

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY

APPLIED GENETICS GROUP

Yfiler® Plus Kit

Improved Haplotype Discrimination using
“Rapidly Mutating” Y-STR Markers in a
Large Multiplex Kit

Katherine B. Gettings, PhD

*Future Trends in
Forensic DNA
Technology*

August 6, 2014

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.**

Outline

Y-STR Kits – Overview

Yfiler Plus Evaluation

Sensitivity

Male:Female Mixtures

Rapidly Mutating (RM) Y-STRs – Overview

Utility for close male relatives

Conclusions

What has happened in the past decade...

January 2003: Selection of core Y-STR loci

June 2003: “Full” Y-chromosome sequence available

- over 700 Y-STR loci identified (only ~20 in 2000)

Commercial Y-STR kits released

- ~~2001-2003: Y PLEX 6,5,12~~
- 09/03: PowerPlex Y
- 12/04: Yfiler
- 06/12: PPY23
- **2014: Yfiler Plus**

Many population studies performed, online databases generated with thousands of Y-STR haplotypes

Forensic casework demonstrations showing value of Y-STR testing along with court acceptance

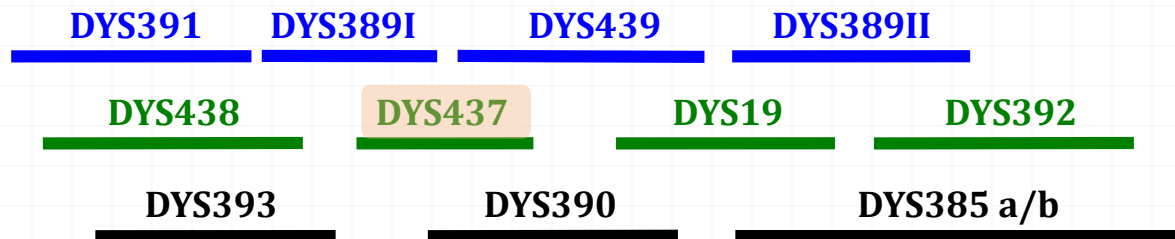
Some renewed interest in Y-STRs to aid familial searching

STR Marker Layouts for Y-STR Kits

100 bp 200 bp 300 bp 400 bp

2003

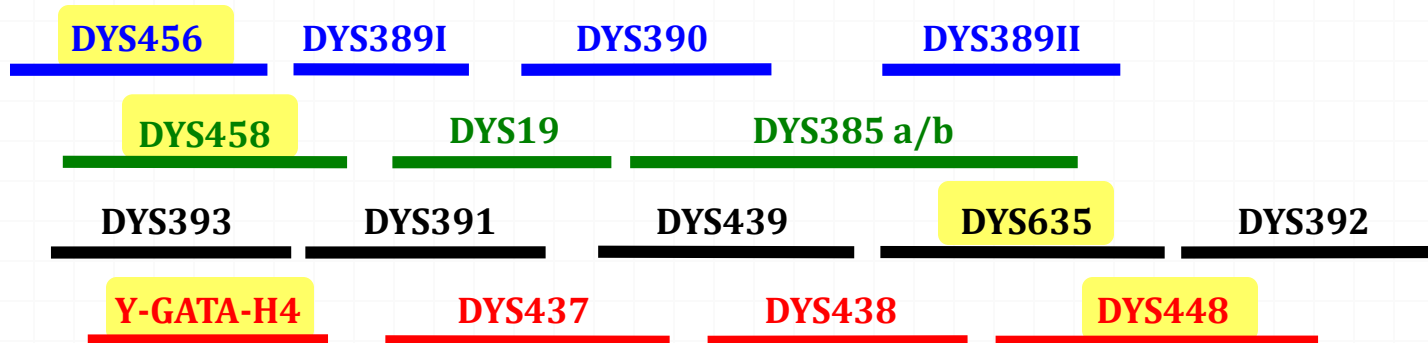
PowerPlex Y



12plex
(4-dye)

2004

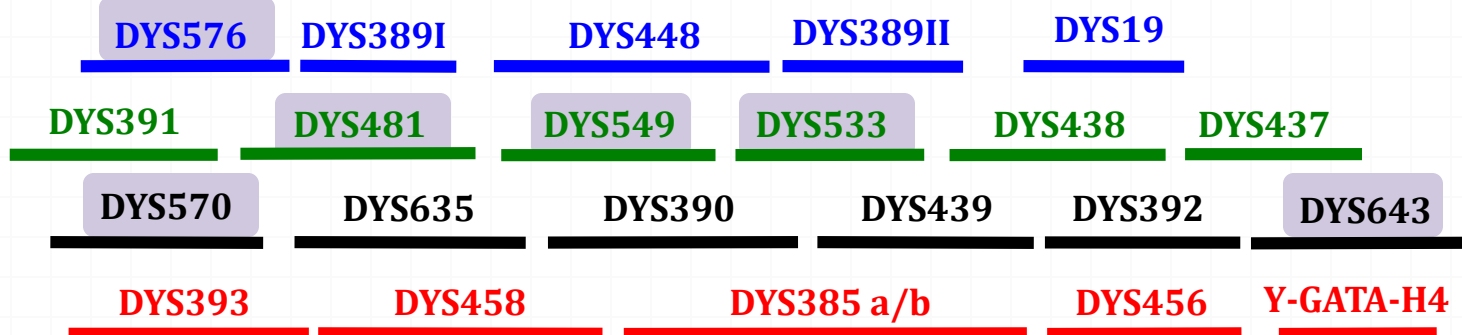
Yfiler



17plex
(5-dye)

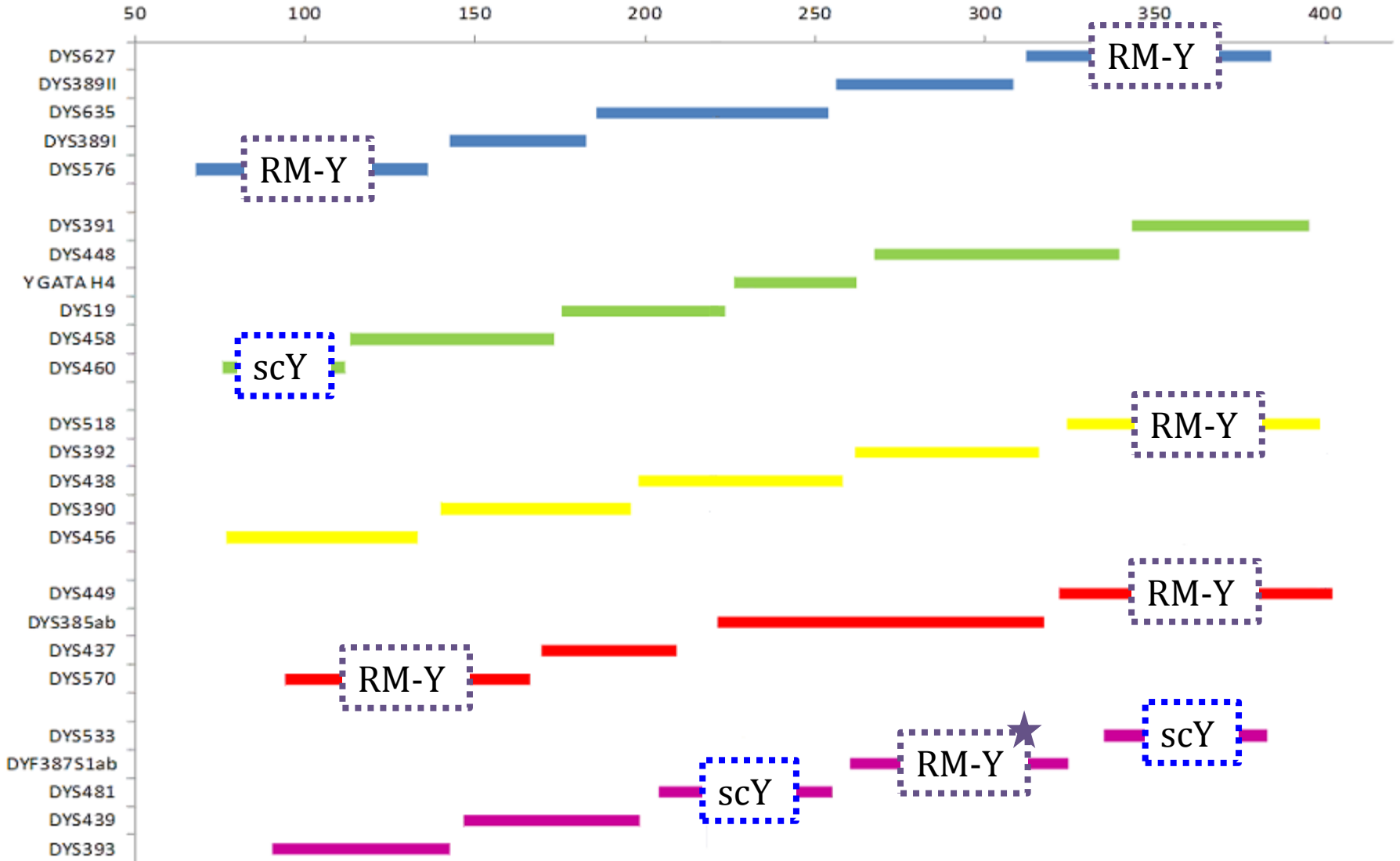
2012

PowerPlex Y23

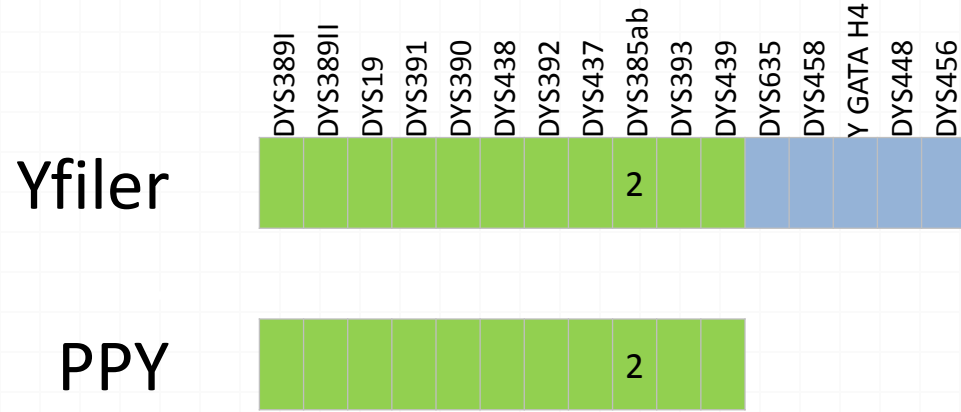


23plex
(5-dye)

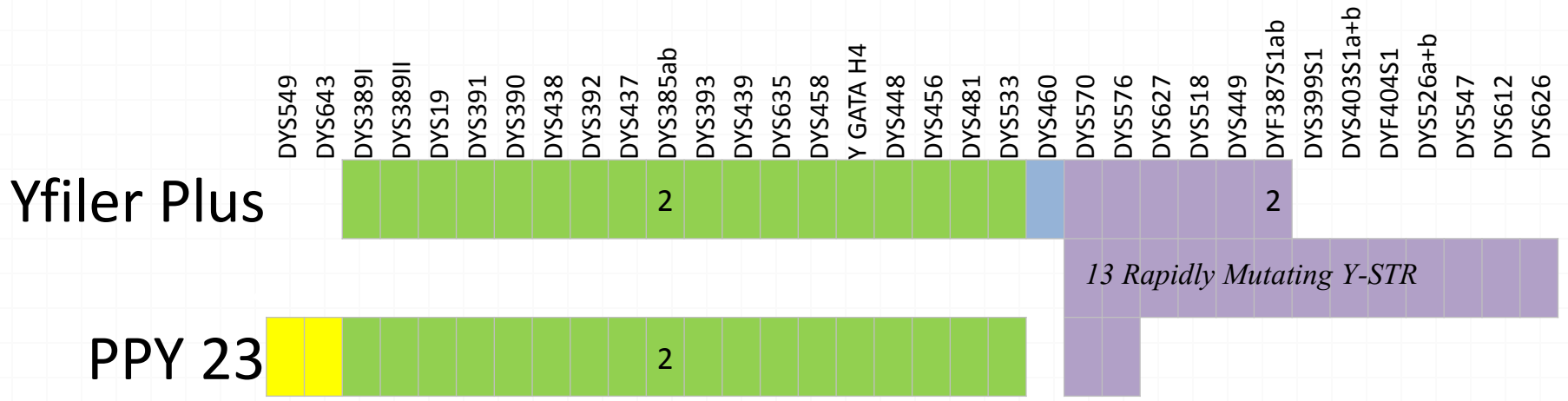
Yfiler Plus Kit



STR Marker Layouts for Y-STR Kits



STR Marker Layouts for Y-STR Kits



Yfiler-Plus Kit Performance

Sensitivity

Sensitivity Experiment

Full (25 μL) and Half (12.5 μL) reaction volumes:

0.5 ng

0.25 ng

0.125 ng

0.0625 ng

0.03125 ng

Sensitivity Experiment

Marker	FULL REACTION VOLUME (25 ul)									
	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	1127	3639		
DYS460		463	304	129	263	328	82	786		
DYS458		404	51	261	466	581	10			
DYS19		173	220	143	335	525	17			
GATA_H4	97		159	226	482	153	1			
DYS448	91		154	204	500	196				
DYS391	342		231	51	1169	721				
DYS456	330	186	269	115	732	241				
DYS390		232	123	87	225	331				
DYS438	144	166	207	205	557	622				
DYS392		209	53	129	700	156				
DYS518	232	118	145	185	638	411				
DYS570		92	163	300	987	327				
DYS437	83	248	364	152	749	485	1429			
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		

Data coded
"pass/fail"
at 175 RFU
(per AB recommendation
for 3500xl)

Validation may
support a different
threshold

≥ 175 RFU
 < 175 RFU

Sensitivity Experiment

Marker	HALF REACTION VOLUME (12.5 ul)									
	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	244		331
DYS458		394	262	2099	2616	1013	4972	22		
DYS19		835	351	1283	1886	1464	2492	24		
GATA_H4		245	475	1690	1766	1274	3980	1		
DYS448			594	1842	1131	1611	4011			
DYS391	199	398	1140	398	2861	1632	3359			
DYS456	122		121	920	1280	1391	2803			
DYS390			214	673	771	1428	2059			
DYS438	321	155	656	717	2163	1269	2201			
DYS392	153		1306	446	1853	858	2242			
DYS518	367	262	844	280	771	1661	1657			
DYS570	196	592	376	1165	1754	1551	3068	2200		
DYS437	145	1152	404	325	1424	1582	6373	1199	6260	
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

