regions generate more stutter
- Each successive stutter product is less intense

  (allele > repeat-1 > repeat-2)
- Stutter peaks make mixture analysis more difficult

Stutter

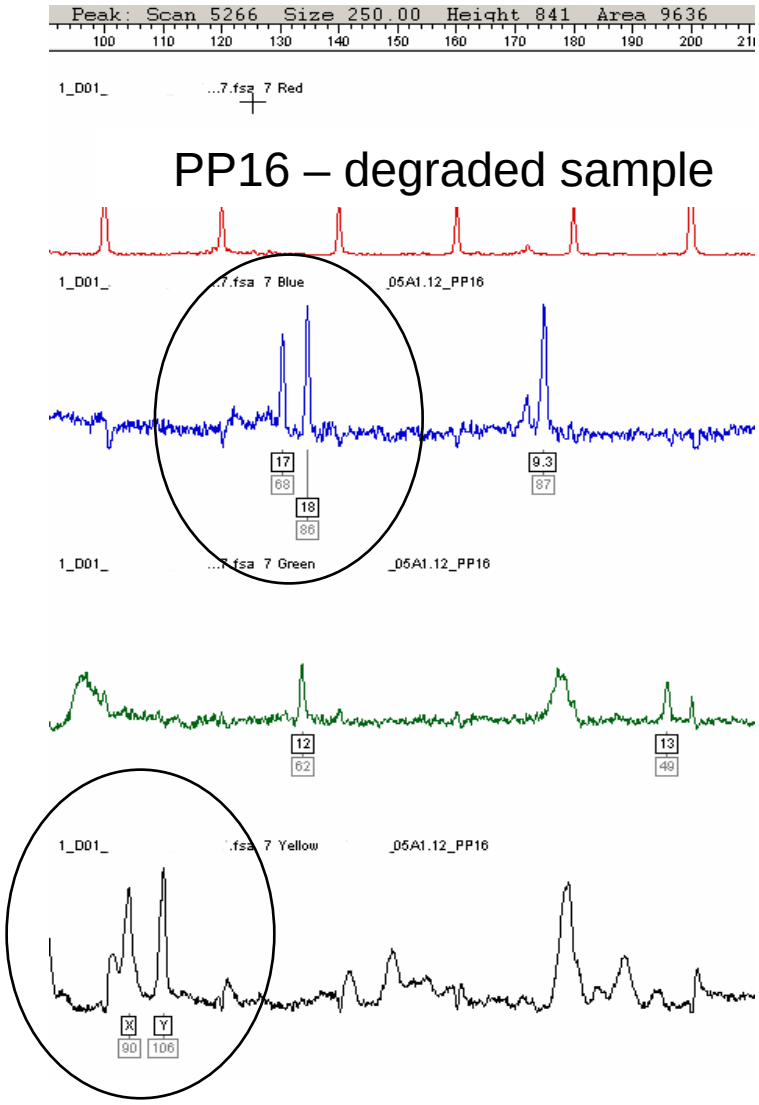

12
6240
13
8087

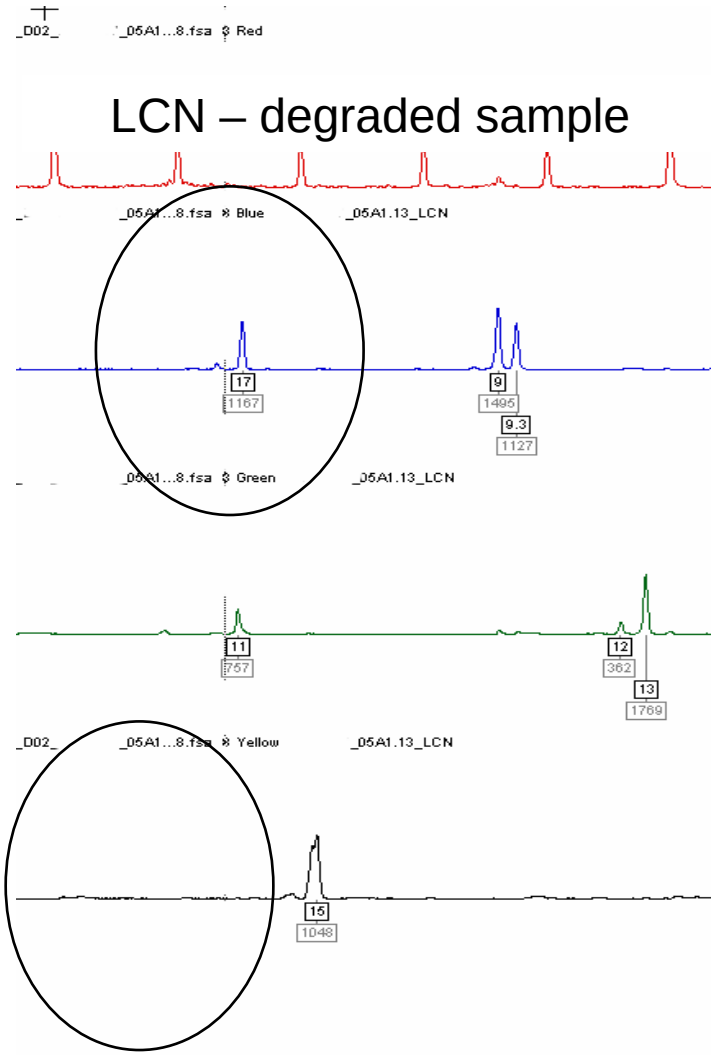

120
Main allele
Main allele
D3S1358
Stutter
Stutter
16

Drop-Out

- Allele is present in the DNA sample but fails to be amplified
- Allele dropout is a problem because a heterozygous sample appears falsely as a homozygote
- This phenomenon impacts DNA databases
- Would expect to see this more frequently in larger alleles that fail to amplify due to low-quality samples
- LCN-STR calling and reporting parameters need to be designed to account for this.

