

# GeneMarker® HID

by SoftGenetics

September 2009

## Analysis of DNA Mixtures in GeneMarker® HID Software: with or without single source reference samples

Haiguo He, Teresa Snyder-Leiby, Rensheng Qi and Jonathan C.S. Liu

### Introduction

Short Tandem Repeat (STR) genotyping is instrumental in data analysis for a wide range of human identification applications. Very often samples from mass disasters and crime scene investigations are mixtures rather than single source; requiring additional analysis after making allele calls.

#### Clayton *et al.* proposed five steps for mixture analysis:

1. Identify the presence of a mixture
2. Identify the number of contributors to a mixture
3. Determine the approximate 'ratio' of the components in the mixture
4. Determine the possible pair-wise combinations for the components of the mixture
5. Compare the resultant profiles for the possible components of the mixture with those from reference samples

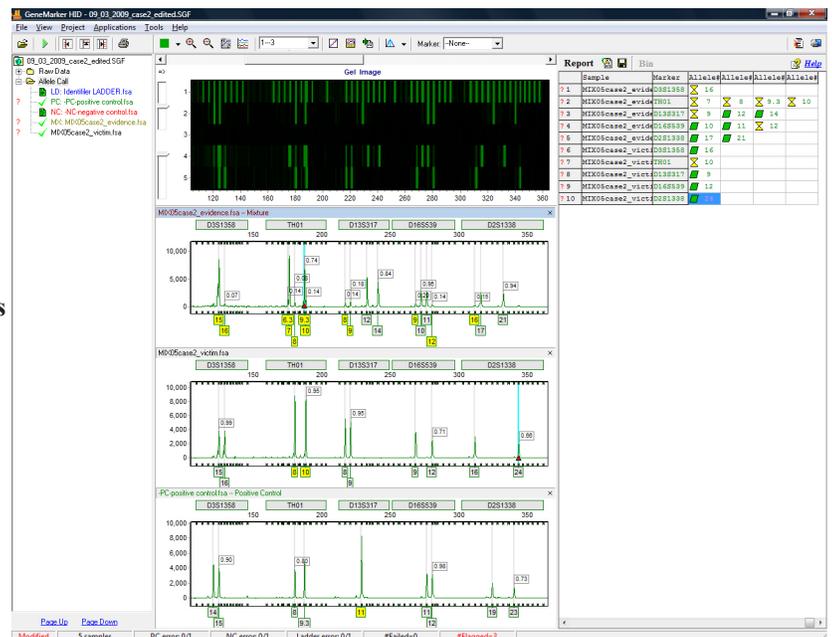
Much of this process is repetitive (for a two-person mixtures there are 50 possible allele combinations for 2, 3, and 4 alleles). Analysis rules may be applied to the genotyping results to automate the repetitive portions of the analysis and provide the analyst with a streamlined workflow for mixture interpretation.

#### GeneMarker HID's Mixture Analysis Features:

1. Identifies the potential mixture samples and any single source contributor samples in a file name tree
2. Considers all possible allele combinations
3. Calculates the mixture ratio, residual, heterozygous imbalance for each genotype combination
4. Calculates the likelihood ratio for single source samples that are potential contributors to the mixture
5. Submit reference files of contributors to the Relationship Testing Database and searches for matches
6. In the absence of a suspect reference sample, submit single source suspect profile based on deconvolution of mixture sample
7. Easily select analysis parameters; including options shown in figure 2

All profiles can be used to screen the database for exact matches and potential close relatives or exported and saved as an excel or .txt file.

**Figure 1:** Main analysis screen identifies potential mixtures and labels peak ratios.



**Figure 1**

GeneMarker HID provides consistent and accurate genotype determination and an embedded mixture analysis application which assists analysts by automating repetitive calculations. The mixture analysis application was developed using recommendations of the DNA commission of the International Society of Forensic Genetics (Gill *et al.*, 2006) and methods of Clayton *et al.*, 1998 and Gill *et al.*, 1998.

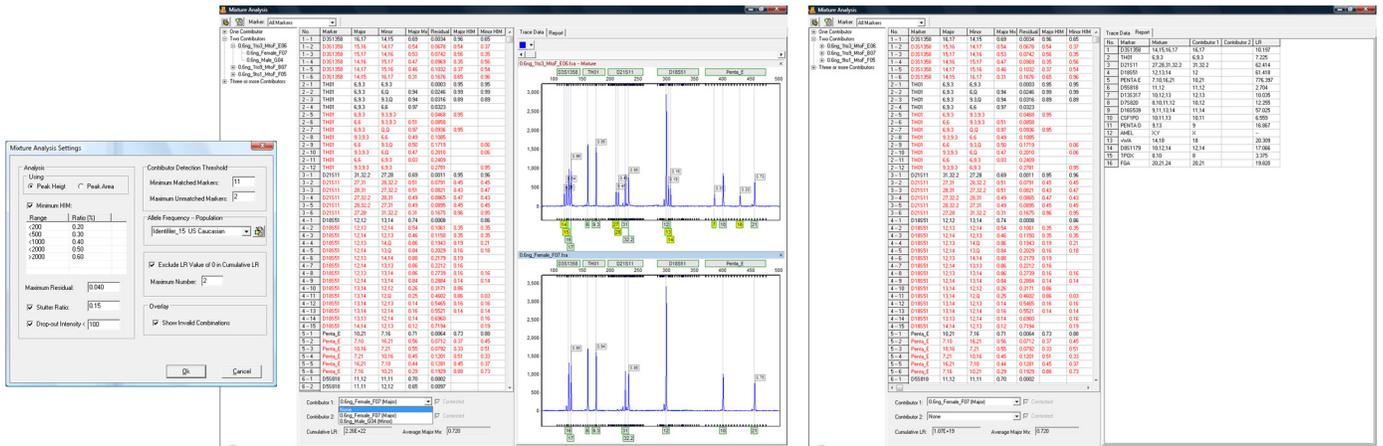
GeneMarker HID identifies the presence of potential mixture samples using two parameters in the Run Wizard: 1) number of alleles / marker and 2) peak area or height ratios. GeneMarker HID makes allele calls, calculates peak ratios and color codes potential mixture samples in the main analysis screen.