Data coded
"pass/fail"
at 175 RFU
(per AB recommendation
for 3500xl)

Validation may
support a different
threshold

≥ 175 RFU
< 175 RFU

Yfiler-Plus Kit Performance

  Mixtures

High Level Female Experiment

Looking for:
Artifacts
Inhibition

Three male samples
combined with same female

Input female DNA constant:
200 ng

Four levels of input male
DNA approx. 1 ng - 0.05 ng

Male only samples
for comparison

24 samples total

200 : 1



200 ng + 1 ng

700 : 1



200 ng + 0.3 ng

2000 : 1



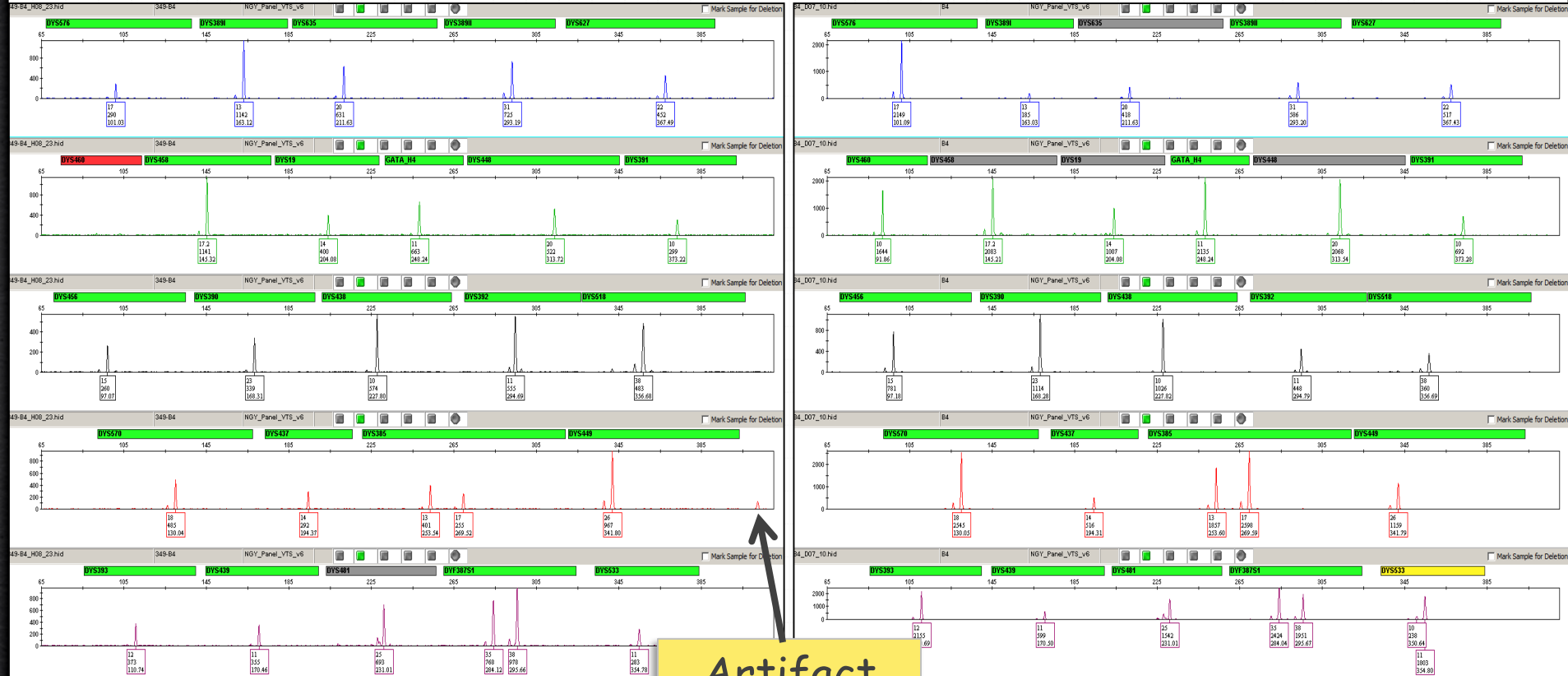
200 ng + 0.1 ng

4000 : 1



200 ng + 0.05 ng

High Level Female Experiment Male Sample A



4000 : 1

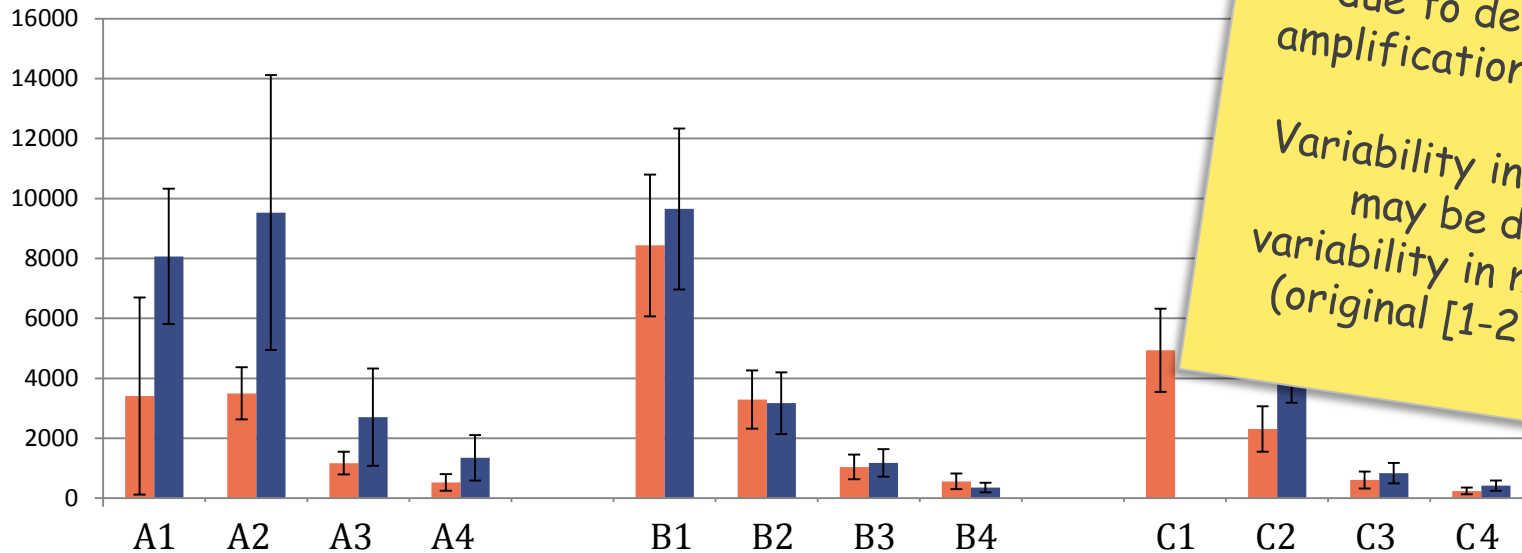
(approx. 200 ng female + 0.05 ng male)



(approx. 0.05 ng)

High Level Female Experiment

Average peak height per allele \pm 1 SD

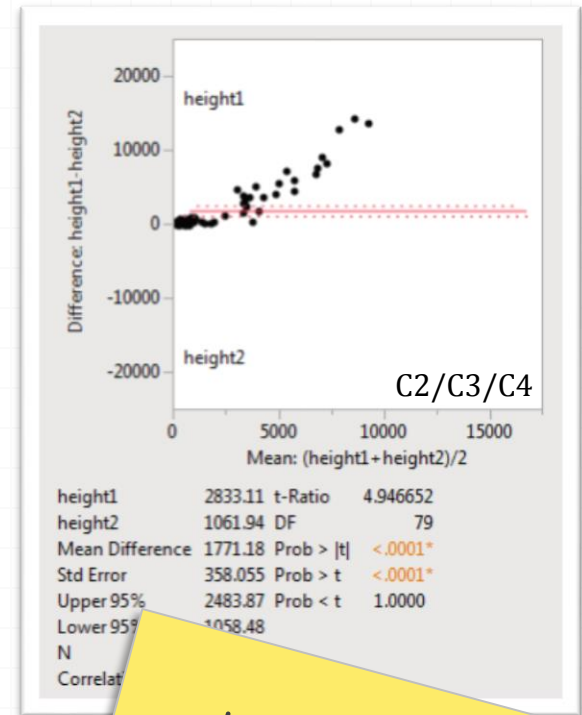
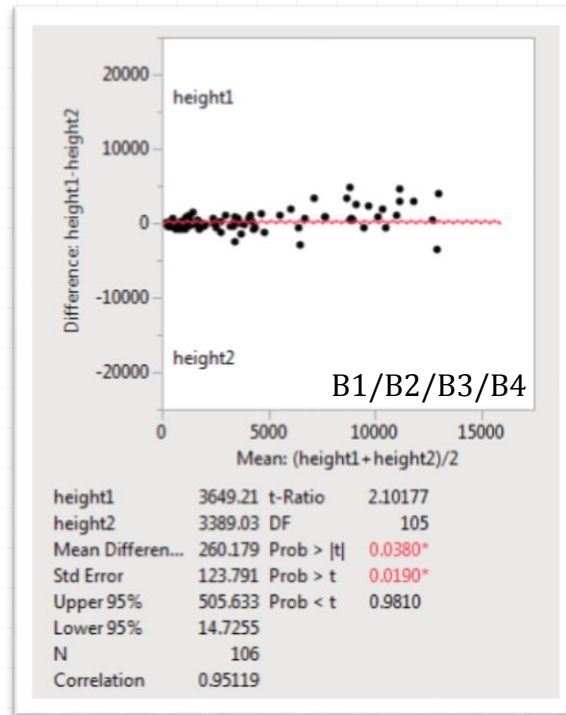
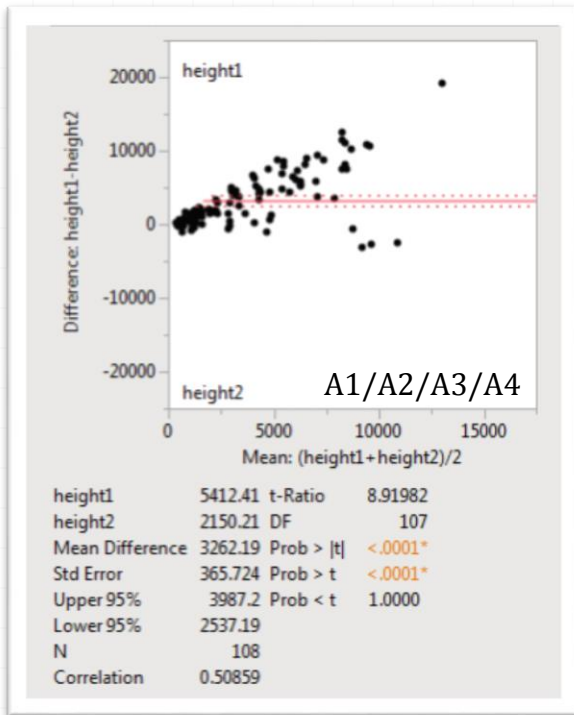


*Female inhibits male,
due to decrease in
amplification efficiency*

*Variability in inhibition
may be due to
variability in male input
(original [1-2 ng/uL])*

A1/B1/C1	200 ng female + 1 ng male	1ng male only
A2/B2/C2	200 ng female + 0.5 ng male	0.5 ng male only
A3/B3/C3	200 ng female + 0.1 ng male	0.1 ng male only
A4/B4/C4	200 ng female + 0.05 ng male	0.05 ng male only

High Level Female Experiment



A1/B1/C1	200 ng female + 1 ng male	1 ng male on
A2/B2/C2	200 ng female + 0.5 ng male	0.5 ng male
A3/B3/C3	200 ng female + 0.1 ng male	0.1 ng male
A4/B4/C4	200 ng female + 0.05 ng male	0.05 ng m

Locus by locus comparison of PH with and without female DNA at all input levels for each sample

200 : 1



700 : 1



2000 : 1



4000 : 1



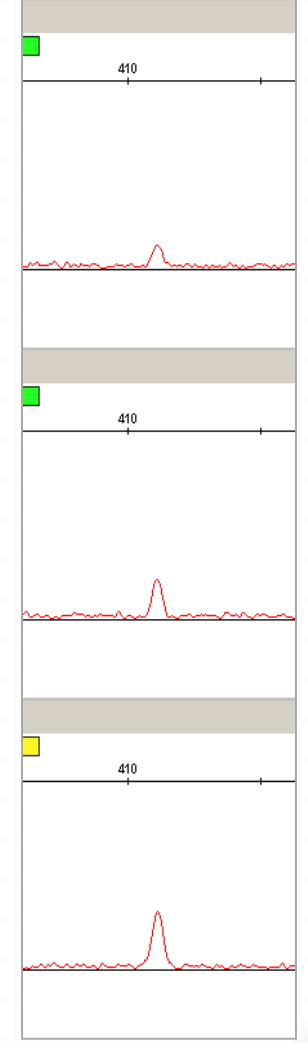
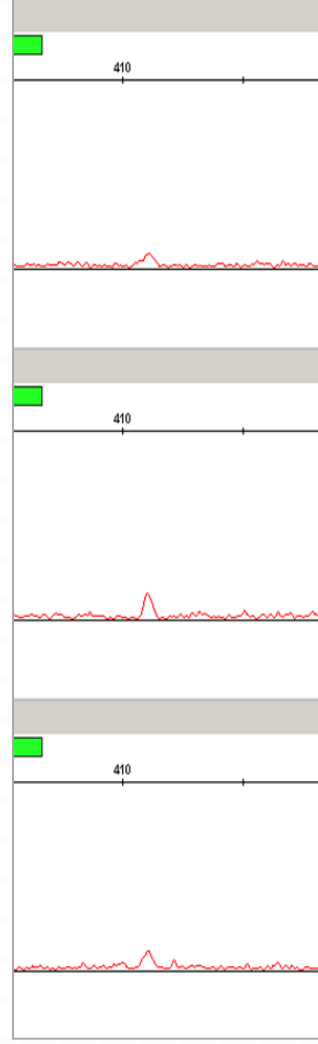
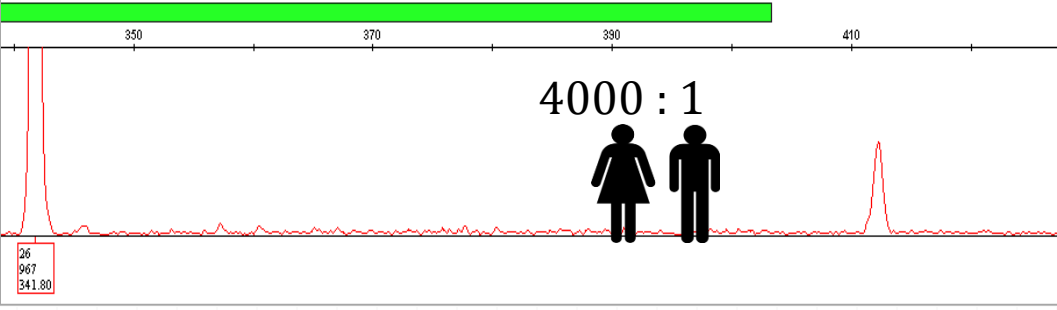
Known artifact
"TAZ 413"
related to presence
of female DNA