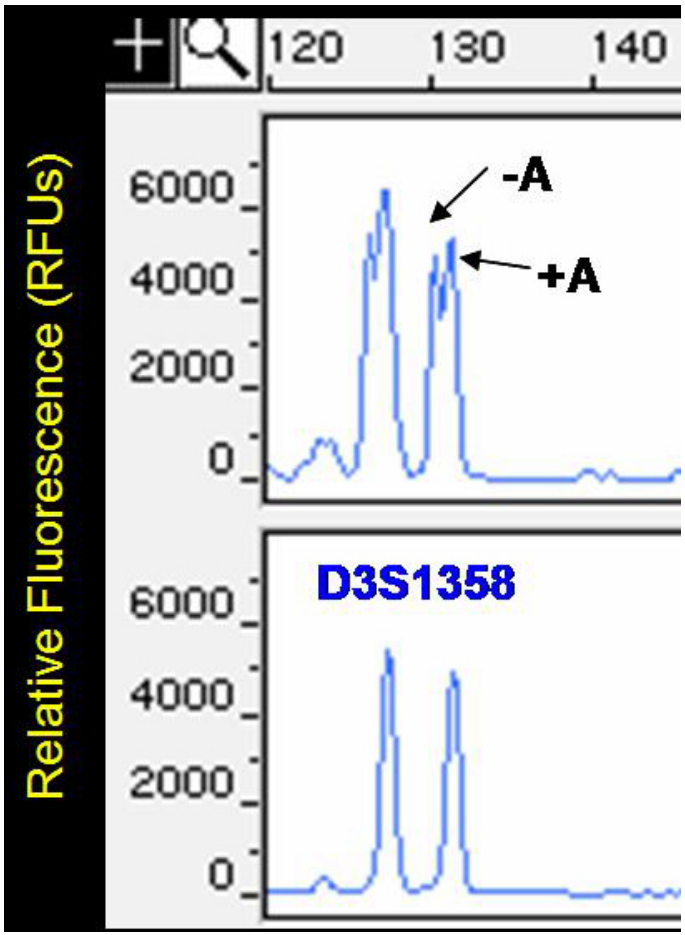
Drop-Out


Peak: Scan 5266 Size 250.00 Height 841 Area 9636
100 110 120 130 140 150 160 170 180 190 200
1_D01_ ...7.fsa 7 Red
PP16 – degraded sample
1_D01_ ...7.fsa 7 Blue _05A1.12_PP16
17 68
18 86
9.3 87
1_D01_ ...7.fsa 7 Green _05A1.12_PP16
12 62
13 49
1_D01_ .fsa 7 Yellow _05A1.12_PP16
X 90
Y 106


_D02_ _05A1...8.fsa Red
LCN – degraded sample
_05A1...8.fsa Blue _05A1.13_LCN
17 1167
9 1495
9.3 1127
_05A1...8.fsa Green _05A1.13_LCN
11 757
12 362
13 1769
_D02_ _05A1...8.fsa Yellow _05A1.13_LCN
15 1048

Overloading

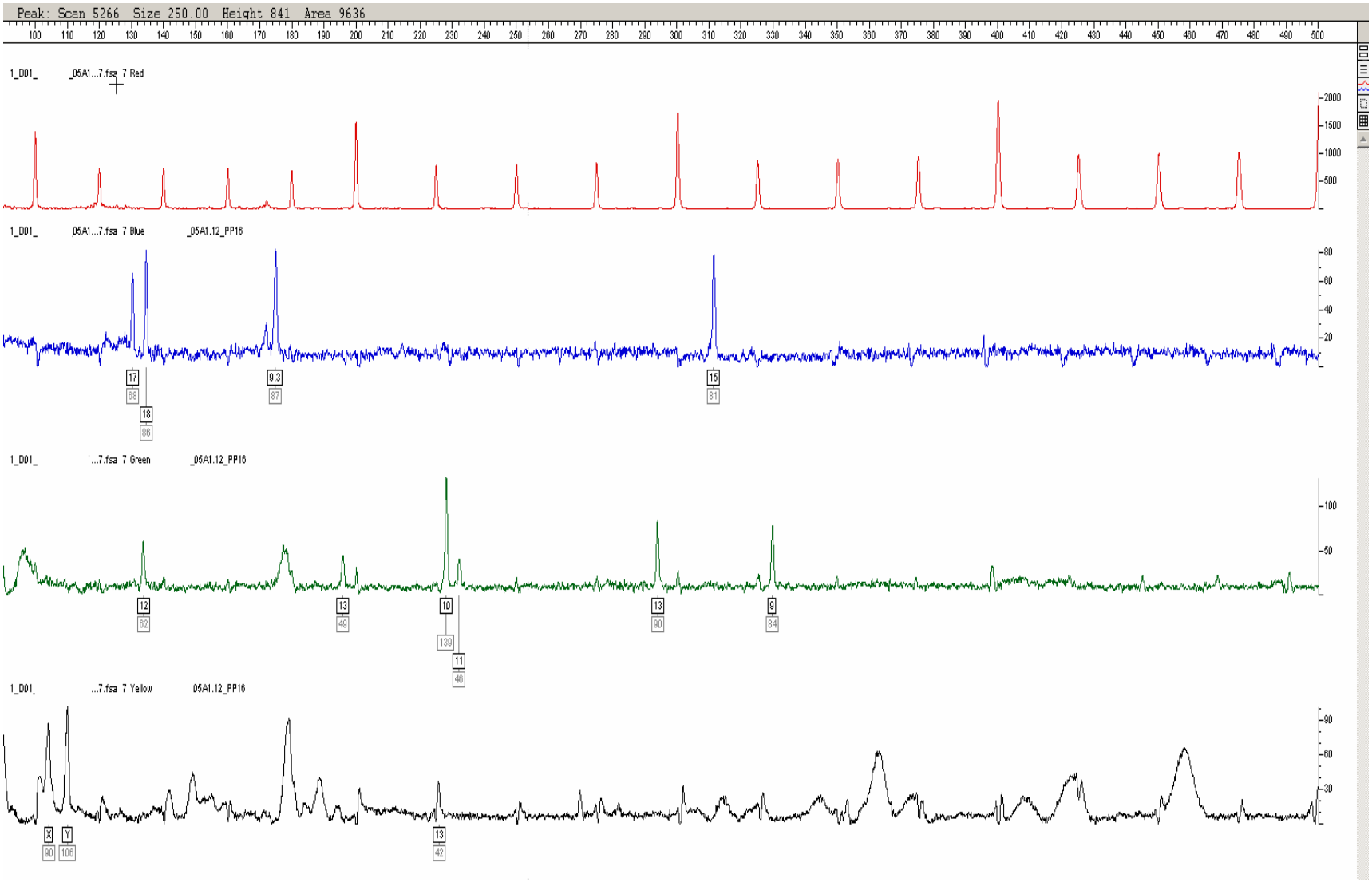
- Occurs with the addition of too much DNA to the reaction.
- Taq polymerase will often add an extra nucleotide to the end of a PCR product; most often an “A”