# Procedure

1. Select File → Open data or Open project → Add → Okay
2. Launch Run Wizard from pop-up menu or from Project drop-down menu
3. Select Appropriate Template and Run Parameters
4. Review allele calls and summary information (figure 1)
5. Applications → Mixture Analysis
6. Launch Mixture Analysis Settings 
7. Review results, select reference sample (figure 2)
8. Export mixture analysis table and likelihood ratio results or
9. Search database for matches

**Figure 2:** Analysis Parameters are easily set in the dialog box, major and minor contributor files are identified, allele combinations that conform to the analysis parameters are in black font, trace comparison and LR results are shown in same analysis screen

Figure 2



# Results

When reference samples are available the likelihood ratio results can be displayed in an exportable table (figure 2). The mixture analysis uses mixture ratio and residual scores to rank the possible allele combinations for major and minor contributors. In some cases only the mixture sample and the victim reference sample are available. DNA isolation from rape cases often have successful male/female cell separation that provides a mixture where the perpetrator is the major contributor and the victim is the minor contributor. Results shown in figure 3 allow the analyst to identify a single profile that corresponds to the perpetrator. This profile can be used to search the database in GeneMarker for an exact match or match to a potential relative; or saved and exported for use with external database searches.

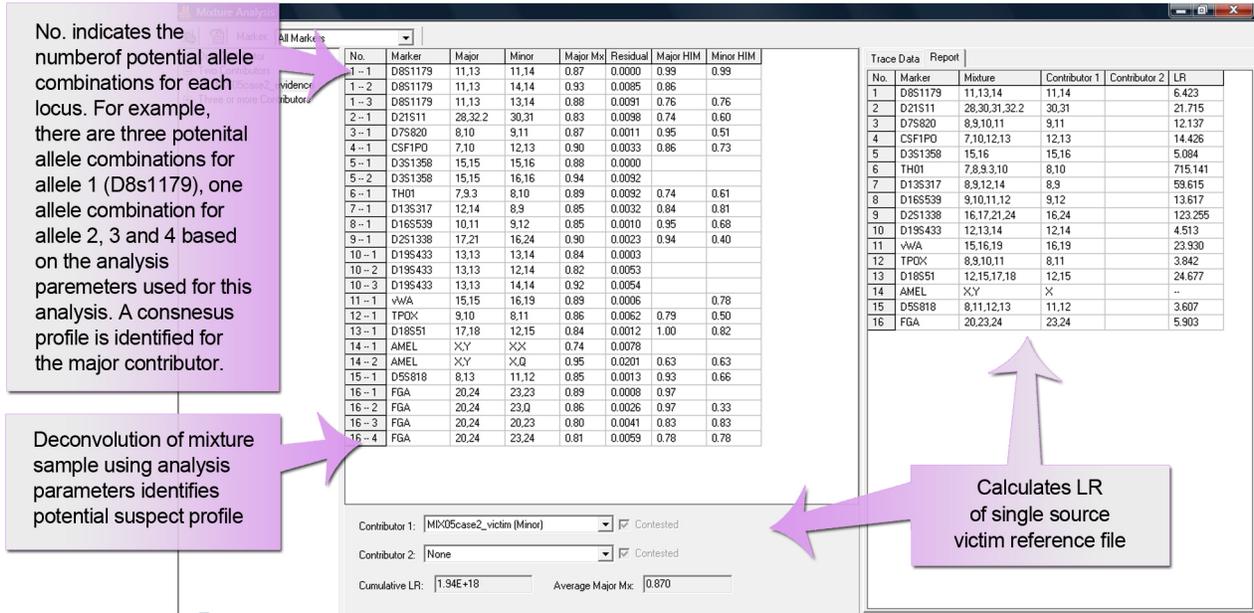


Figure 3

**Figure 3:** A single profile for the major contributor is identified even in the absence of a major contributor reference file.