But not directly
affected by
female : male

26
4291
341.81

26
763
341.80

26
967
341.80

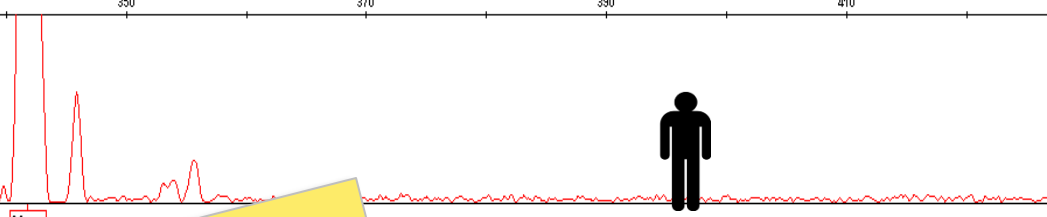


B1

NGY_Panel_VTS_v6

Male A

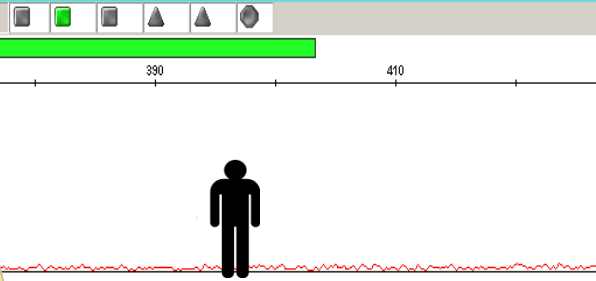
350 370 390 410



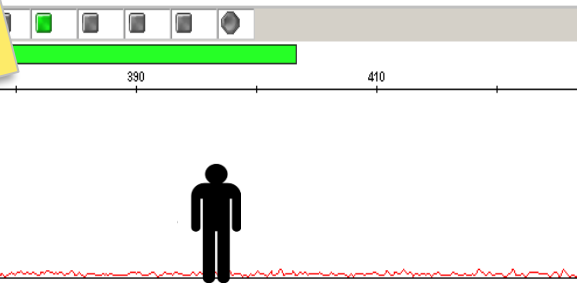
26
12420
341.79

v6

390 410



390 410

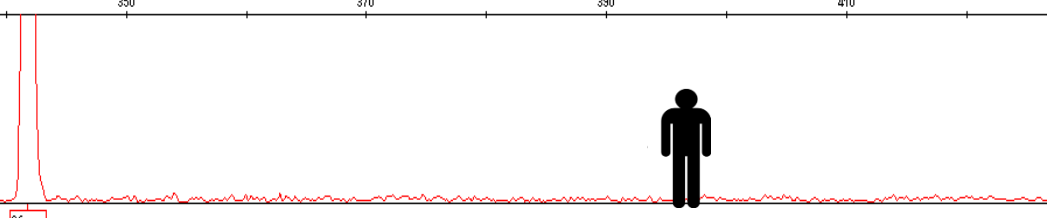


26
20009
341.80

B4

NGY_Panel_VTS_v6

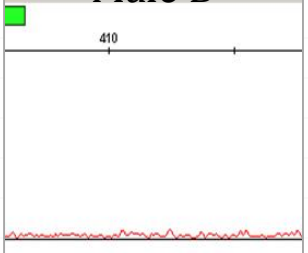
350 370 390 410



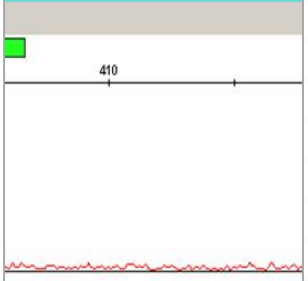
26
1159
341.79

Male B

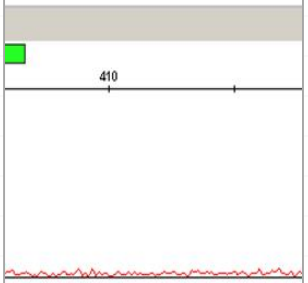
410



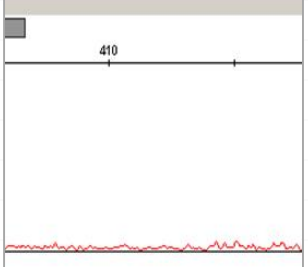
410



410

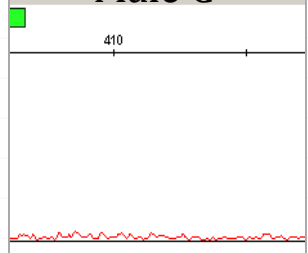


410

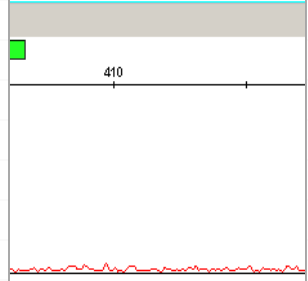


Male C

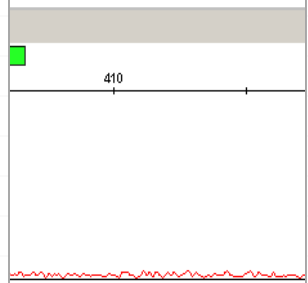
410



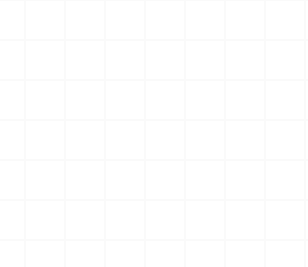
410



410



410



Known artifact
"TAZ 413"
related to presence
of female DNA

But not directly
affected by
female : male

Male:Female Mixture Summary

A high level of female DNA does inhibit Y-STR amplification
(as expected)

No reproducible artifacts were seen within loci

One reproducible, known artifact was seen outside of bin

Related to presence of female DNA

Not diagnostic of female DNA presence or level

Performance with Unrelated Males

NIST U.S. Population Samples

Y-STRs for haplotype differentiation

Discrimination
capacity

measures the number
of unique haplotypes
in a population

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

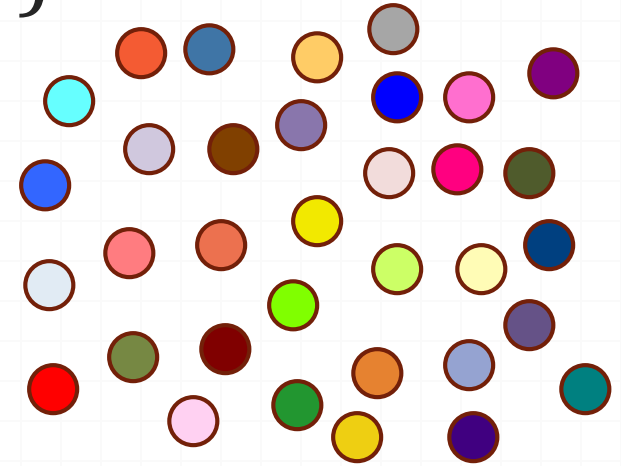
of Haplotypes

Population size

Results so far (582 individuals)

582
Males

Yfiler Plus
27 Y-STRs

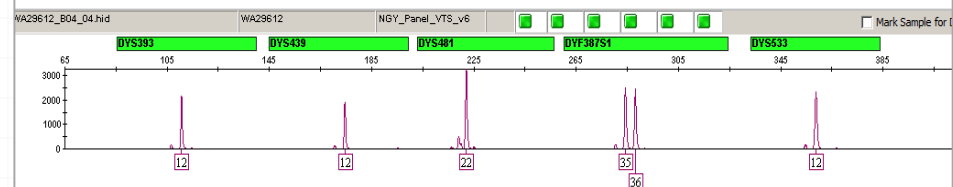
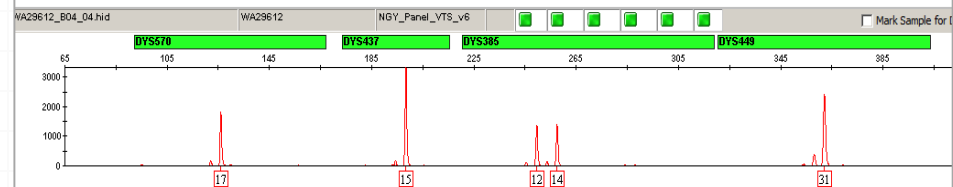
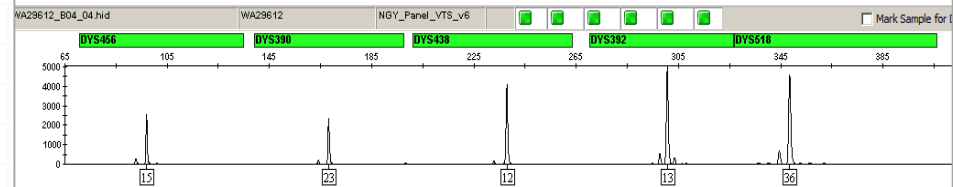
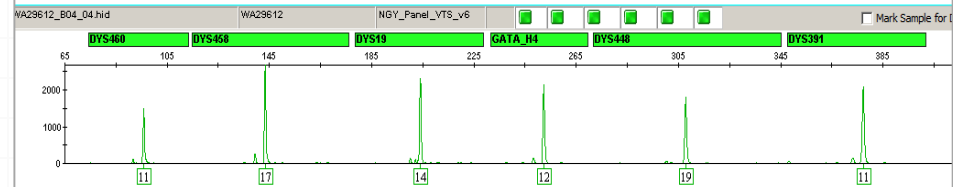
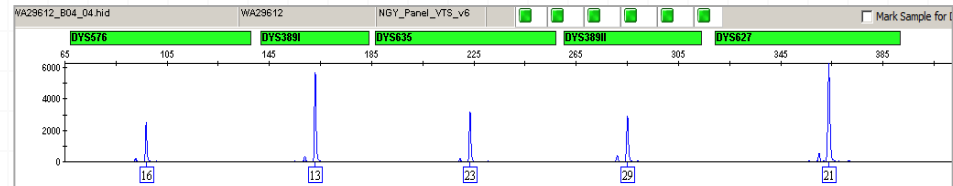
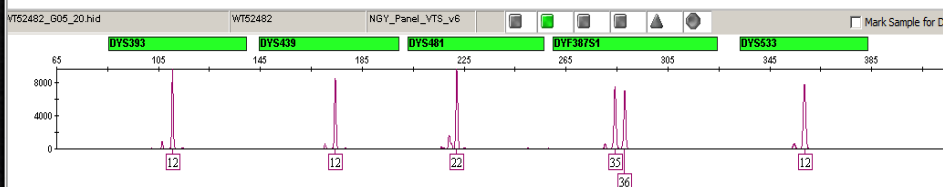
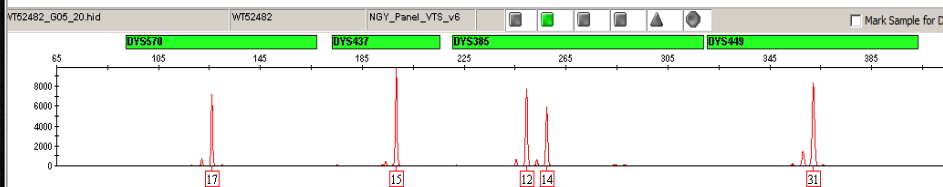
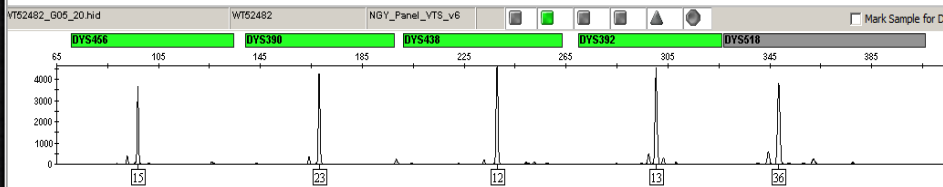
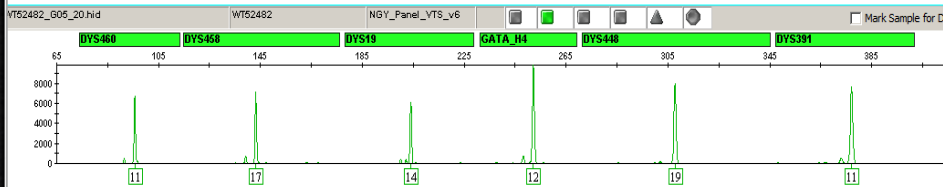
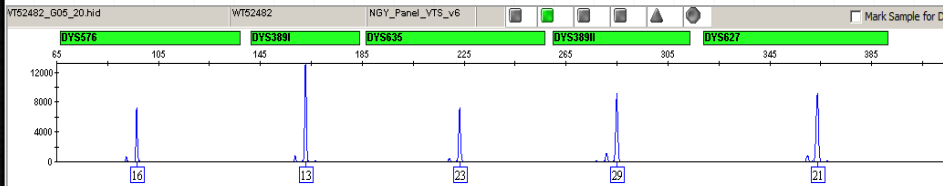