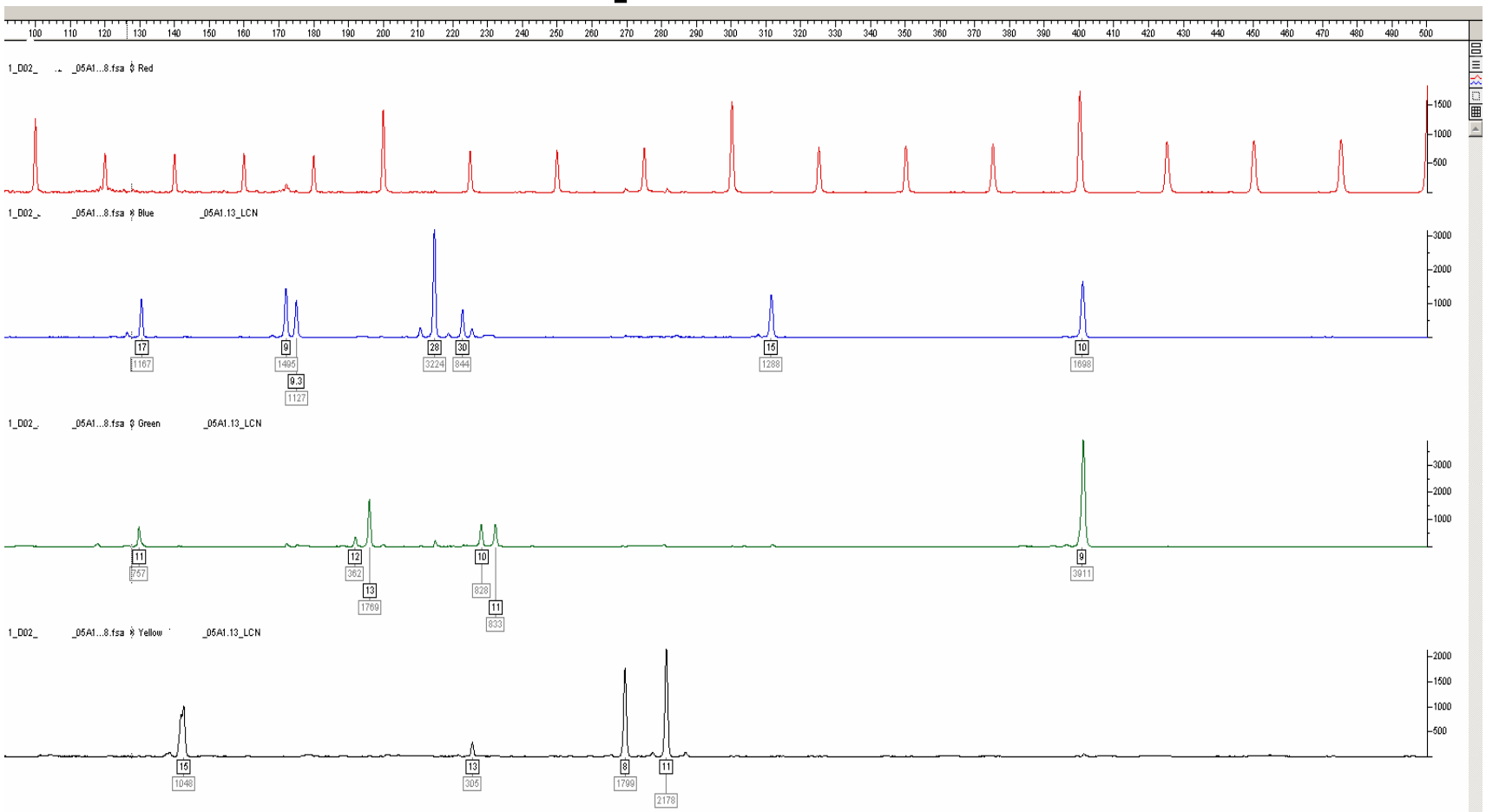
Quantitation

- LCN-STR analysis is very sensitive to input DNA templates.
- qPCR doesn’t always give an accurate measure of the DNA present in the solution when the sample is degraded.
- Samples that quantify as “no DNA present” will often produce STR results.

