# GeneMarker<sup>®</sup>HID

# Crime Scene Suspect Inclusion/Exclusion using GeneMarker<sup>®</sup>HID with Relationship Testing

Teresa Snyder-Leiby, He Haiguo, Tammy Serensits, Chengsheng Jonathan Liu

### Introduction

The DNA of all humans is more than 99% identical. However, the short tandem repeat (STR) core loci selected for forensic genotyping are areas of variation that can be used to exclude or include an individual in a pool of suspects. The likelihood that unrelated people will share the same STR profile can range from 1 in a billion or more, depending on the number of loci compared between the two samples. Kinship formulas established in the literature calculate the likelihood that the sample tested is a random match to another individual from the same population. Data from large population studies provides allele frequency data used in the likelihood ratio calculations<sup>1</sup>.

GeneMarkerHID calculates these likelihood ratios to provide a random match probability<sup>1,2</sup>. This random match probability is the chance that a randomly selected individual from a population will have an identical STR profile at the DNA markers tested. The save to database function in GeneMarkerHID accepts Identifiler<sup>®</sup> CODIS profiles. It can be modified to accept profiles from other standard kits including, Cofiler<sup>®</sup>, Profiler <sup>®</sup>, ProfilerPlus<sup>®</sup> Minifiler<sup>®</sup>, SGMPlus<sup>®</sup>, PowerPlex (12,16,S5,ES,Y)<sup>®</sup>, Yfiler, providing easy updates of the closed system as additional profiles become available.

### Procedure

- 1. File→Open data→Project→Run→to make allele calls (See Chapter 2, GM HID User Manual) or open previously saved GM HID project
- 2. Select Relationship Testing
- 3. Select Family Group Tool and 'Okay'
- 4. Select the appropriate allele frequency<sup>1</sup>
- 5. Select individual node and choose find family

### Results

**Figure 2:** The database search and calculations are performed with the 'find family command'. Sample PED\_1\_C has the same STR profile as the suspect. The random match probability is 1 in  $5.13 \times 10^{21}$  between sample PAT 1 C and PED 1 C.

Panel:	Identifiler_D	DC		•	i 🖻				
Population:	US Caucasi	an		ĩ					
	US African /	American		1					
	US Caucasi	an			1	1		7	
CSF1P0	US Southea	est Hispanic est Hispanic		9	10	10.3	11		
	0.012300	0.012300	0.012300	0.012300	0.019700	0.253700	0.012300	0.30	
TPOX	<6	6	i 7		9	10	11	12	
	0.012300	0.012300	0.012300	0.544300	0.123200	0.036900	0.253700	0.03	
TH01	<5	5	6	7	8	8.3	9	9.3	
	0.012300	0.012300	0.226600	0.172400	0.125600	125600 0.012300		0.30	
vW/A	11	12	13	14	15	16 17		18	
	0.012800	0.012800	0.012800	0.102000	0.112200	0.201500 0.26280		0.22	
D165539	<8	8	9	10	11	11.3	12	13	
	0.012400	0.019800	0.104000	0.066800	0.272300	0.012400	0.339100	0.16	
D7S820	6	6.3	7	8	8.1	8.1	9	9.1	
	0.012300	0.012300	0.017200	0.162600	0.012300	0.012300	0.147800	0.01	
0136317	1<8	8	81	9	10	11	12	13	
Befresh		ave	Export	1		Ok	1	Cance	-

**Figure 1:** Select the appropriate panel and population to supply allele frequency data for the calculations.

iearch: 44 &	▼ Search:		\$	All Marker	vidual Sample PAT_1_C.fsa PAT_1_F.fsa PAT_1_M.fsa PAT_2_C.fsa	1	Να. 1 2 3 4 5 6 6 7 7	Select Node Deselect Node Select Parents Select Siblina	1
ais	tion Details	2	9 10	Select Family Edit Node					
ed Alleles   Matched Markers   PI/KI	Matched Alleles	X/Y	Name	ID	File Name		11 12	Add Child	
					Same-Person		13	Add Mate	
16/16 5.13E+2	32/32	XY		723	PED_1_C.fsa	1	14	Find Family	
					Enther		15		
16/16 2.11E+0	20/32	XY		724	PED 1 E.fsa	1			
	20102				100_1_1.00	÷			
					Mother				
16/16 1.12E+1	18/32	××		725	PED_1_M.fsa	1			
					E.J.Ch.				
16/16	18/32	×		725	PED_1_M.fsa Full-Sibs	1			

Within seconds the Relationship Testing tools of GeneMarkerHID locate duplicate matches and calculate the likelihood that a random person from the same allele frequency population would have a matching STR profile. In the example above the file PED\_1\_C has the same STR profile as suspect sample PAT\_1\_C. There is a 1 in  $5.13 \times 10^{21}$  chance that a random person in the population has that same profile. Likelihood ratios are also generated from samples with high kinship probability to the profile used in the search. This is beneficial in situations where information on the case suggests that the perpetrator may be related to someone with a prior criminal record.

