SOFTWARE APPLICATION ENHANCEMENTS IMPROVE EFFICIENCY OF STR ANALYSIS DURING VALIDATION AND CASEWORK

Sarah Copeland, Teresa Snyder-Leiby, Edward Bouton, Jacie Wu

State College, PA
Booth No. 305

Abstract

Improvements are being made in STR human identification chemistries and instrumentation to increase the information gathered from evidence. To minimize analyst intervention when processing this additional information, tools are being developed for forensic STR analysis software.

STR kits have increased sensitivity, enabling detection and analysis of minor contributors in forensic mixtures. Allele specific stutter filters are preferred for complex repeats in loci spanning a large size range or effected by longest uninterrupted sequence (LUS). Allele specific, additive stutter filters in GeneMarker®HID software have additional N-2 repeat stutter position filters (in addition to the previous N-1, N+1, N-0.5, and N+0.5 stutter positions).

Probabilistic Mixture analysis often requires calibration data that includes edited single source dilutions with and without the filtered stutter peaks to give the software context for analyzing casework mixture samples. Previously, the calibration data was edited to remove artifacts and stutter before exporting results. Then, that same calibration data was analyzed without stutter filters (artifacts were still removed) to retain the stutter peaks in the exported results. A new report table toggle option in GeneMarker®HID software that enables the export of both tables from the reviewed/edited calibration data results greatly reduces the time and potential for inconsistencies.

Amplification inhibitors and DNA degradation can complicate forensic STR analysis by producing incomplete genotypes from casework samples. PowerPlex®35GY, Globalfiler™IQC, Investigator 26plex, and Investigator 24plex are examples of PCR kits that use internal quality sensor peaks (also referred to as internal control probes or quality sensor probes). The two internal quality sensor markers at low and high molecular weights saves time and improves workflow by assisting the analyst to discern between failed PCR amplification, absence of DNA, inhibited DNA, and degraded DNA. An embedded module in GeneMarker®HID software provides further efficiency by providing parameters to automatically flag samples that break any of the rules for small and large size internal quality sensor peaks. This automation provides additional time savings and avoids potential human bias when evaluating samples that have unacceptable or decreasing peak heights.

Validation of these new chemistries, instruments, and SOPs is an extremely time-consuming process. Many labs do not have a dedicated validation team, necessitating fitting validation work in with casework. GeneMarker®HID has new tools in the validation assistance module, which now include Limit of Detection (LOD) and calculation of Stutter Ratios to streamline SOP development.

This poster presents examples of each of the above tools using 6 and 8 dye .fsa files analyzed with GeneMarker®HID v. 3.3.0 software.

Conclusion

Linked applications speed up analysis time and remove opportunities for possible variation between analysts by applying consistent, reproducible rules.

The stutter filter and report toggle make genotyping more accurate and also provide information for probabilistic genotyping (PG) software (MaSTR, STRmix, EuroForMix) mixture analysis models.

Quality sensor rubrics from kit manufacturers were previously performed manually. The Quality Sensor Flagging tool automates this process so analysts can quickly identify PCR inhibition, amplification failure, degradation, or insufficient template.

Validation tools automatically apply consistent, reproducible rules to increase efficiency and decrease potential variation in the validation process.

Acknowledgments

SoftGenetics would like to thank Promega and Qiagen for providing data and valuable conversations during development.

References

- Adamowicz, M.S.; Rambo, T.N.; Clarke, J.L. *Internal Validation of MaSTR™ Probabilistic Genotyping Software for the Interpretation of 2-5 Person Mixed DNA Profiles.* Genes 2022, 13, 1429 https://doi.org/10.3390/genes13081429.
- Investigator® Quality Sensor Technical Information, www.Qiagen.com.
- John E. Schienman, et al. *The Evaluation and Implementation of GeneMarker®HID (Softgenetics), PowerPlex® Fusion 6C (Promega), and STRmix™ (ESR) for Forensic Casework*. ISHI 2016 Convention, Minneapolis, MN, United States.
- OSAC 2021-S-0003, *Standards for Setting Analytical and Stochastic Thresholds for Application to Forensic DNA Casework Using Electrophoresis Platforms*, Organization of Scientific Area Committees for Forensic Science, 2021-S-0003