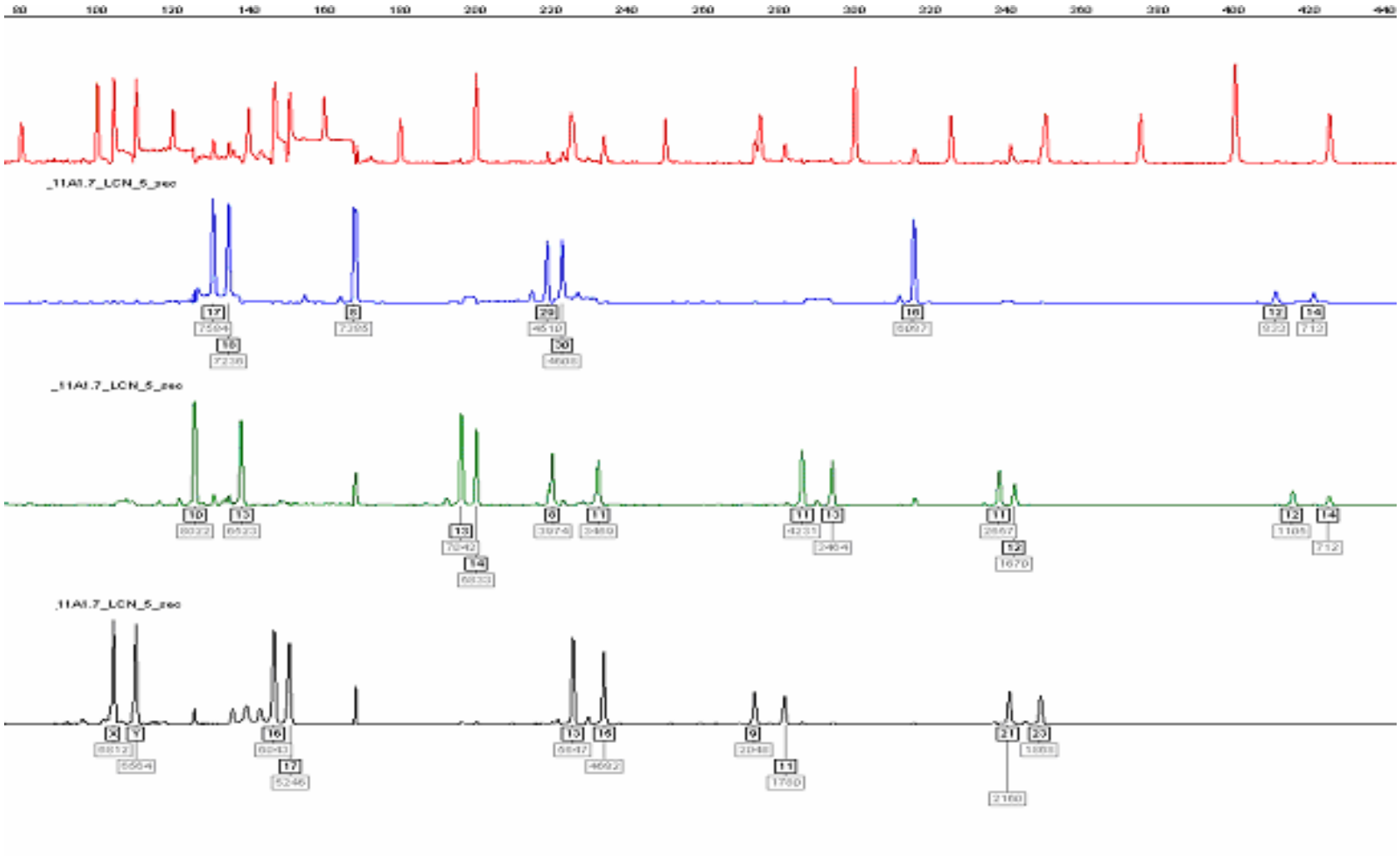
PP16 amp of 0 DNA



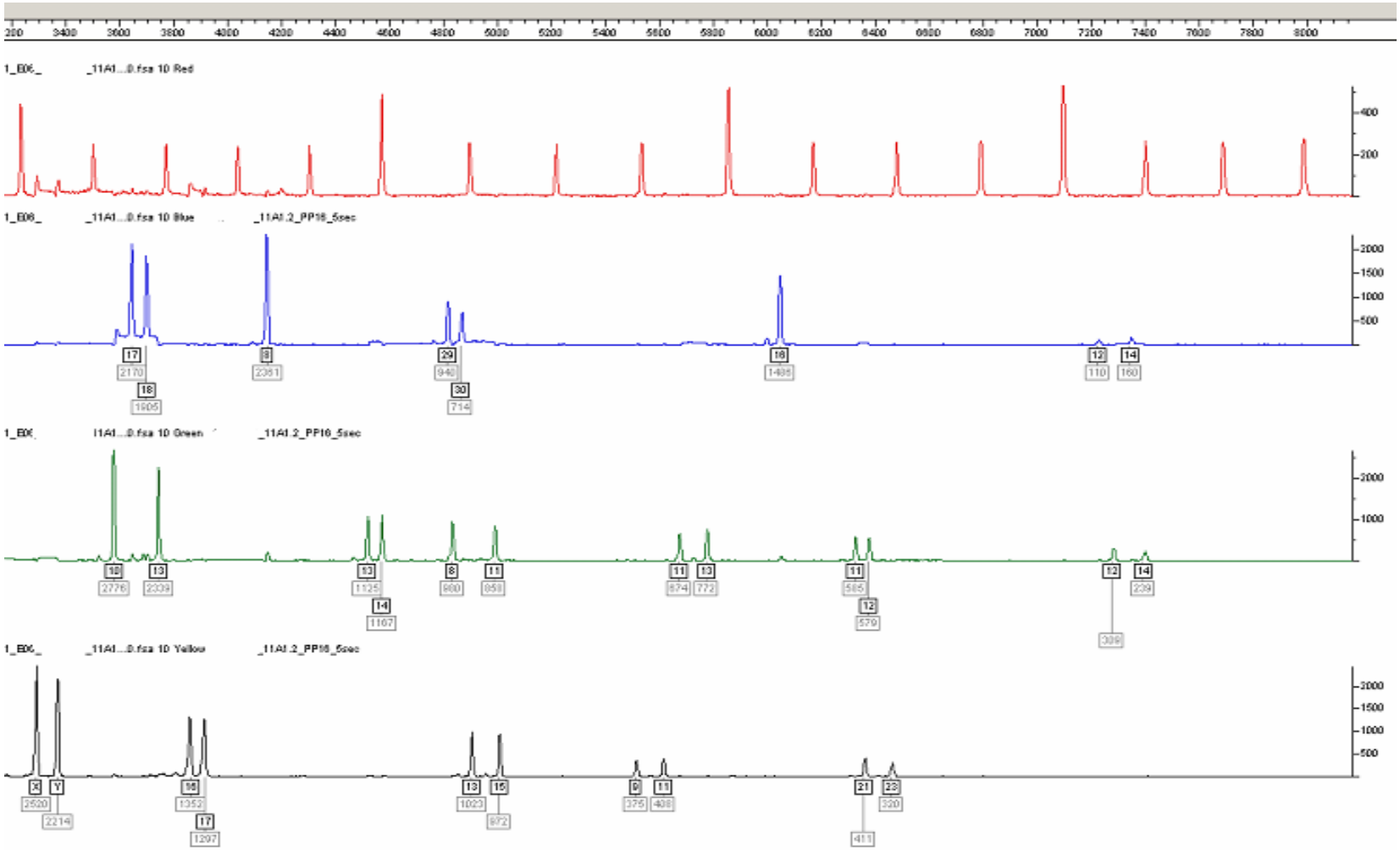
LCN amp of 0 DNA



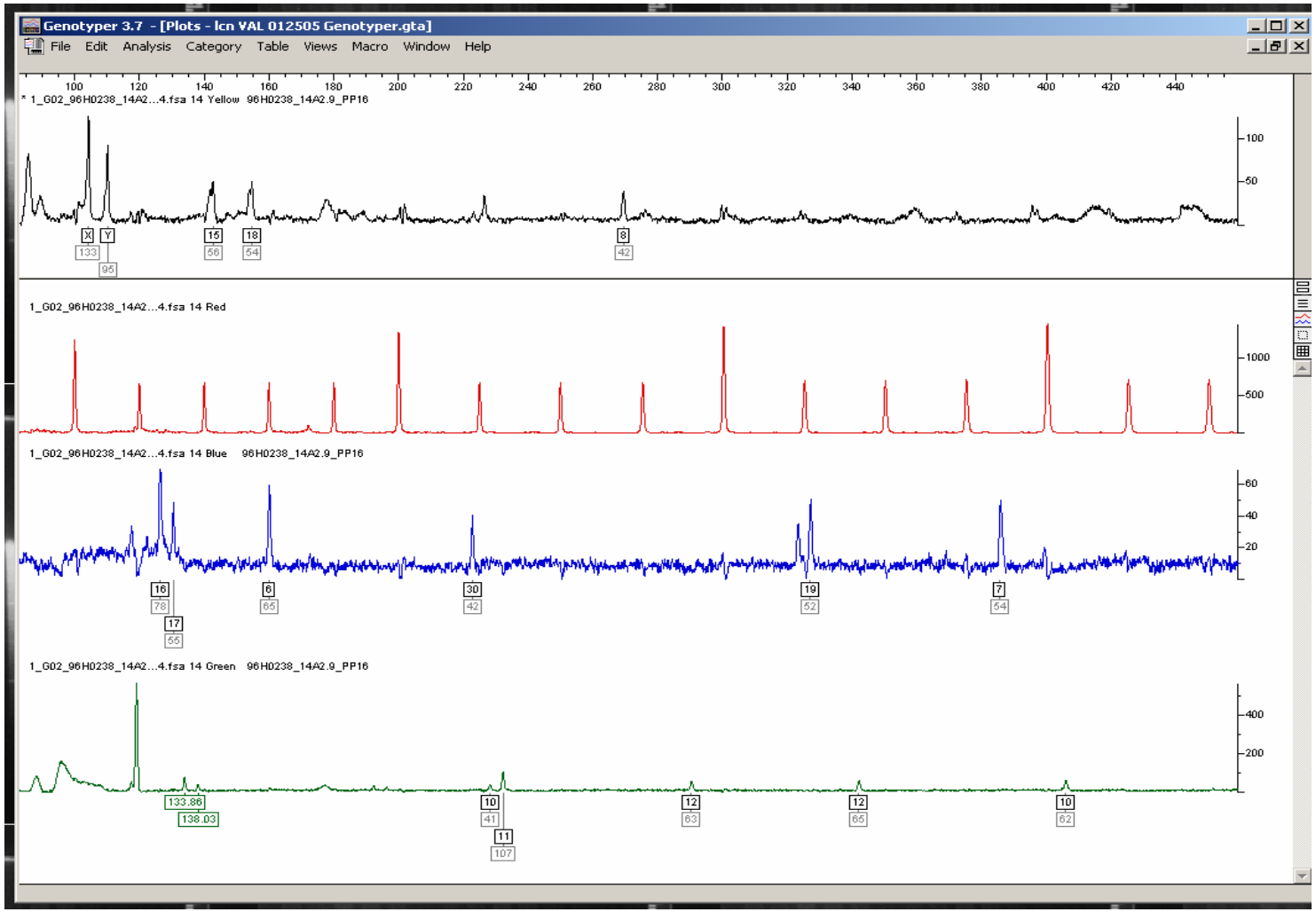