**Figure 3:** Kinship analysis of these two files indicates that they are 1,950,000 times more likely to be full siblings rather than random individuals from the population.

### Discussion

Time required to process and analyze DNA profiles is one of the greatest limitations to DNA data use for detecting repeat offenders through CODIS and in the exoneration of

Marker	Person A			Perso	Person B							Relationships	Report Content	
	PAT_	TU_CT.Isa	-	PAT.	10_03.158	_	Unrelated (P)	Parent-Child (P)	Parent-Child (LR)	Full-Sibs (P)	Full-Sibs (LR)	Unrelated		C Likelihood Rati
	12	13		12	13		0.03288	0.17740	5.39538	0.34692	10.55110	Parent-Chi	ld	
TPOX	9			9			0.03305	0.18180	5.50055	0.34916	10.56429	🔽 Full-Sibs		C Probability
THOI	5	6		5	6		0.00261	0.06070	23.29151	0.28100	107.82454	☐ Hall-Sibs		
ww.a	17	19		17	18		0.04989	0.13610	2.72777	0.08052	1.61388	Uncle-Nep	hew	<ul> <li>Both</li> </ul>
D16S539	12			10	12		0.04105	0.11000	2.67953	0.06526	1.58976	Cousins 🗌		
D7S820	11	13		9	10		0.10174	0.00000	0.00000	0.02543	0.25000	E Grandpare	nt-Grandchild	
D13S317	12	14		10	12		0.04861	0.05030	1.03477	0.03730	0.76738		Ok	Cancel
D55818	11			11			0.06817	0.26110	3.82995	0.39759	5.83210			
FGA	19	25		19	25		0.01056	0.07640	7.23485	0.29084	27.54167			
D8S1179	12	14		13	14		0.14812	0.22220	1.50015	0.14813	1.00008			
D18551	15			14	17		0.02095	0.00000	0.00000	0.00524	0.25000			
D21511	29	33.2		28	33.2		0.01441	0.21510	14.92537	0.11115	7.71269			
D3S1358	16			16	18		0.03366	0.05480	1.62813	0.03581	1.06407			
				·	Produ	+ Grove	3.11E-20	0.00E+00	0.00E+00	6.06E-14	1.95E+06			

wrongly convicted individuals<sup>4</sup>. Within three 'clicks' from the main analysis window, GeneMarkerHID uses robust, proven statistics to search the database for exact matches and calculate the random match probability. In addition, GeneMarkerHID has excellent application potential for genotyping, missing person identification, detecting lab personnel contamination, detecting cell line contamination, animal breeding programs and natural population kinship analysis.

GeneMarkerHID Relationship Testing has all of the strengths of GeneMarker including; unique pattern recognition and sizing technology providing >99% accuracy, easy linked navigation, management control and tracking, exportable CODIS and LIMS reports, bulk printing capabilities, instrument compatibility with ABI, MegaBACE and Beckman-Coulter and compatibility with STR kits (including: Cofiler<sup>®</sup>, Profiler<sup>®</sup>, Identifiler<sup>®</sup>, Minifiler<sup>®</sup>, SGMPlus<sup>®</sup>, PowerPlex<sup>®</sup>)

## Acknowledgements

We would like to thank Tom Reid (DNA Diagnostic Center, Oxford OH) for use of data and collaboration; and Arthur Eisenberg, John Planz and Rhonda Roby (UNT-HSC, Fort Worth, TX) for collaborating in the development of this relationship testing tool.

Registered trademarks are the property of the respective owners.

#### References

- 1. Budowle, B., B Shea, S Niezgoda, R. Charkaraborty. CODIS STR loci data from 41 sample populations Journal of Forensic Sciences 2001. 46(3)453-489.
- 2. Li, CC and L Sachs. The derivation of joint distribution and correlation between relatives by the use of stochastic matrices. 1954. Biometrics 10:347-360
- 3. Eisenberg, AJ and JV Planz. 2007. Applied Forensic Statistics, 18th International Symposium on Human Identification
- 4. President's DNA Initiative, http://www.dna.gov