Residual score is calculated

$$\sum (p_a \text{ observed} - p_a \text{ expected})^2$$

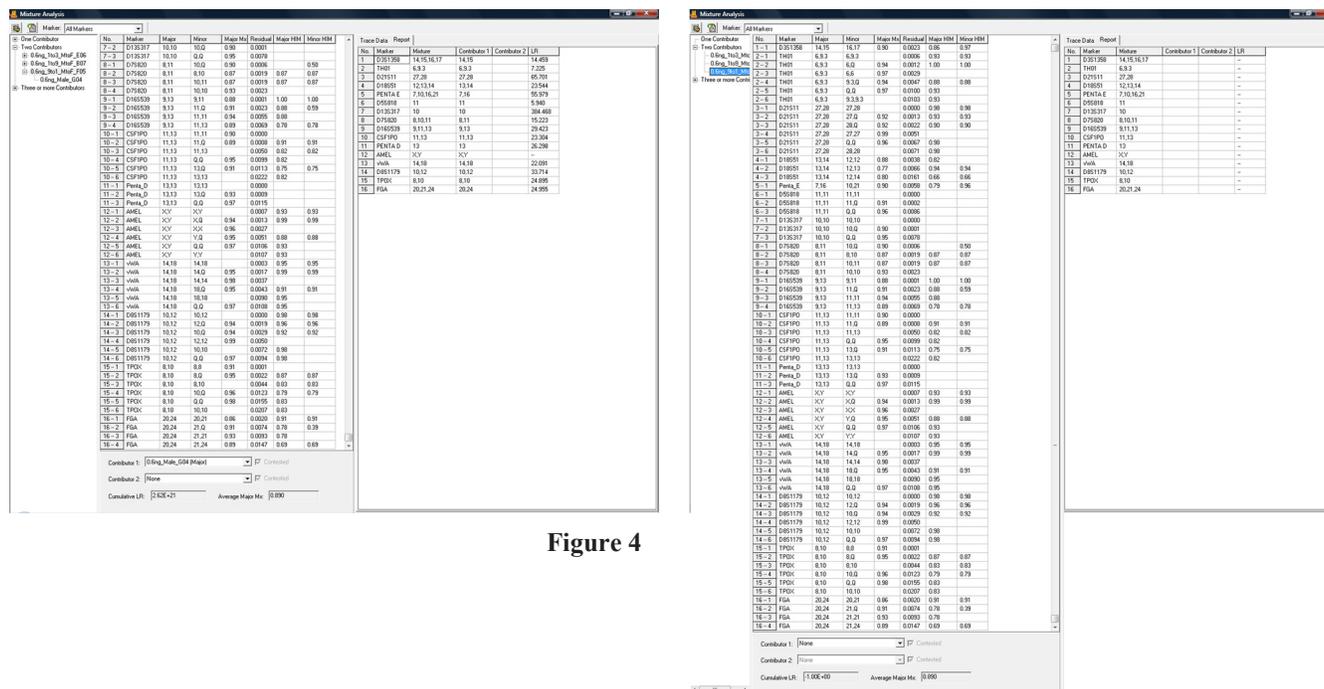


Figure 4

Figure 4: Potential perpetrator profile is identified when a reference sample is available (left) and in the absence of a reference sample

## Discussion

The mixture analysis application in GeneMarker HID is a powerful assistant to analysts in the essential task of mixture interpretation encountered in crime scene and missing persons cases. The program uses standard rules for data analysis and automates repetitive calculations with linked navigation and a user friendly interface. In addition to the five standard steps for mixture interpretation, GeneMarker HID mixture analysis also provides a searchable genotype for a potential contributor in the absence of a reference sample.

GeneMarker HID with Mixture Analysis has the same strengths of GM HID 1.76, including: expert system capabilities, unique pattern recognition and sizing technology providing >99% accuracy, easy linked navigation, management control and tracking provides an audit trail of analyst decisions, 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

Data was generously supplied by Bob McLaren and Doug Storts, Promega Corporation, Madison, WI and John Butler and Margaret Kline through the NIST Interlaboratory Mixture Interpretation Study <http://www.cstl.nist.gov/biotech/strbase/interlab/MIX05.htm>.

## References

1. T.M. Clayton, J.P. Whitaker, R. Sparkes, P. Gill. 1998. Analysis and interpretation of mixed forensic stains using DNA STR profiling. *Forensic Science International* 91:55-70
2. P. Gill, R. Sparkes, R. Pinchin, T. Clayton, J Whitaker, J. Buckleton. 1998. Interpreting simple STR mixtures using allele peak areas. *Forensic Science International* 91:41-53.
3. John Butler. 2008 Strategies for Mixture Deconvolution. [http://www.cstl.nist.gov/biotech/strbase/training/AAFS2008\\_MixtureWorkshop.htm](http://www.cstl.nist.gov/biotech/strbase/training/AAFS2008_MixtureWorkshop.htm)
4. Gill, Brenner, Buckleton, Carracedo, Krawczak, Mayr, Morling, Prinz, Schneider, Weir 2006 DNA commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures
5. P. Gill, B. Sparkes, J.S. Buckleton. 1998. Interpretation of simple mixtures when artefacts such as stutters are present. *Forensic Science International* 213-224

Trademarks are property of their respective owners.