LCN amp of ‘Inhibited’ Sample



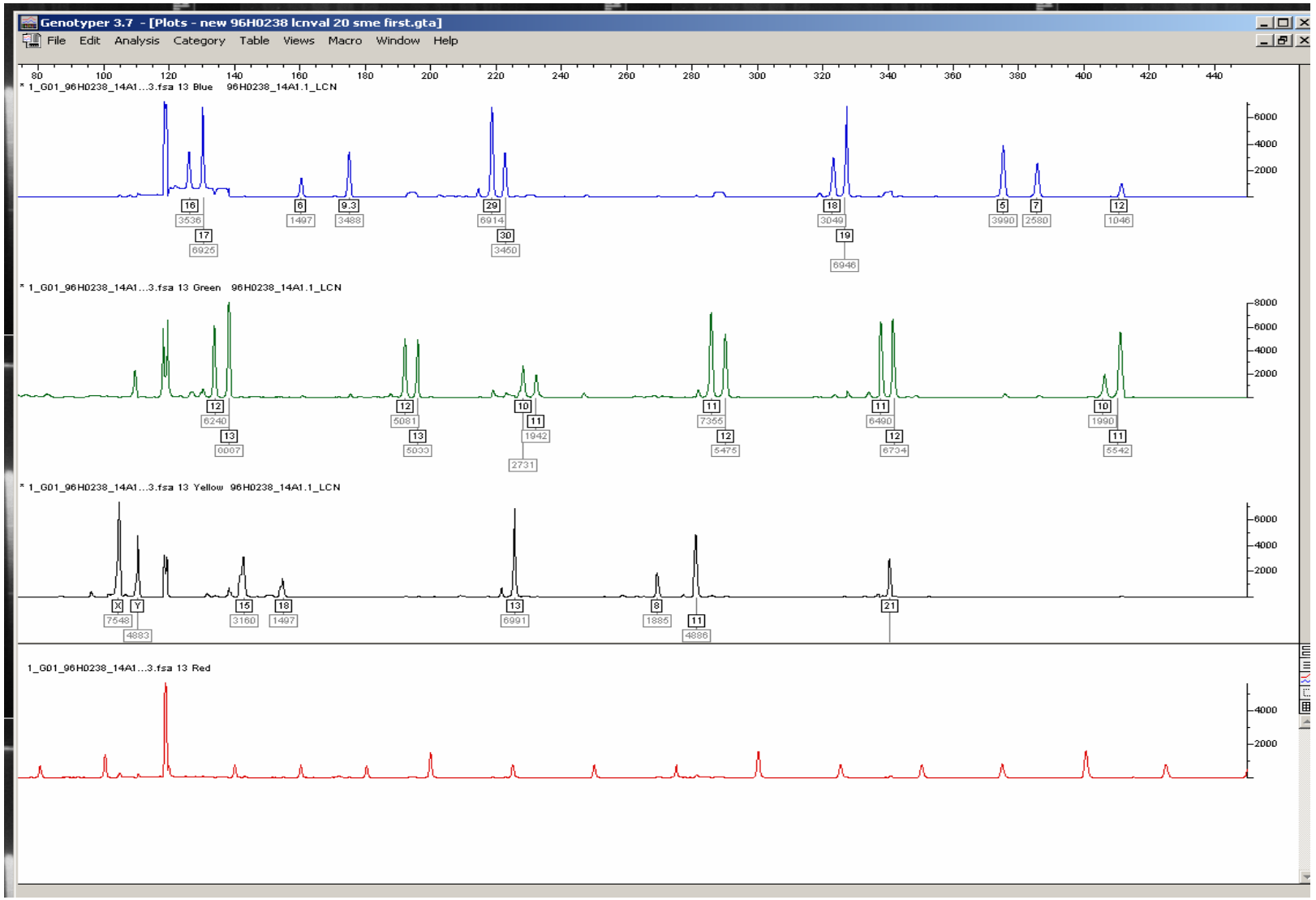
PP16 amp of ‘Inhibited’ Sample



PP16 with an input of 137pg