580 Unique Haplotypes



1 Shared Haplotype

DC = 0.998



**Two Caucasians
Same Haplotype at Yfiler Plus Loci**

Kinship stats over 24 auSTR

	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



N = 948 males

PowerPlex Y

haplotypes

816

discrimination capacity

0.8608

times haplotype
observed

PPY
(12 loci)

1	751
2	42
3	12
4	4
5	2
6	2
7	.
8	1
9	.
10	.
11	1
12	.
13	.
14	.
15	.
16	.
17	.
18	.
19	.
20	1

Number of unique and shared haplotypes observed with various combinations of Y-STR loci across 948 U.S. population samples

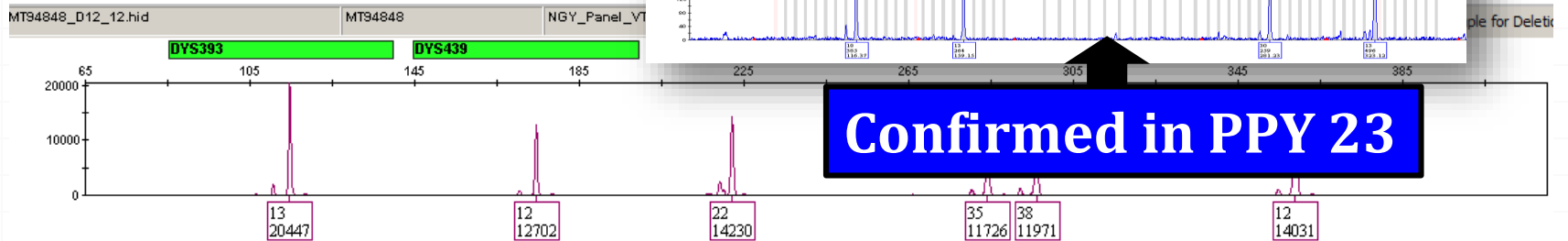
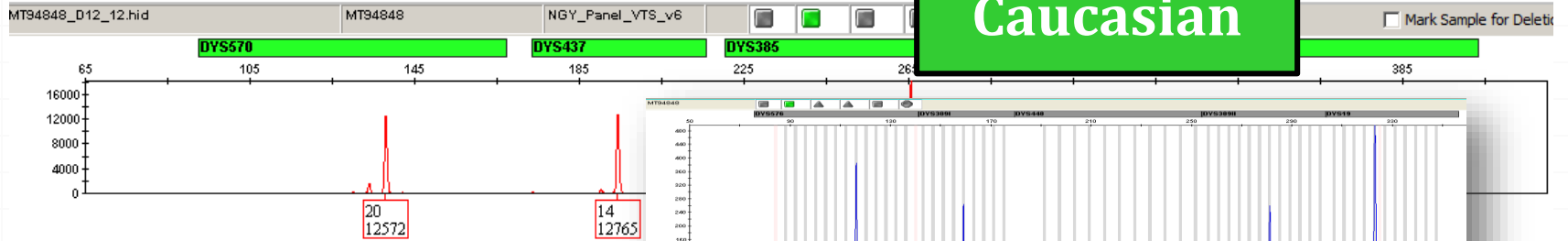
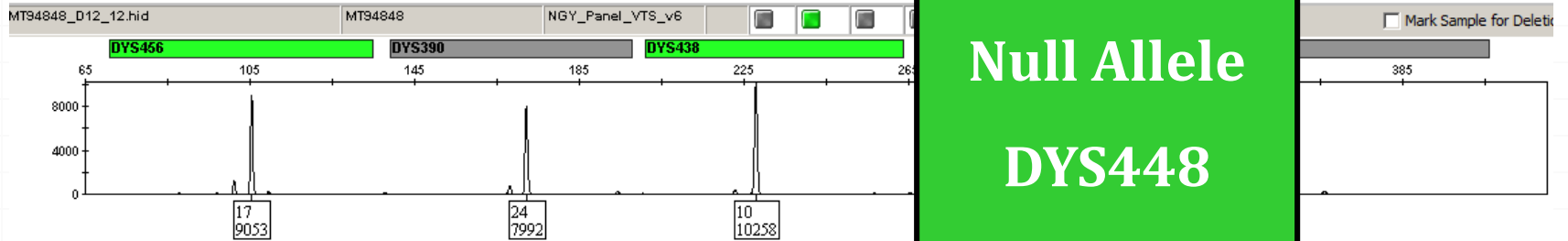
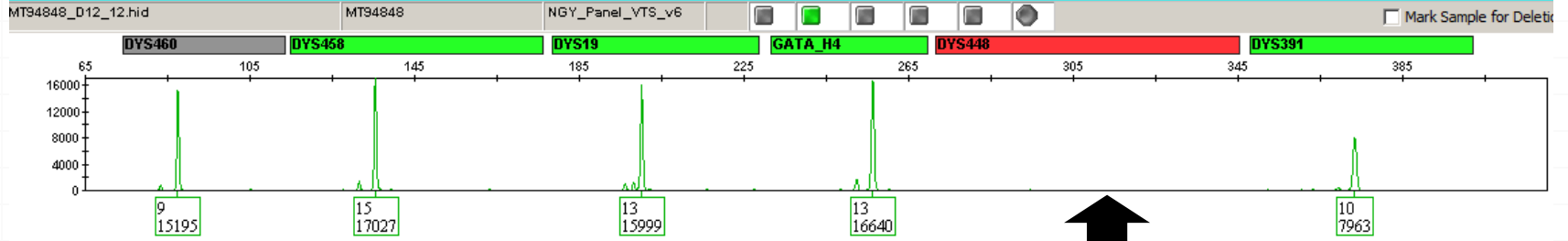
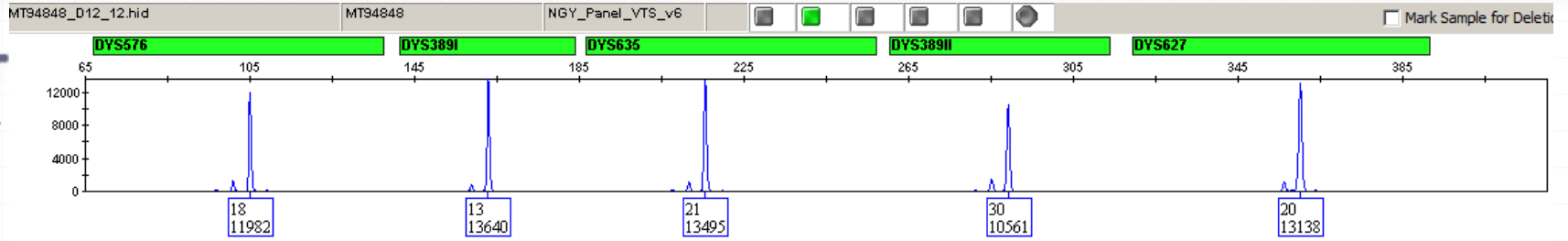
N = 948 males

	Yfiler	New Loci*	Yfiler Plus*
# haplotypes	930	945	946
discrimination capacity	0.9810	0.9842	0.9979
# times haplotype observed	Yfiler (17 loci)	New Loci* (10 loci)	Yfiler Plus (27 loci)
1	916	918	944
2	11	15	2
3	2	.	.
4	1	.	.
5	.	.	.
6	.	.	.
7	.	.	.
8	.	.	.
9	.	.	.
10	.	.	.
11	.	.	.
12	.	.	.
13	.	.	.
14	.	.	.
15	.	.	.
16	.	.	.
17	.	.	.
18	.	.	.
19	.	.	.
20	.	.	.

