

Examination of Rapidly Mutating Y-STR Loci for Increased Resolution of Common Haplotypes Using a Large Multiplex Kit

Michael Coble, PhD; Katherine Butler Gettings, PhD; Becky Hill Steffen, MS; and Peter Vallone, PhD National Institute of Standards and Technology

AAFS Annual Meeting – Criminalistics Session February 21, 2015

NIST and NIJ Disclaimer

Past and Present Funding: Interagency Agreement between the National Institute of Justice and NIST Office of Law Enforcement Standards

Points of view are mine and do not necessarily represent the official position or policies of the US Department of Justice or the National Institute of Standards and Technology.

Certain commercial equipment, instruments, software 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 the National Institute of Standards and Technology nor does it imply that any of the materials, instruments or equipment identified are necessarily the best available for the purpose.

What has happened in the past decade...

- Selection of core Y-STR loci (SWGDAM Jan 2003)
- "Full" Y-chromosome sequence became available in June 2003; over 700 Y-STR loci identified (only ~20 in 2000)
- Commercial Y-STR kits released

 Y_PLEX 6,5,12 (2001-03), PowerPlex Y (9/03), Yfiler (12/04), PPY23 (6/12) Yfiler Plus (since 2014)
- Many population studies performed and online databases generated with thousands of Y-STR haplotypes
- Forensic casework demonstrations showing value of Y-STR testing along with court acceptance
- · Renewed interest in Y-STRs to aid familial searching

	STR Marker Layouts for Y-STR Kits	
2003	I I	12plex (4-dye)
Yfiler 7004	DY\$456 DY\$3891 DY\$390 DY\$3891 DY\$458 DY\$19 DY\$385 a/b DY\$393 DY\$391 DY\$439 DY\$635 Y-GATA-H4 DY\$437 DY\$438 DY\$448	17plex (5-dye)
PowerPlex Y23	DYS576 DYS389I DYS448 DYS389II DYS19 DYS391 DYS481 DYS549 DYS333 DYS438 DYS437 DYS570 DYS655 DYS390 DYS439 DYS392 DYS643 DYS393 DYS458 DYS395 a/b DYS456 Y-GATA-H4	23plex (5-dye)



















Disadvantages of the Y-Chromosome

- Loci are not independent of one another and therefore rare random match probabilities cannot be generated with the product rule; must use haplotypes (combination of alleles observed at all tested loci)
- Not as informative as autosomal STR results

 More like addition (10 + 10 + 10 = 30) than multiplication (10 x 10 x 10 = 1,000)
- Paternal lineages possess the same Y-STR haplotype (barring mutation) and thus fathers, sons, brothers, uncles, and paternal cousins cannot be distinguished from one another

Rapidly Mutating (RM) Y-STRs

Trying to <u>separate</u> close male relatives















Performance with unrelated males

NIST U.S. Population Samples













N - 948 males	PowerPlex Y	Yfiler	PowerPlex Y23	Yfiler Plus*	
# haplotypes	816	930	945	946	Number of unique shared haplotyr
discrimination capacity	0.8608	0.9810	0.9968	0.9979	observed with va combinations of Y
# times haplotype	PPY	Yfiler	PPY23	Yfiler Plus	loci across 948 l
observed	(12 loci)	(17 loci)	(23 loci)	(27 loci)	population samp
1	751	916	942	944	
2	42	11	3	2	
3	12	2			
4	4	1			
5	2				
6	2				
7					
8	1				
9					
10					
11	1				
12					
13					
14					
15					
16					
17					
18					
19					
20	1				



N = 948 males	Yfiler	New Loci*	Yfiler Plus*	
# haplotypes	930	945	946	
discrimination capacity	0.9810	0.9842	0.9979	The new loci alone
# times haplotype	Yfiler	New Loci*	Yfiler Plus	perform slightly better
observed	(17 loci)	(10 loci)	(27 loci)	than Yfiler
1	916	918	944	
2	11	15	2	
3	2			
4	1			
5				
6				
7				
8	•			
9		•	-	
10	•			
11	•			
12		•	•	
13	•	•	-	
14	•	•	-	
15	•	•		
10	•	•	-	
10	•	•	•	
10	•	•	-	
20	•	•	•	
20				



Gene Diversity

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

$$\begin{array}{ll} \mathsf{GD} = & (1 - \sum_i x_i^2) \\ & & \uparrow \\ & &$$













Gene Diversity	Locus	Gene Diversity
0 0 0 2 0 . 0 j	DYF387S1a/b	0.919
of the YFP	DYS385a/b	0.919
	DYS627	0.8584
Markers	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











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^{4,4}, Mark Vermeulen^{4,5}, Hans Knoblauch⁶, Herbert Schuster⁴, Michael Krawczak⁶, Luzt Roewer⁴ Forensic Science International: Genetics 1 (2007) 125–128.

Summary

- Rapidly Mutating Y-STRs are highly diverse markers that can discriminate common haplotypes and close relatives.
- These markers may create interpretational issues for paternity/missing persons cases, but LRs can be useful for evaluating these situations.