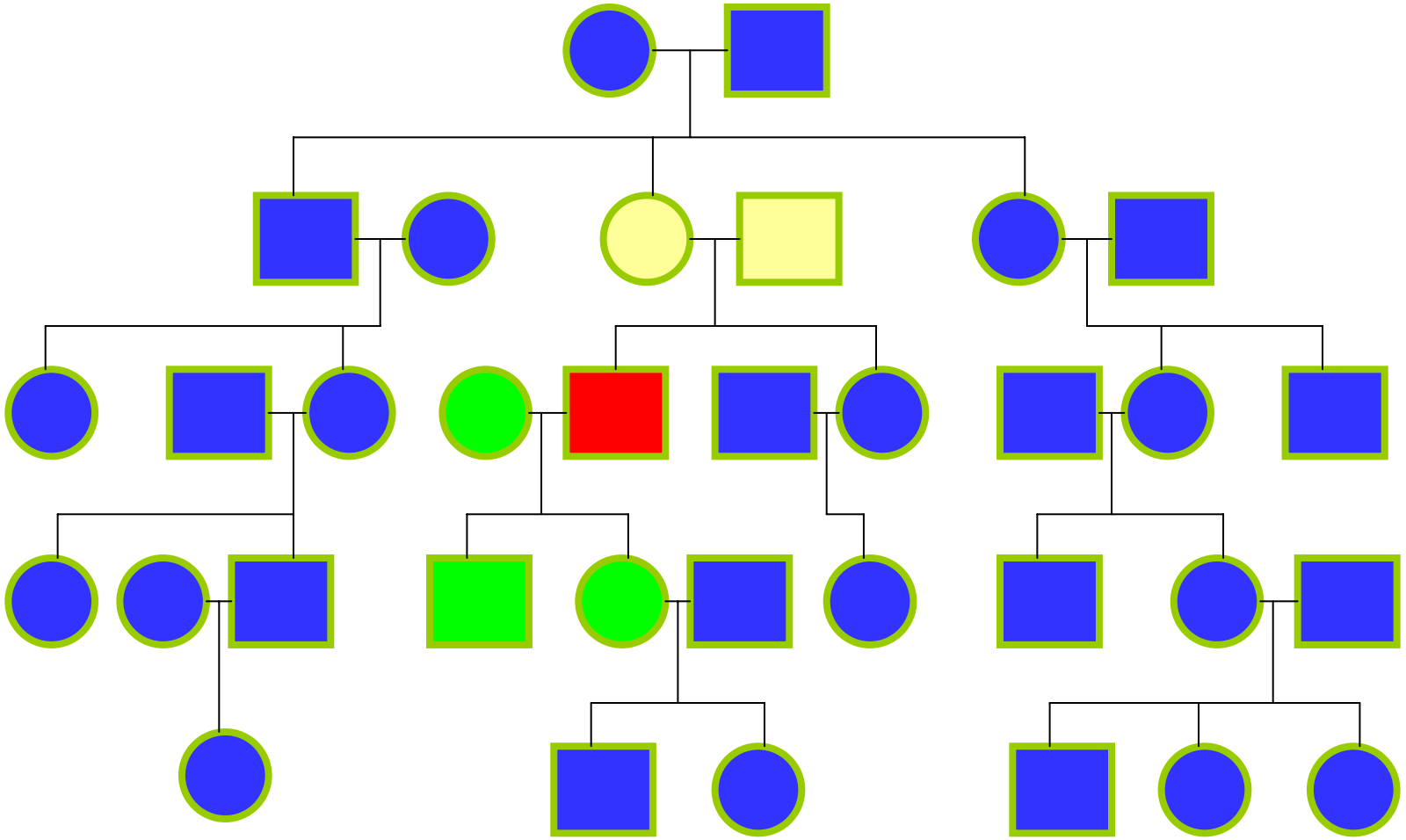
LCN with an input of 137pg



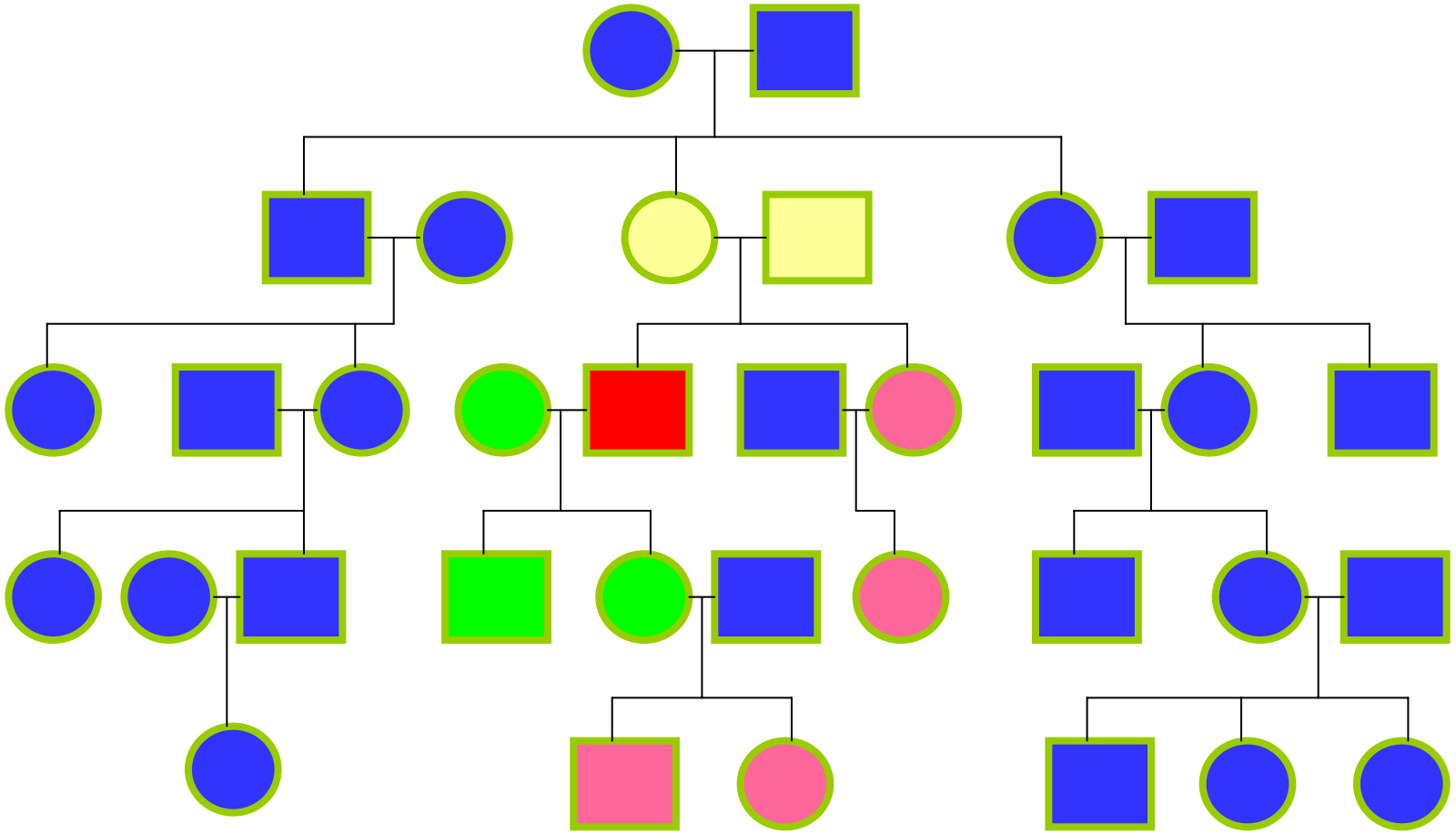
How does this help?