The new loci alone perform slightly better than Yfiler

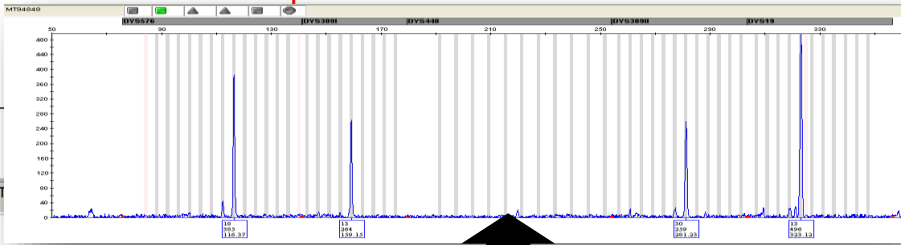
Interesting Samples

NIST U.S. Population Samples

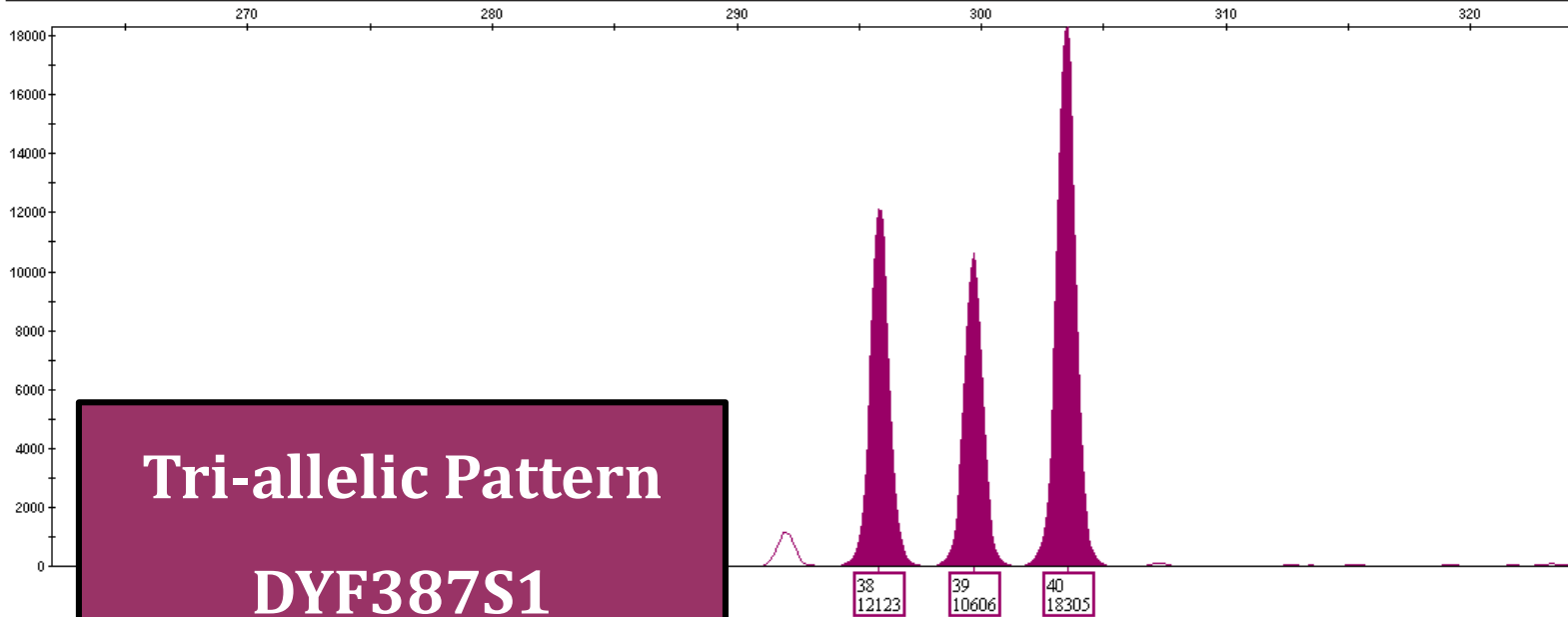


**Null Allele
DYS448
Caucasian**

Confirmed in PPY 23



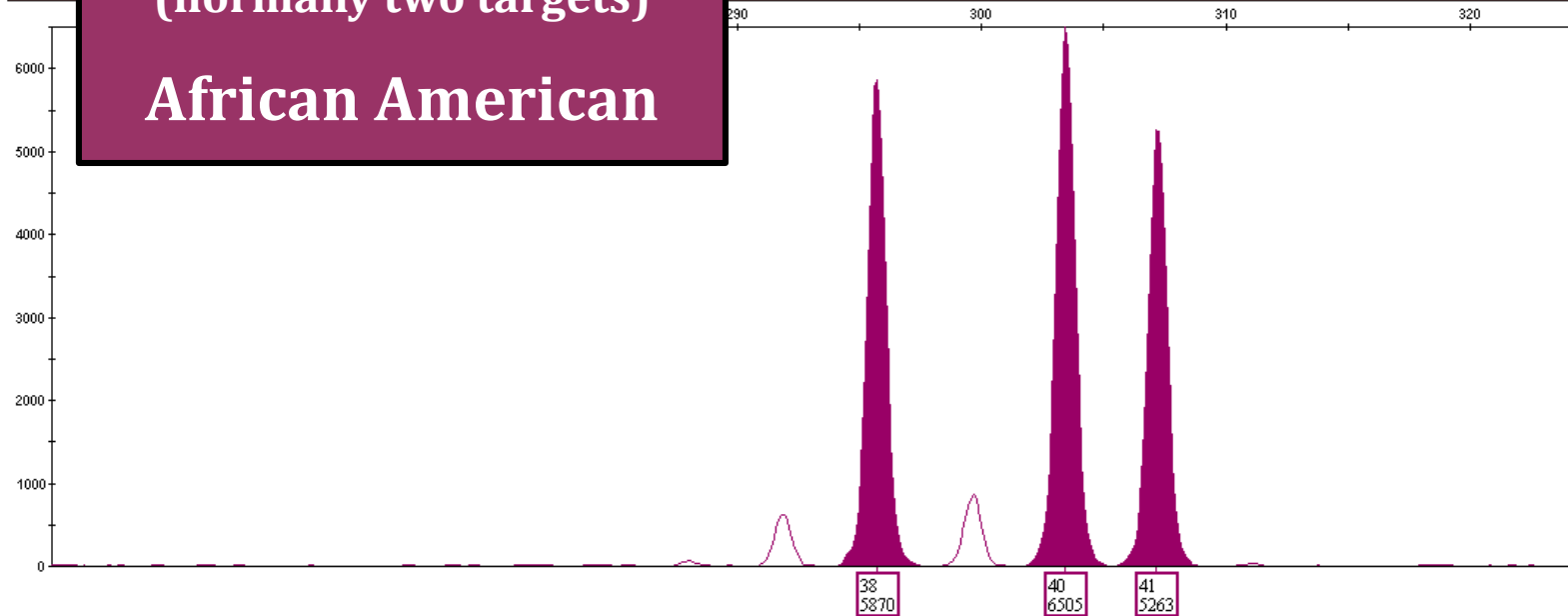
DYF387S1



Tri-allelic Pattern
DYF387S1
(normally two targets)
African American

PT84213_B07_

DYF387S



PT83899_C11_08.hid

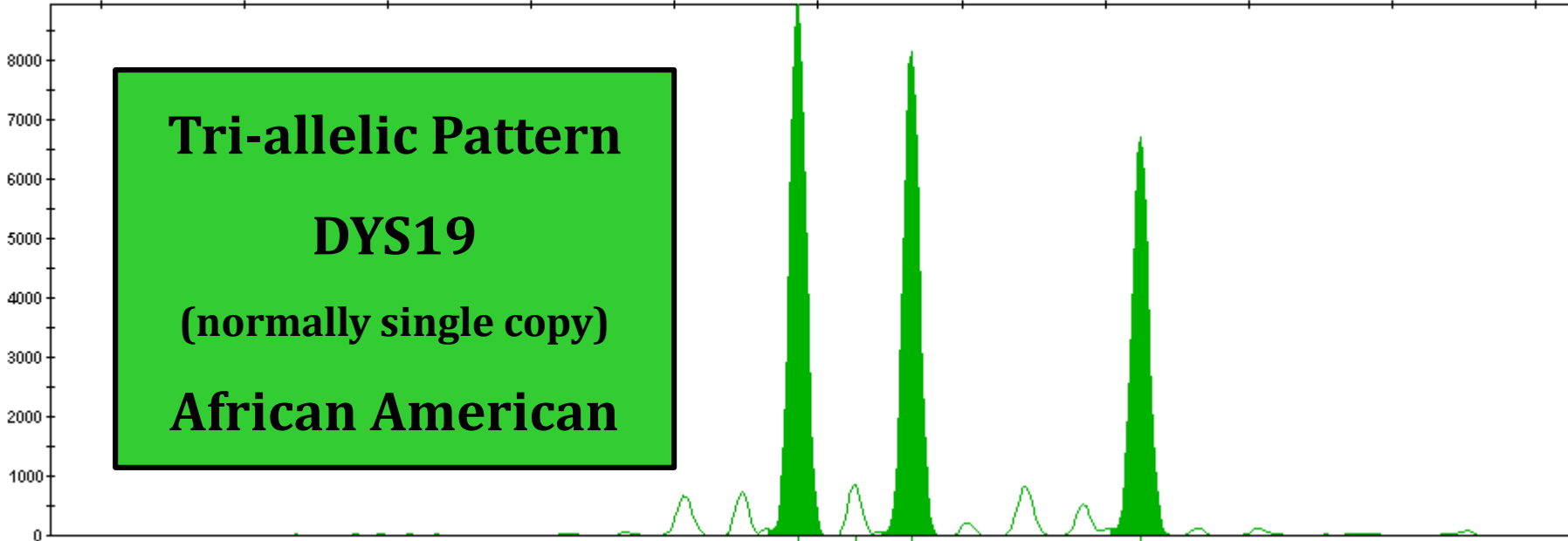
PT83899

NGY_Panel_VTS_v6



DYS19

190 200 210 220 230



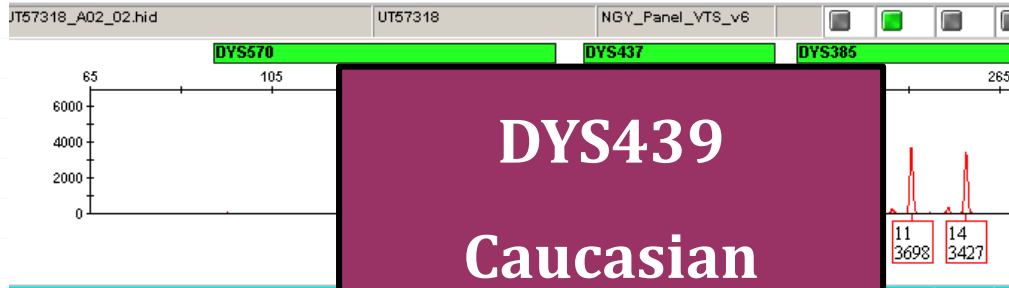
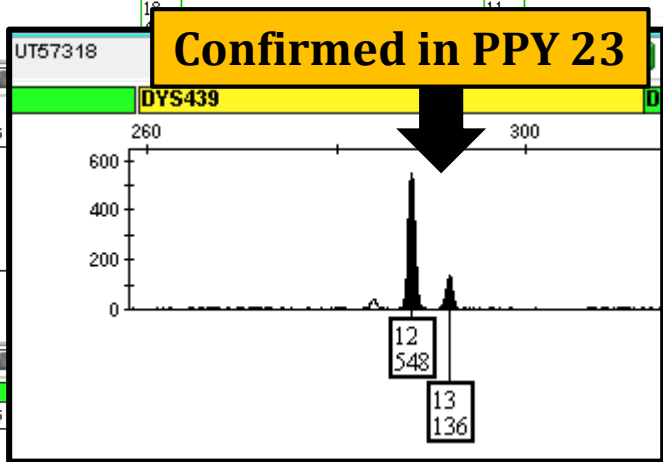
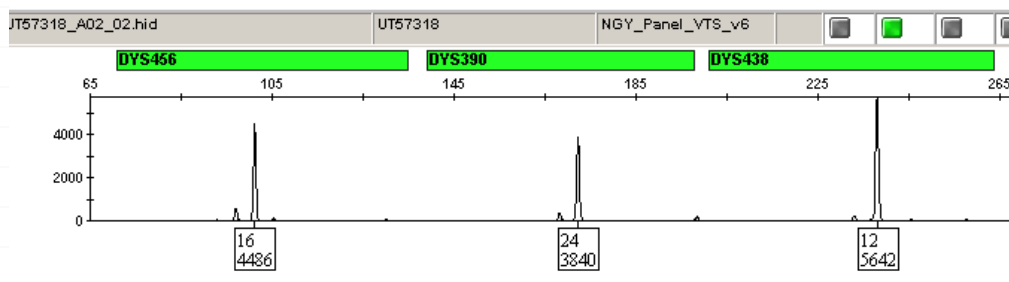
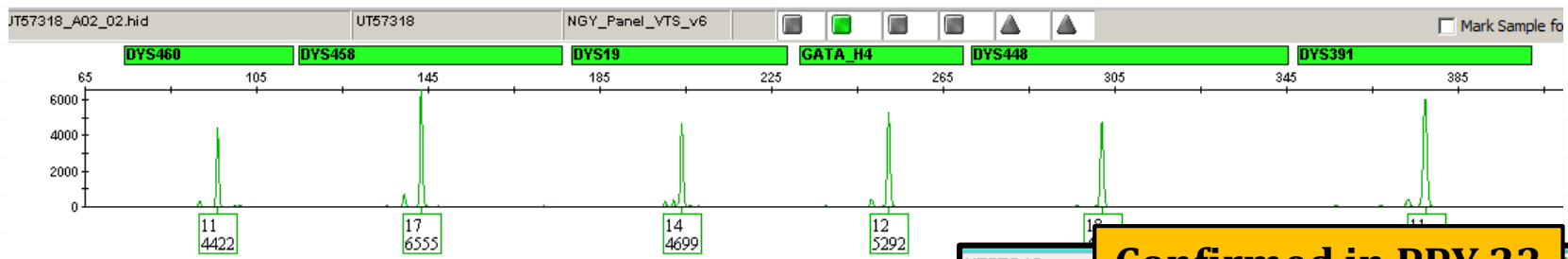
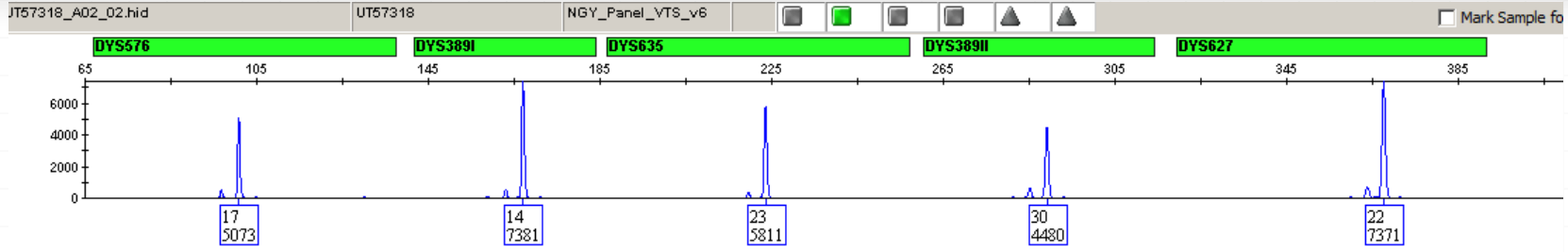
Tri-allelic Pattern
DYS19
(normally single copy)
African American

14
204.33
8952

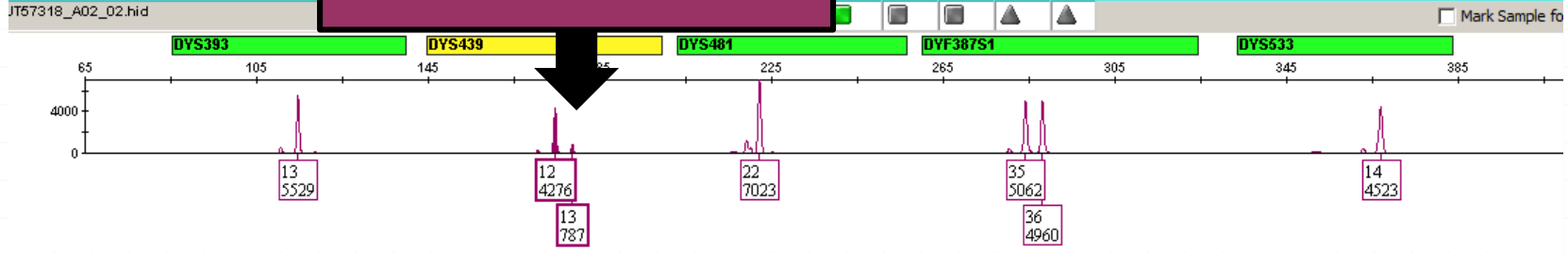
14.2
206.30
861

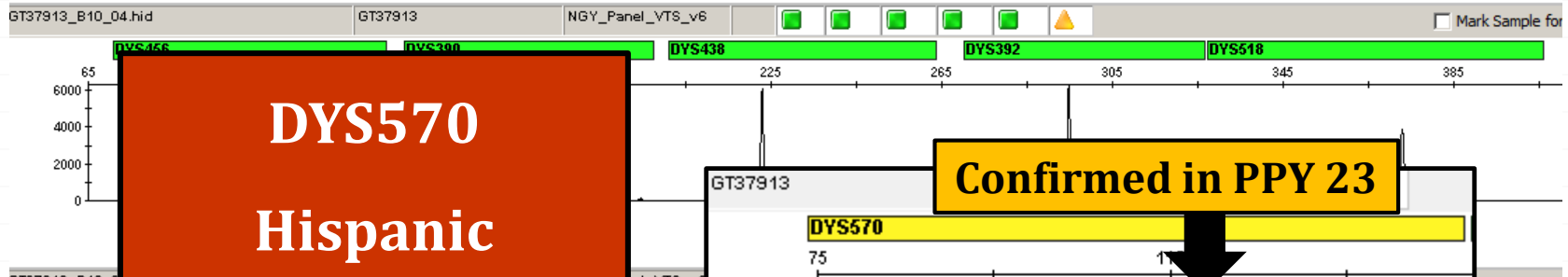
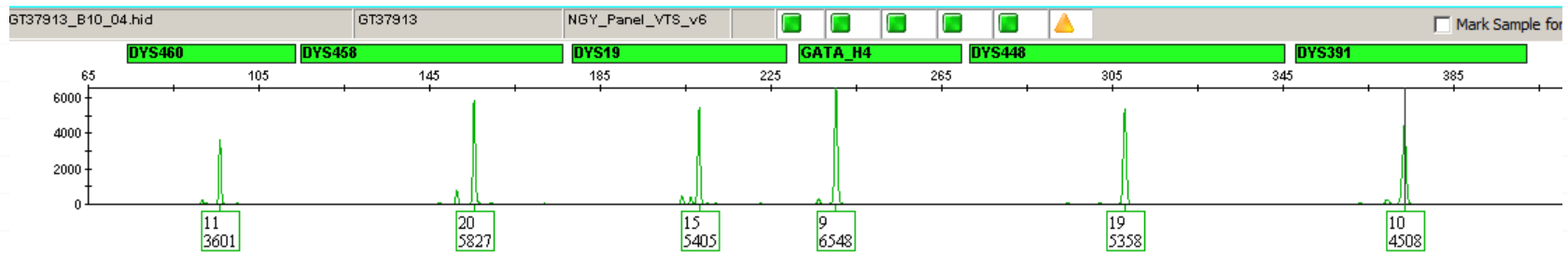
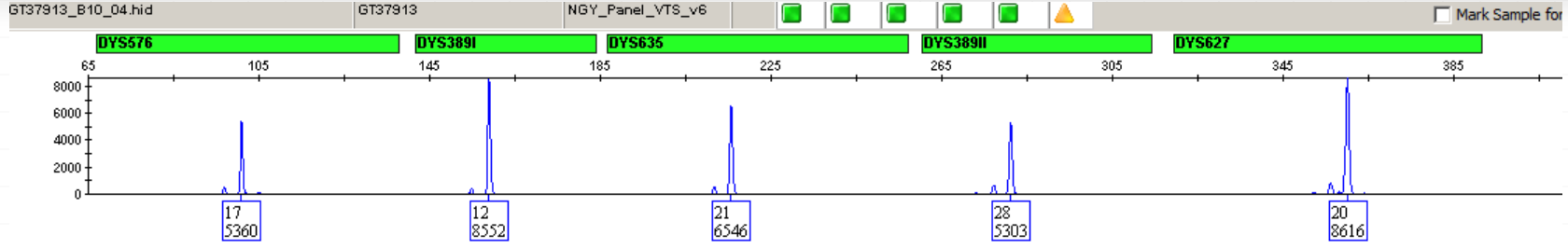
15
208.27
8149

