GeneMarker®HID
by SoftGenetics
July 2008

Crime Scene Suspect Inclusion/Exclusion using GeneMarker®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[1].

GeneMarkerHID calculates these likelihood ratios to provide a random match probability[1,2]. 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® CODIS profiles. It can be modified to accept profiles from other standard kits including, Cofiler®, Profiler ®, ProfilerPlus® Minifiler®, SGMPlus®, PowerPlex (12,16,S5,ES,Y)®, 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[1]
5. Select individual node and choose find family

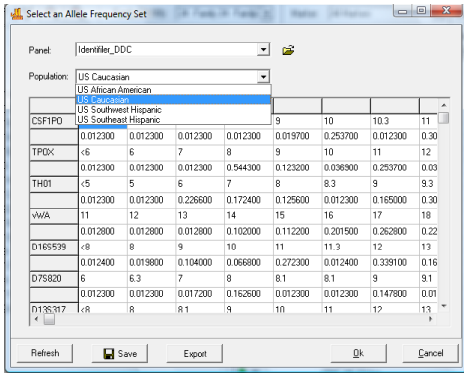
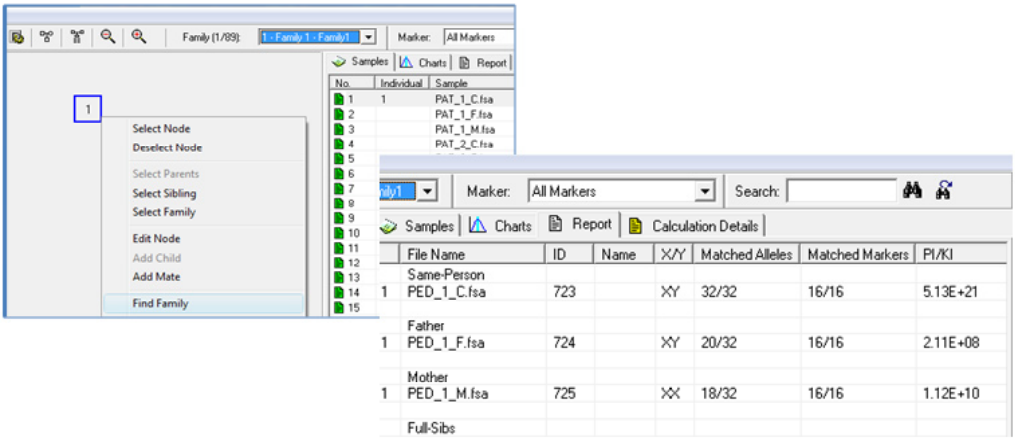


Figure 1: Select the appropriate panel and population to supply allele frequency data for the calculations.

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 x10^{21} between sample PAT_1_C and PED_1_C .



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 $5.13x10^{21}$ 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.

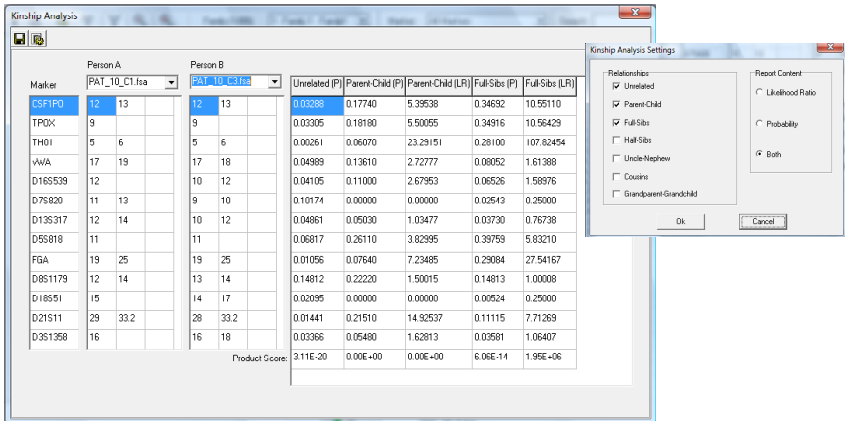


Figure 3: Kinship analysis of these two files indicates that they are 1,950,000 times more likely to be full siblings rather than random individuals from the population.

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 wrongly convicted individuals[4]. 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®, Profiler®, Identifiler®, Minifiler®, SGMPlus®, PowerPlex®)

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.

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

SoftGenetics LLC 200 Innovation Blvd. Suite 235 State College, PA 16803 USA
Phone: 814/237/9340 Fax 814/237/9343
www.softgenetics.com email: info@softgenetics.com