- Sets of commingled remains with common mitotype can be separated with other techniques such as LCN-STR.
- But what about relatives for comparison?

References Available for Nuclear DNA Analysis



Statistical analysis could increase the pool…



Case study

- Aircraft was lost at the beginning of the Vietnam War
- 4 people on board
  - We have mitochondrial DNA references for 2 of the 4.
  - The other two have no available maternal relatives.

Case Study

- One mitotype from the remains:

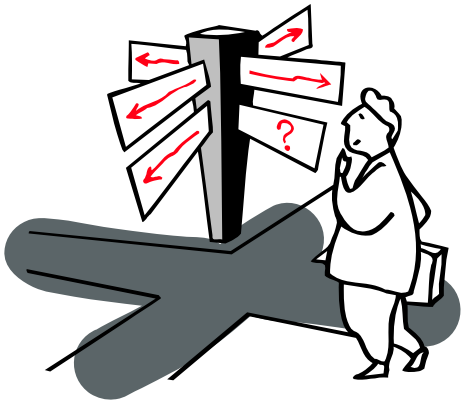
263G

315.1C

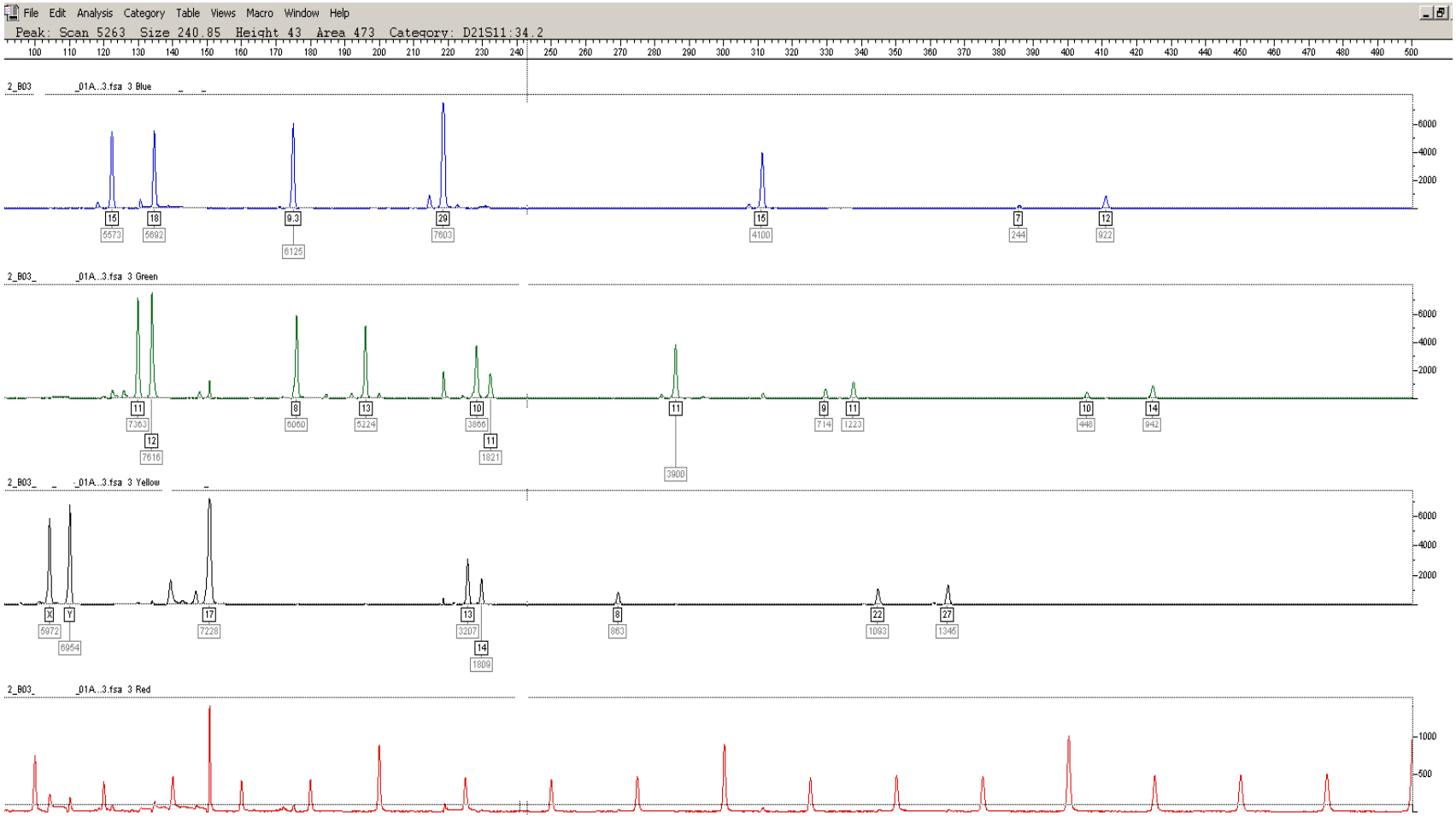
- Matches one of the references.