17
216.24
6710

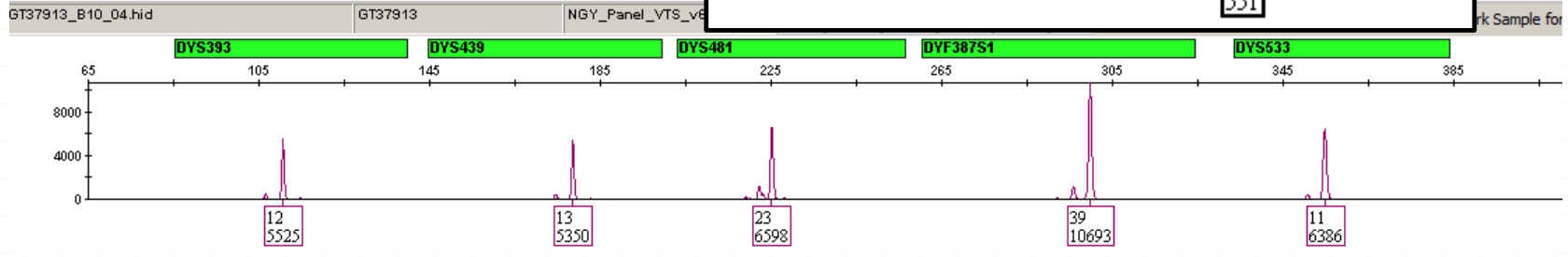
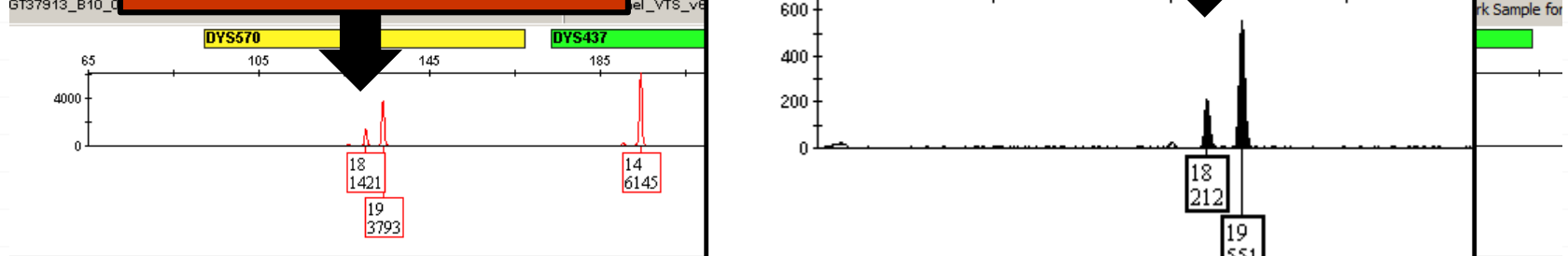
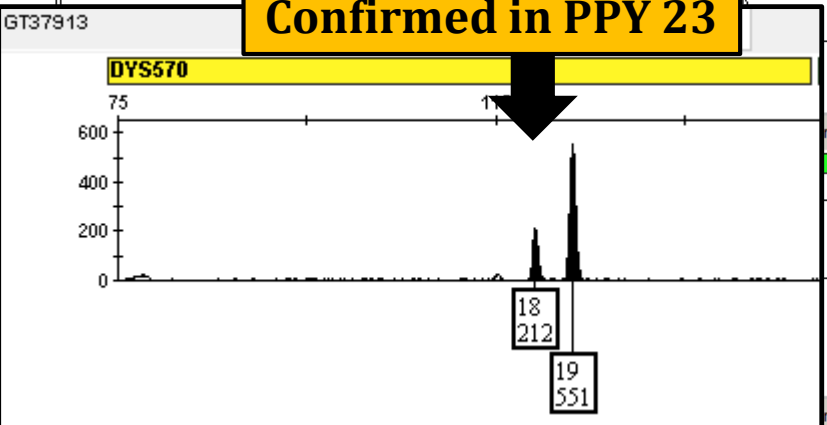


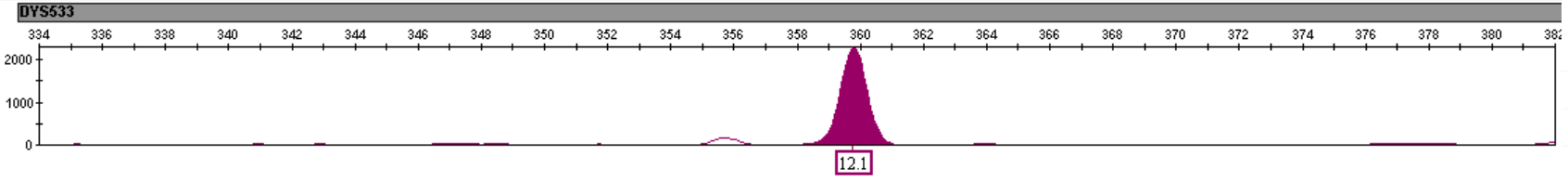
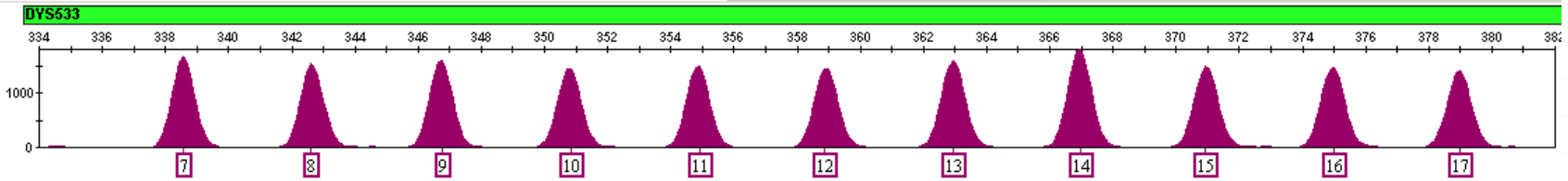
DYS439
Caucasian





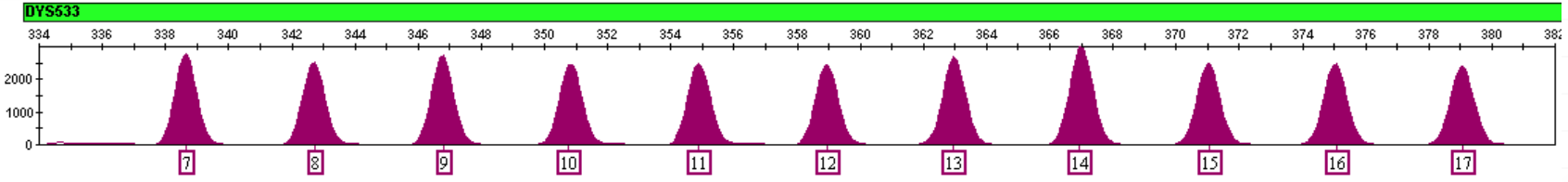
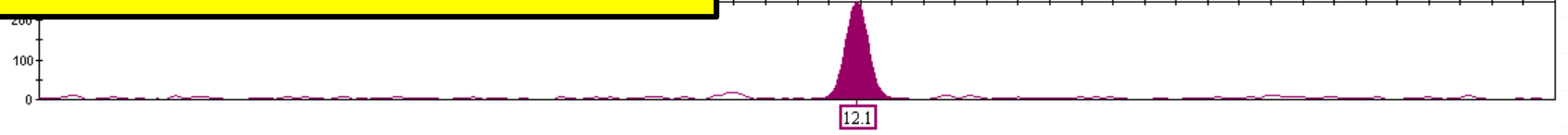
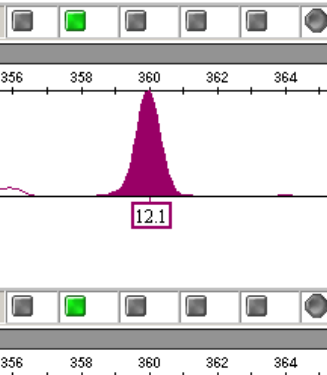
**DYS570
Hispanic**

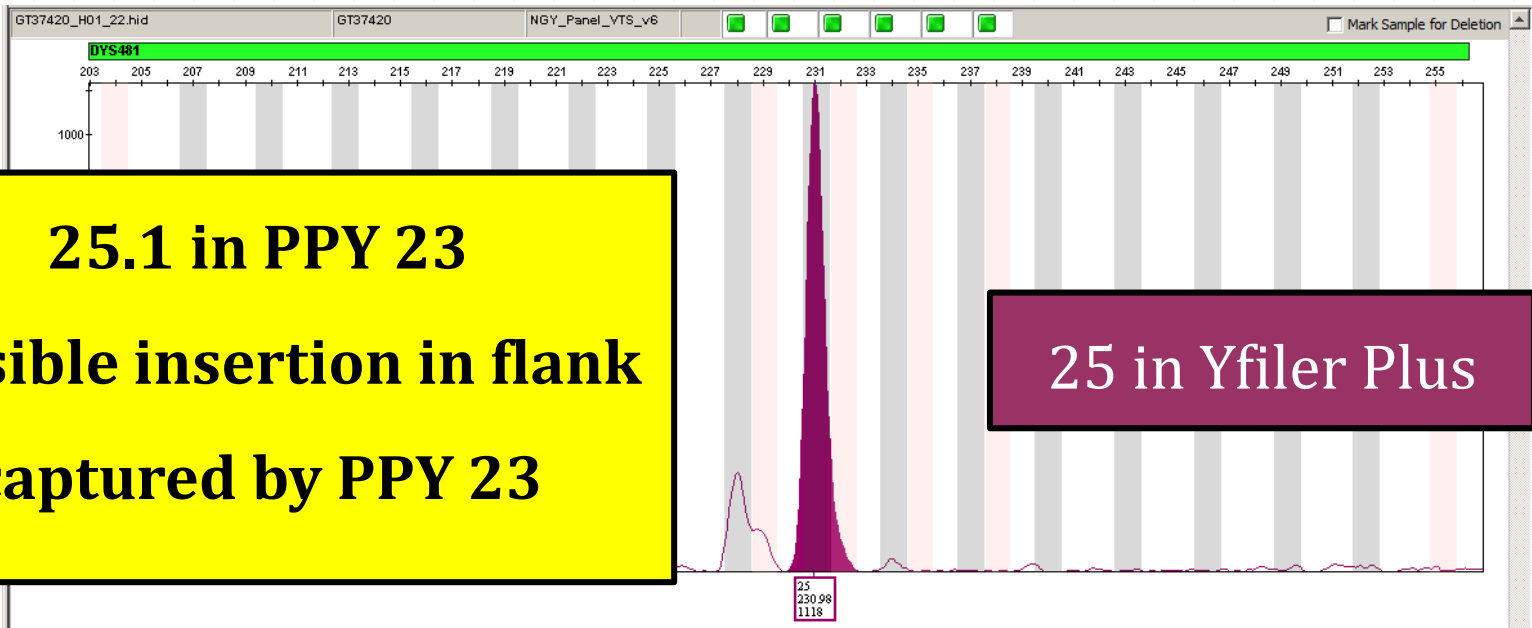




12 in PPY 23
Possible insertion in flank
captured by Yfiler Plus

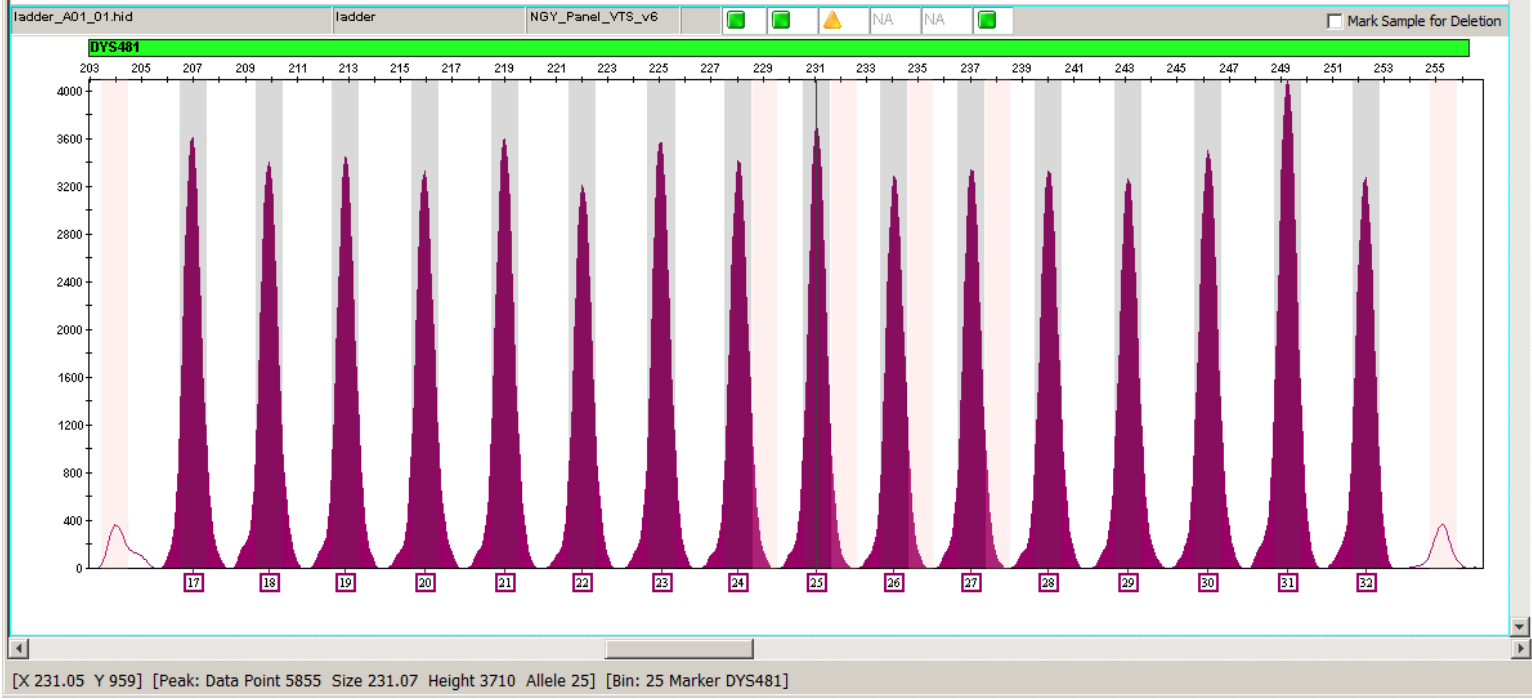
Off-Ladder Allele
DYS533 12.1
Caucasian





25.1 in PPY 23
Possible insertion in flank
captured by PPY 23

25 in Yfiler Plus



Disadvantages of the Y-Chromosome

Loci are not independent →

random match probabilities cannot be generated with product rule,
must use haplotypes (combination of alleles observed at all tested loci)

