

Missing Persons Identification using GeneMarker[®]HID with Relationship Testing

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

Introduction

Missing person identification may be necessary in many situations: mass natural disasters (earthquakes, tsunamis), human attacks (Sept 11 World Trade Center), war, or in the cases of thousands of missing children and adults reported every year. Identification of remains from physical characteristics (skeletal features, dental comparisons, fingerprints, scars/tattoos, personal effects) is often impossible due to degradation or damage to the body. Short Tandem Repeats (STR) analysis has the ability to provide complete individual profiles, even when the DNA samples have degraded from time or exposure to the elements. STRs are variable regions in genomic DNA which are amplified with specific primers by Polymerase Chain Reaction (PCR). 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. Related individuals have more shared loci than unrelated people. Kinship formulas have been established in the literature to calculate the relatedness between individuals based on shared loci.

GeneMarkerHID includes tools that allow accurate and rapid comparison of an STR profile from an unidentified sample to an STR data base. The rigorous statistical analysis to determine levels of kinship follows the methods of Brenner¹ and uses stochastic matrices of Li and Sachs³. Features of the program include: user friendly linked navigation, management control and tracking, individual peaks are quality flagged including failure rationale, supervisor review module, exportable CODIS and LIMS reports, and bulk printing capabilities.

Procedure

1. Open data file or previously saved project
2. Run Wizard to call alleles
3. Select Relationship Testing
4. Select Family Group Tool
5. Select the appropriate allele frequency²
6. Select individual node and choose find family to display the nearest relatives from the data base

Results

Figure 1: The kinship calculations are performed when the “find family command” is selected for a sample. These values are presented in a report format, listing the samples with the highest kinship probability.

The report provides the identifiers of samples with the highest kinship scores to the unidentified sample, number of matched alleles and the probability of a specific type of relationship versus unrelated (Fig. 1). Figure 2 displays the kinship analysis for comparison of two samples of interest. Kinship analysis displays the kinship equation result for each locus and the product of all loci for a given relationship. The final row in the analysis is the probability that the samples share a given relationship versus unrelated.

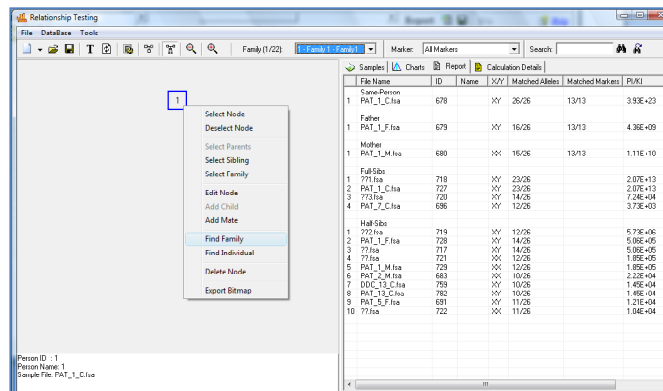


Figure 1

Figure 2: Kinship analysis comparing the unidentified sample with a sample that was identified as having a high probability of parent-child relationship with the find family tool.

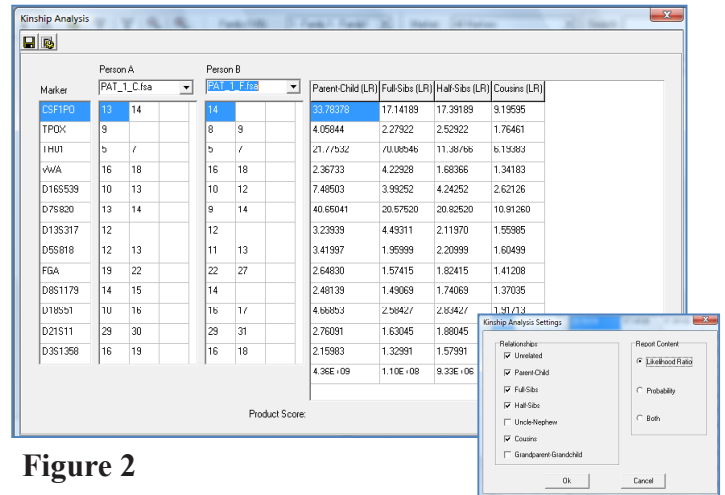


Figure 2

Discussion

The GeneMarkerHID with Relationship Testing uses rigorous statistical methods to determine kinship between an unidentified sample and members within a database. STR profiles from unidentified human remains can be compared to the database to locate exact matches or samples with high kinship scores. Text files can be used to update the DataBase and Allele Frequency functions. The Relationship Testing automatically generates reports that include: likelihood ratios of nearest relatives in the database, identifiers, number of alleles matched and a list of exact copies located when comparing the unidentified sample to the database.

GeneMarkerHID with 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®)

In addition, GeneMarkerHID with Relationship Testing has excellent application potential for genotyping, quality control to detect cell line contamination, animal breeding programs and natural population kinship analysis. Paternity issues in animal breeding programs are common with livestock and in aquaculture. Aquaculture also would have the ability to detect escaped domestic individuals in wild populations, as is a concern with salmon. Kinship analyses are strong tools, either separately or in complementation with observation/behavioral data, to track emigration/immigration and inbreeding. The instant listing of duplicate copies of a genotype in the report provides excellent verification of the actual number of animals in a natural population.

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. Brenner, C. Symbolic kinship program, *Genetics* 145(1997):535-542
2. Budowle, B., B Shea, S Niezgod, R. Charkarabarty. CODIS STR loci data from 41 sample populations *Journal of Forensic Sciences* 2001. 46(3)453-489.
3. 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
4. Eisenberg, AJ and JV Planz. 2007. *Applied Forensic Statistics*, 18th International Symposium on Human Identification