Case Study

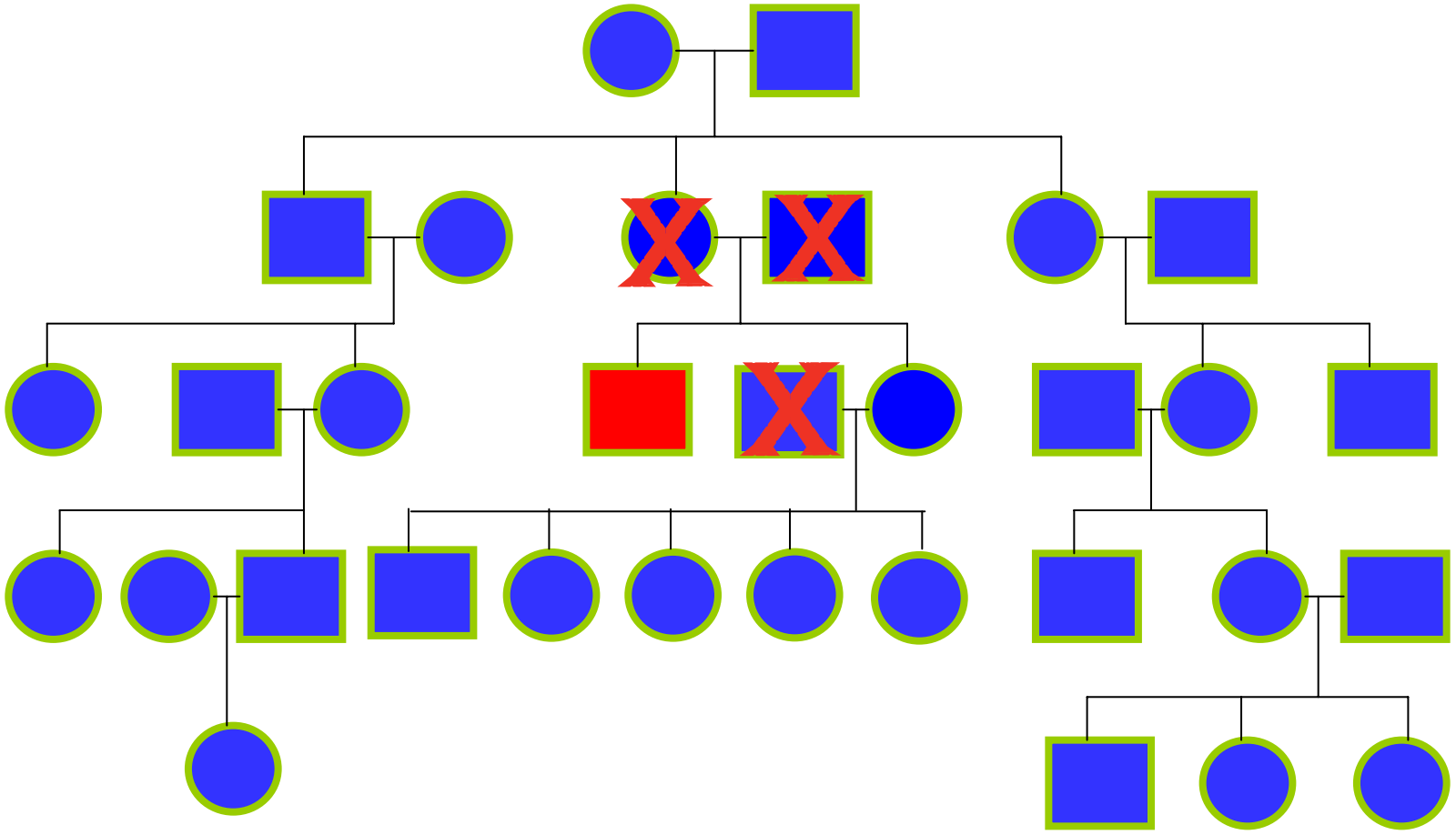
- HOWEVER,
  - Most common Caucasian type
  - Two individuals without references also Caucasians
  - Anthropology not consistent with medical records.



There’s enough nucDNA



Who to use for a Reference?



Conclusions

- mtDNA is a very useful tool for identification
- Larger databases with good quality control are necessary.
- Additional techniques are required to resolve most common mitotypes and aid in identification.

Acknowledgements

- September 11th Phylogenetics is taken from:
  - SM Edson, JP Ross, MD Coble, TJ Parsons, and SM Barritt (2004). Naming the Dead – Confronting the Realities of Rapid Identification of Degraded Skeletal Remains. *Forensic Sci. Rev.* 16: 63.
- STR background information acquired from:
  - Dr. Timothy McMahon, background lectures in STR Analysis for George Washington University DNA Profiling Class.
- LCN-STR work is currently on-going under the auspices of Jodi Irwin, Suni Edson, AFDIL Research Dept. and Mitochondrial DNA Section.

Additional Reading

- LCN Analysis:
  - Whitaker, et al. (2001). A comparison of characteristics of profiles produced with the AMPFlSTR SGM Plus Multiplex system for both standard and low copy number (LCN) STR DNA analysis. Fors. Sci. Intl. 123 (2-3): 215-23.
  - Gill, P. (2002). Role of short tandem repeat DNA in forensic casework in the UK – past, present, and future perspectives. Biotechniques 32(2): 366-8, 370, 372.
  - Dixon, et al. (2006). Analysis of artificially degraded DNA using STRs and SNPs – results of a collaborative European (EDNAP) exercise. For. Sci. Intl. *In press.*

- http://www.cstl.nist.gov/div831/strbase/

Acknowledgements

- Suzanne Barritt
- Christine Boyer
- Amanda Coute
- Heather Thew
- Brion Smith
- Mark Leney
- Franklin Damann
- Michael Coble
- Jim Ross
- Thomas Parsons
- Demris Lee
- Kerry Maynard
- Jennifer O'Callaghan
- Timothy McMahon
- Jim Canik
- Jodi Irwin
- Rebecca Just
- Mark Wadhams

- All the staff, past, present and future, of AFDIL and JPAC-CIL, without whom, none of this would be possible.

AFDIL
Armed Forces DNA Identification Laboratory

JOINT POW/MIA ACCOUNTING COMMAND
JPAC

Contact Information and Disclaimer

- Suni M. Edson

  Suni.Edson@afip.osd.mil

  301-319-0193

- Disclaimer

  The views expressed herein are the those of the author and not necessarily those of the Armed Forces Institute of Pathology, the US Army Surgeon General, nor the Department of Defense.