Paternal lineages have same Y-STR haplotype (barring mutation) →

fathers, sons, brothers, uncles, paternal cousins cannot be
distinguished from one another

∴ not as informative as autosomal STR results

More like addition
 $10 + 10 + 10 = 30$

than multiplication
 $10 \times 10 \times 10 = 1,000$

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 H el ene V ezina,¹⁰ Hans Knoblauch,¹¹ R udiger Lessig,¹² Lutz Roewer,¹³ Rafal Ploski,¹⁴ Tadeusz Dobosz,¹⁵ Lotte Henke,¹⁶ J urgen Henke,¹⁶ Manohar R. Furtado,² and Manfred Kayser^{1,*}



Manfred Kayser

Kayser lab

RM Y-STRs
-13 loci
-3 multiplexes

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



ELSEVIER

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

Y-STRs with higher mutation rate may distinguish close male relatives

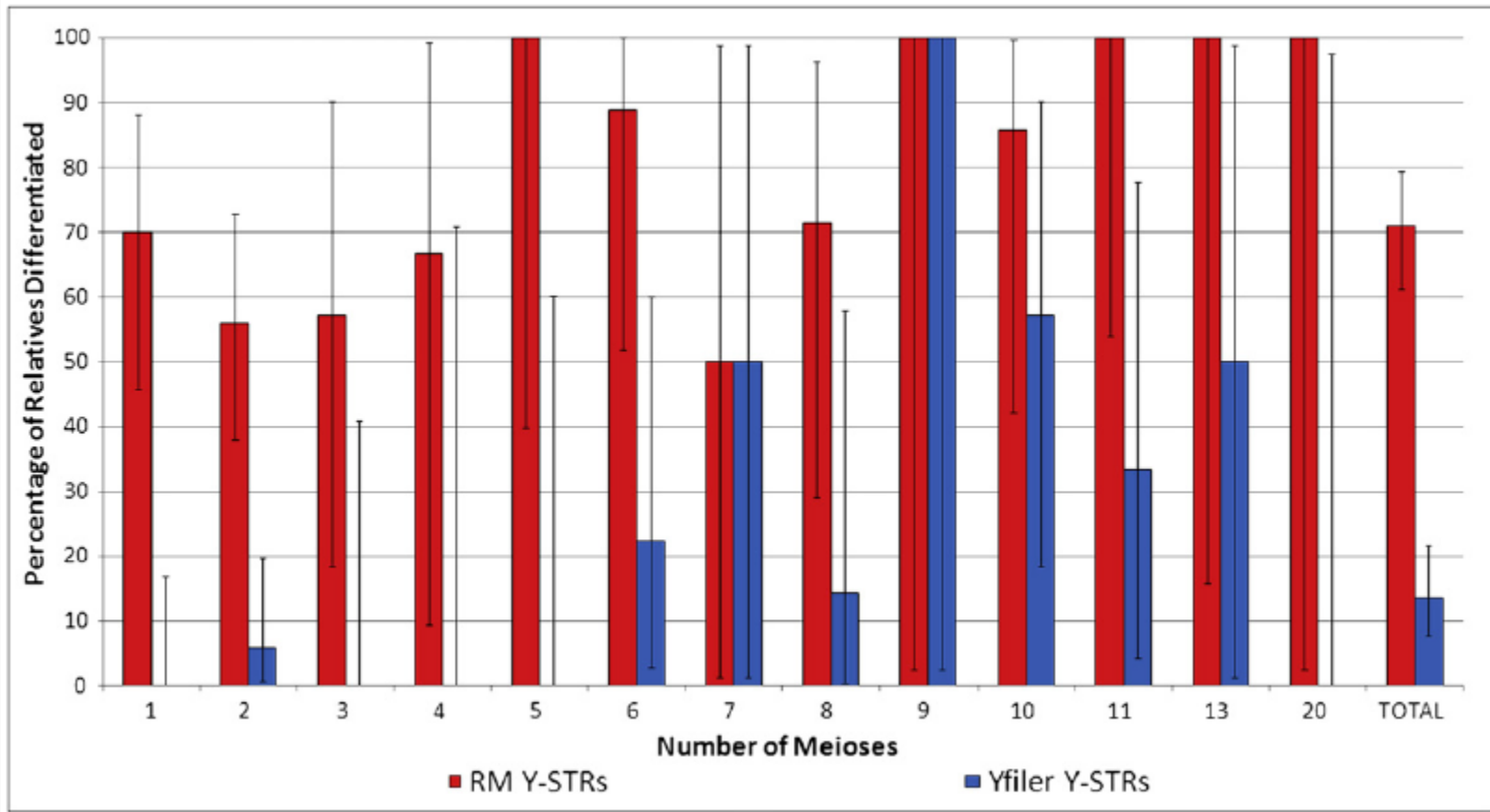
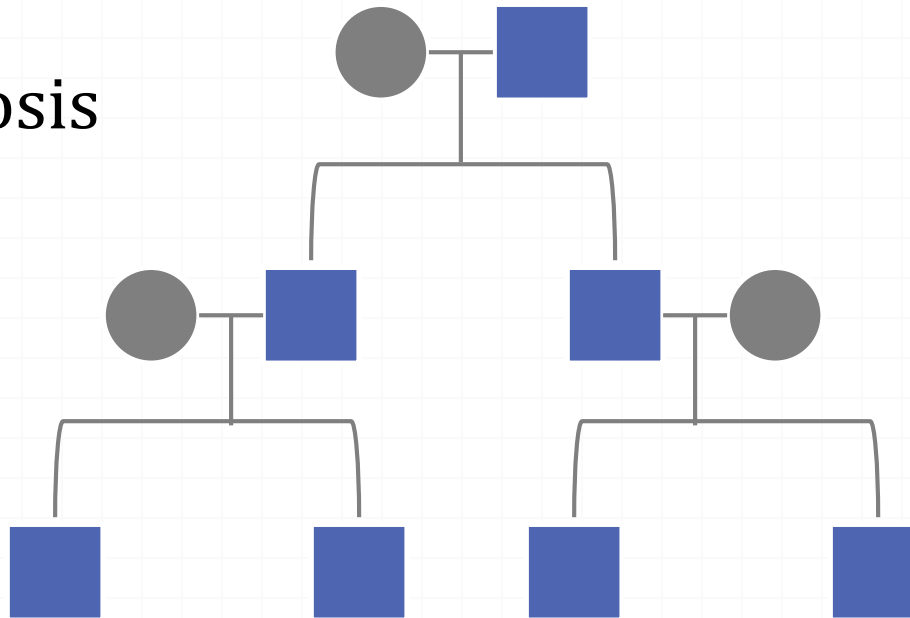


Figure 4. Male Relative Differentiation with Newly Identified 13 RM Y-STRs and Commonly Used 17 Yfiler Y-STRs

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.

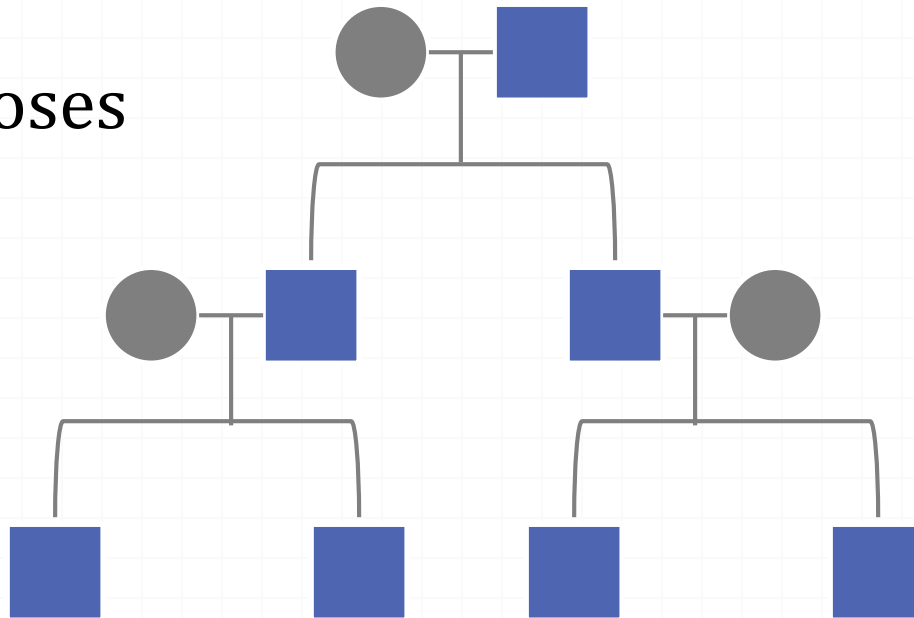
Y-STRs with higher mutation rate may distinguish close male relatives

One meiosis



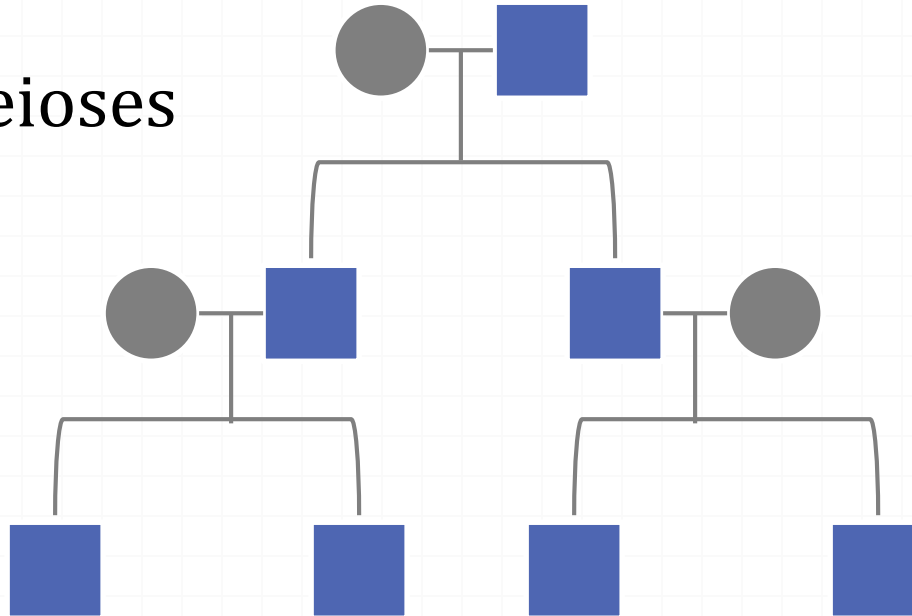
Y-STRs with higher mutation rate may distinguish close male relatives

Two meioses

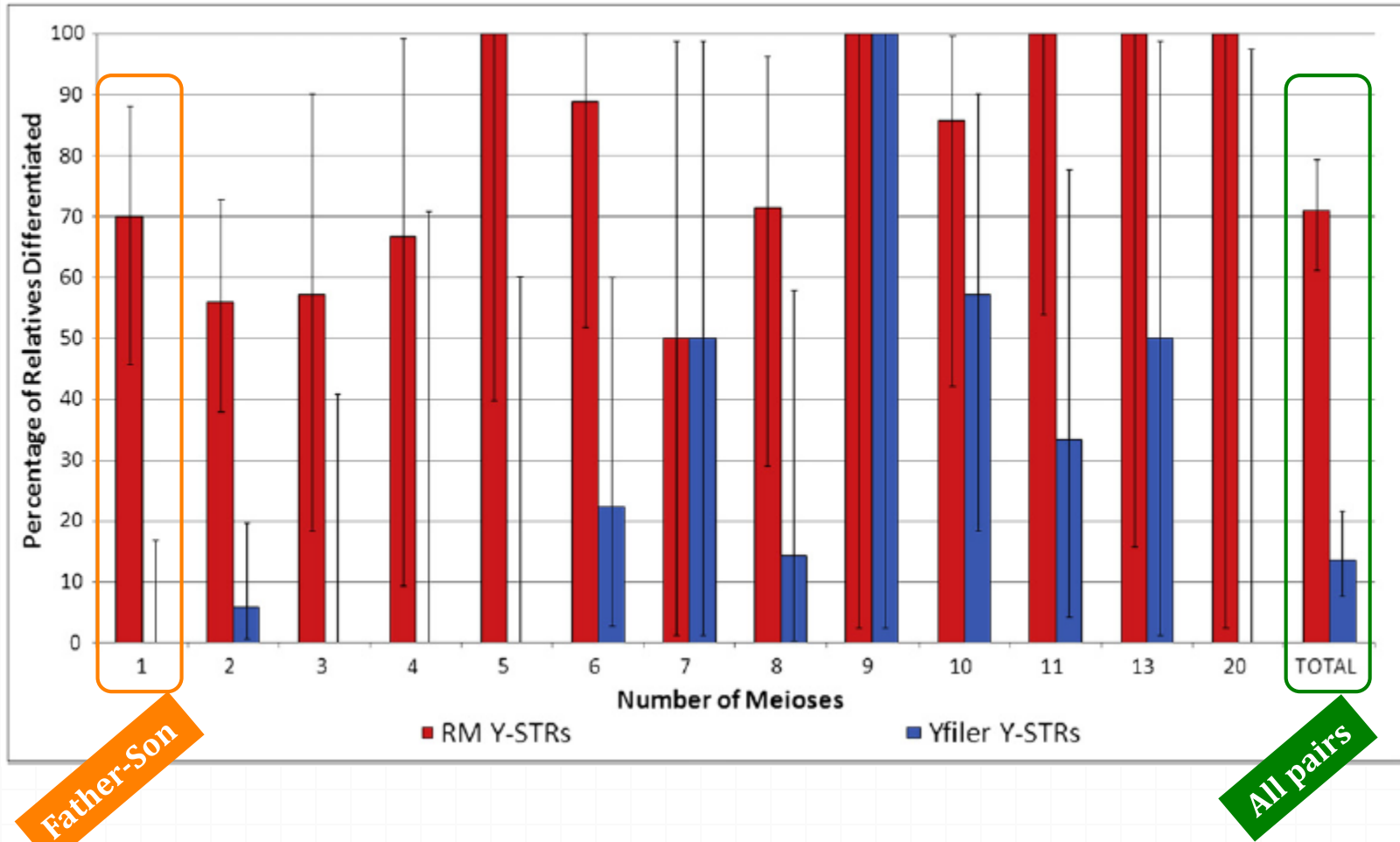


Y-STRs with higher mutation rate may distinguish close male relatives

Three meioses



Y-STRs with higher mutation rate may distinguish close male relatives



Father-Son

All pairs

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.

Performance with Unrelated Males

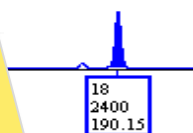
NIST U.S. Population Samples

Rapidly Mutating (RM) Y-STRs

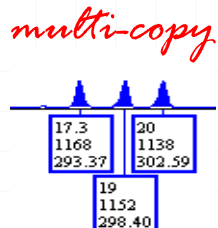
RM Y-STR Study Group
led by Manfred Kayser

NIST supplied data
from:
1,296 U.S. samples
- 634 population
- 331 father/son
pairs

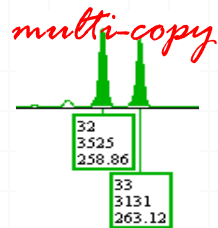
DYS576



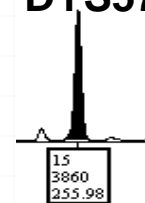
DYF399S1



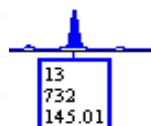
DYF387S1



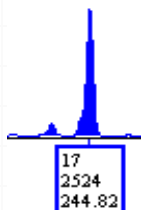
DYS570



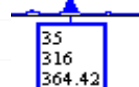
DYS526a



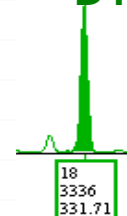
DYS626



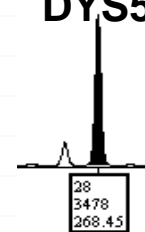
DYS526b



DYS627

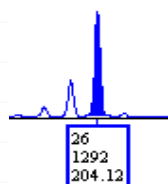


DYS518

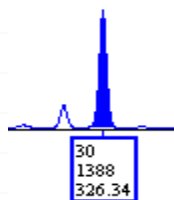


Multiplex 2

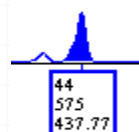
DYS612



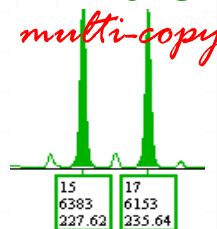
DYS449



DYS547



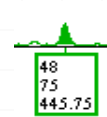
DYF404S1



DYF403S1a

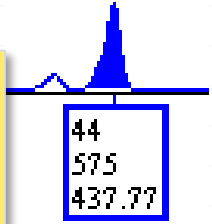
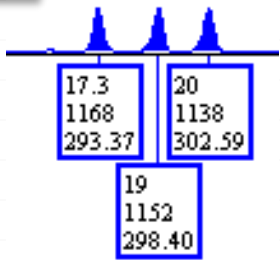


DYF403S1b

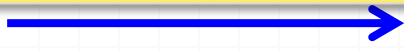


Multiplex 3

Why do these markers mutate “rapidly”?

	Locus	(average mutation rate)	Large number of repeats	
Markers in Yfiler Plus	DYS449	(1.2%)	 <p>DYS449</p>	
	DYS518	(1.8%)		
	DYS547	(2.4%)		
	DYS570	(1.2%)		
	DYS576	(1.4%)		
	DYS612	(1.4%)		
	DYS626	(1.2%)		
	DYS627	(1.2%)		
	DYF387S1	(1.6%)		 <p>DYF387S1</p>
	DYF399S1	(7.7%)		
DYF403S1 a/b	(3.1/1.2%)			
DYF404S1	(1.3%)			
DYS526 a/b	(1.3%)			

Point of Comparison:
DYS458 has highest mutation rate in Yfiler at 0.64%
Yfiler average ~0.2%



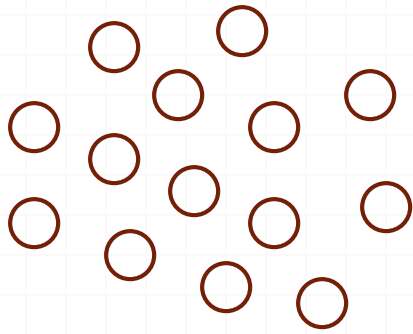
i-copy Markers

RM Y-STRs for close male relatives

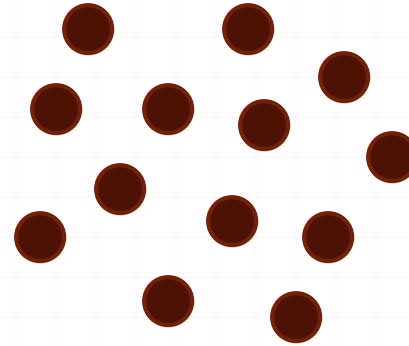
$$GD = \left(1 - \sum_i x_i^2\right)$$

↑
Relative frequency
of each allele

Gene Diversity
measures the
uniqueness of a
marker in a
population



Marker Y



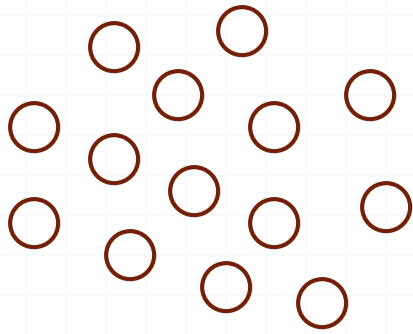
1 type = 100%

$N = 100$

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

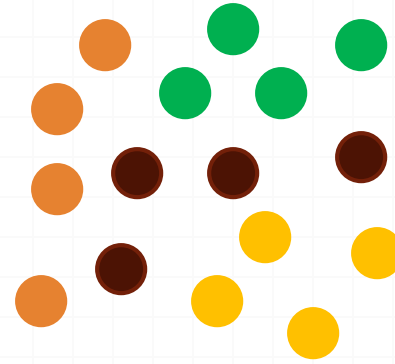
0

$$DC = 1/100 = 0.01$$



$N = 100$

Marker Y
→

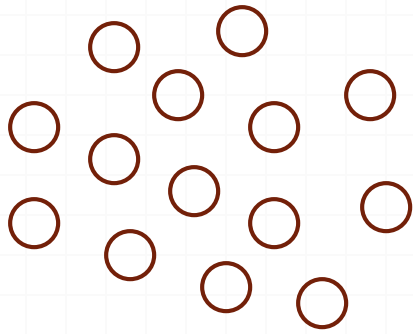


4 types = 25%

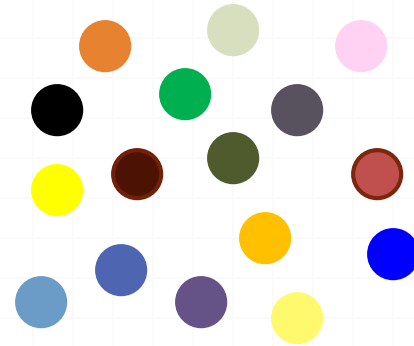
$$GD = (1 - \underbrace{\sum_i x_i^2}_{0.75})$$

0.75

$$DC = 4/100 = 0.04$$



Marker Y



100 types = 1%

$N = 100$

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

0.99

$$DC = 100/100 = 1.0$$

Gene Diversity of the 13 RM Y-STRs

Marker	GD	DC
DYS576	0.766	0.035
DYF399S1	0.993	0.587
DYF387S1	0.870	0.098
DYS570	0.743	0.035
RM-01 (all)	0.9998	0.9764

Marker	GD	DC
DYS526a/b	0.923	0.138
DYS626	0.794	0.043
DYS627	0.848	0.043
DYS518	0.791	0.039
RM-02 (all)	0.9985	0.8661

Point of Comparison:
DYS385a/b
has gene diversity of
0.929

Marker	GD	DC
DYF403S1a/b	0.923	0.791
DYF404S1	0.902	0.110
DYS612	0.832	0.043
DYS449	0.796	0.043
DYS547	0.798	0.039
RM-03 (all)	1.000	0.9984

All 13 RM Y-STRs
resolved 948
males!

Mutation Rate Information

Meioses	Mutations	Group
63	15	AfAm
89	25	Asian
91	11	Caucasian
88	20	Hispanic
331	71	total
		(21.4%)

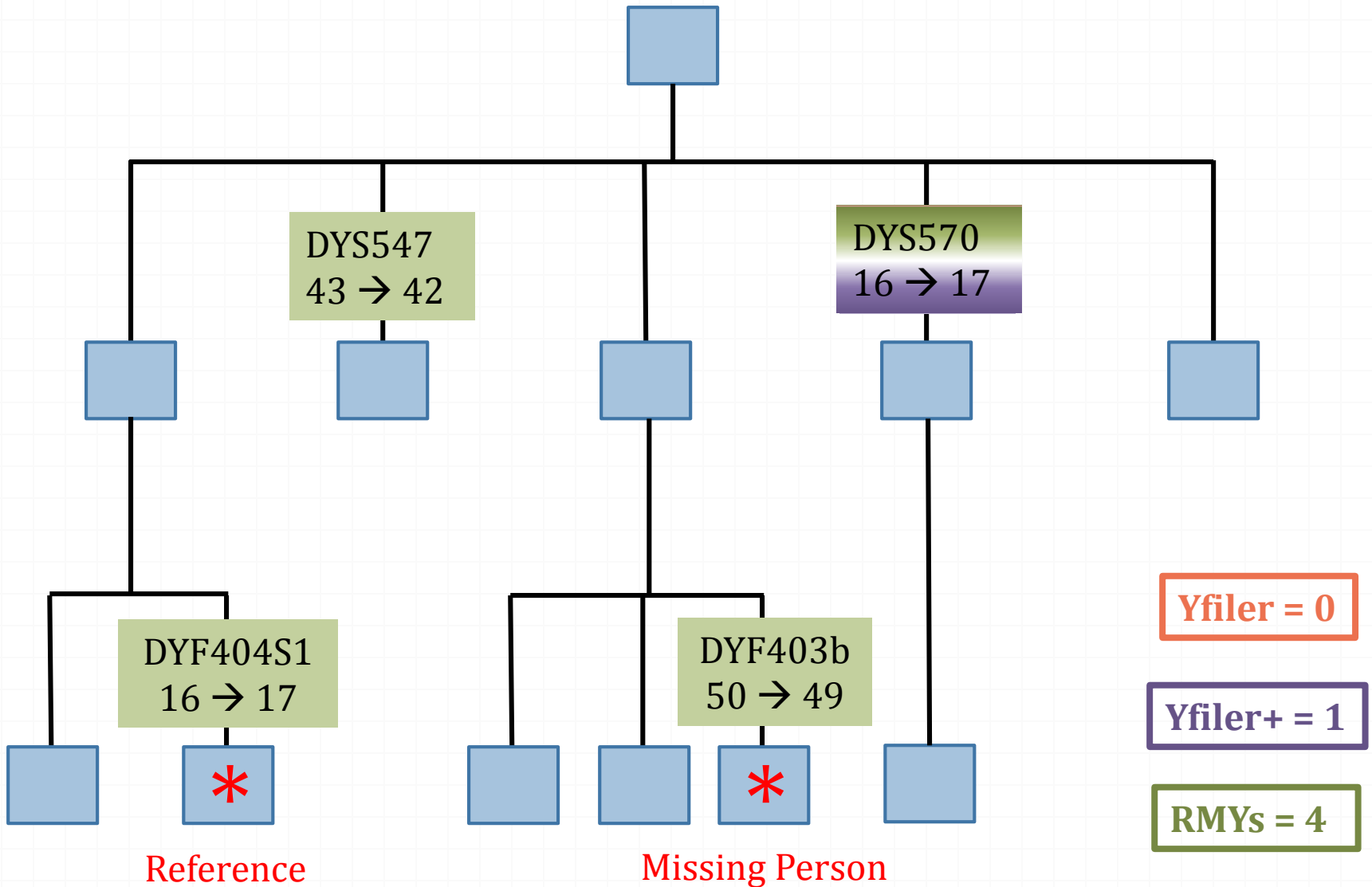
+1 Repeat (Son)	-1 Repeat (Son)	
8	6	AfAm
11	13	Asian
5	6	Caucasian
8	12	Hispanic
+2 Repeat (Son)	-2 Repeat (Son)	
0	1	AfAm
1	0	Asian
0	0	Caucasian
0	0	Hispanic

21.4% of father-sons distinguishable at 13 RM Y-STRs

9.4% father-sons distinguishable at 6 RM Y-STR loci in Yfiler Plus

Marker	# of Mutations
DYF399S1	15
DYF403S1a/b	11
DYS627	7
DYS612	7
DYS518	6
DYS570	5
DYS626	5
DYS547	4
DYS526a/b	3
DYS576	3
DYS449	3
DYF404S1	1
DYF387S1	1

Y-STR mutations in a Paternal Lineage



Interpretational Issues

Move away from “excluding” based on discordant markers

Likelihood Ratio weighs competing propositions

Requires haplotype frequency, # meioses and mutation rate

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.

Summary

27 Yfiler Plus loci show improved haplotype resolution

- 10 new loci alone performing better than Yfiler

Preliminary testing shows Yfiler Plus has improved sensitivity

- Full profiles possible at 62.5 pg input DNA (half reaction)

Rapidly Mutating Y-STRs are highly diverse markers that can discriminate common haplotypes and close relatives

RM Y-STRs may create interpretational issues for paternity/missing persons cases

- LR's can be useful for evaluating these situations

Acknowledgments



**Mike
Coble**



**Peter
Vallone**

2014 “Future trends in
forensic DNA technology”
seminar series



HID
UNIVERSITY
life technologies



ThermoFisher
SCIENTIFIC



life
technologies™

Contact Info:

katherine.gettings@nist.gov

301-975-6401

Jonathan Tabak
Andrea Carbonaro
Ariana Wheaton

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 those of the presenters and do not necessarily represent the official position of the National Institute of Standards and Technology.