Methods and Results

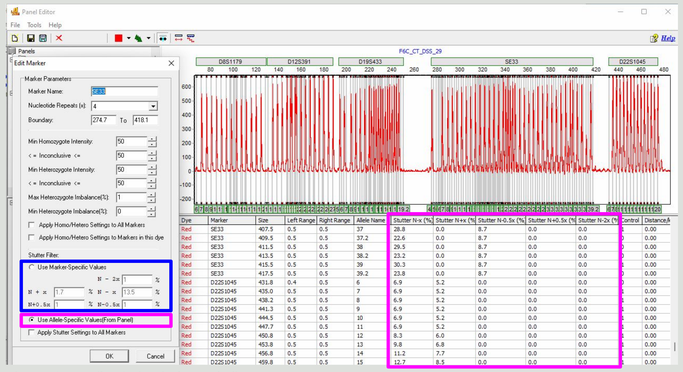


Figure 1: Stutter Filter Options in the Panel Editor

- The Panel Editor has options for setting marker-specific (blue box) or allele-specific (magenta boxes) stutter filters.
- Using the lab's SOP or published guidelines, enter maximum stutter percentages for alleles or stretches of alleles at all positions.
- GeneMarkerHID automatically filters out peaks at these positions if their height ratio is below these settings.

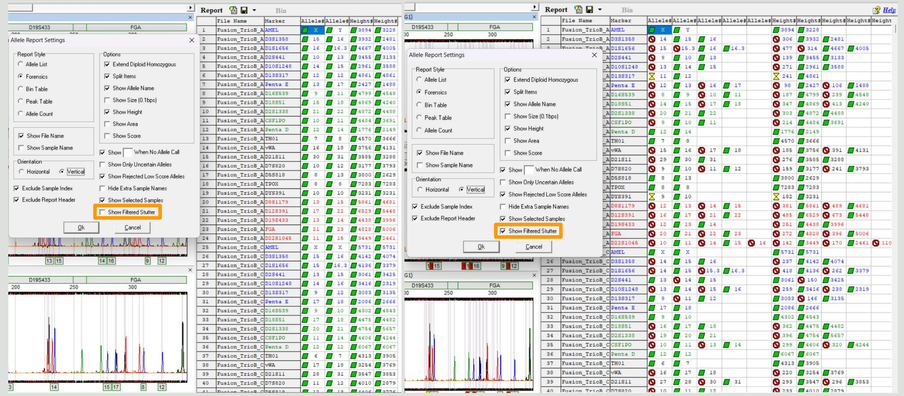


Figure 2: Report Toggle for Stutter

- Users need to generate reports for probabilistic genotyping (PG) software containing the edited samples' profiles with and without filtered stutter peaks.
- Instead of editing the artifacts twice, this setting allows the user to edit the data once and then select this option (orange boxes) to generate the second report.

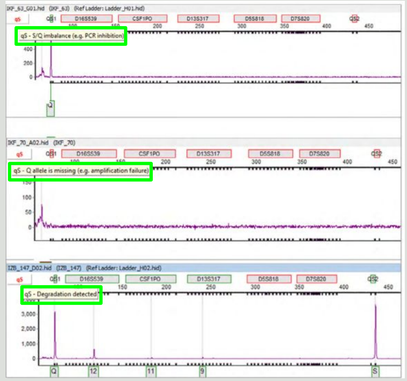


Figure 3: Quality Sensor Flagging

- Certain PCR kits include quality sensor peaks that help the analyst determine if there is degradation, inhibition, or amplification failure.
- GeneMarkerHID automatically analyzes the two qS peaks in each sample and provides a pop-up message with the findings (green boxes).
- In addition to the pop-up message, qS designation is also displayed in the file name tree and the Allele Report table.

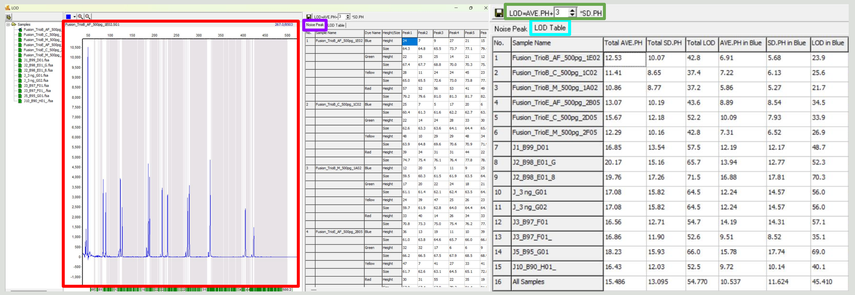
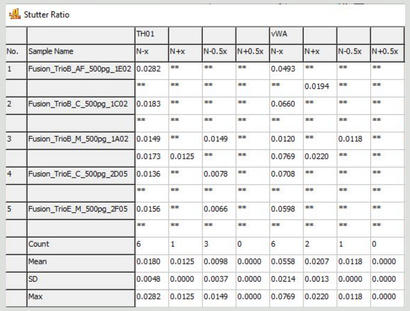


Figure 4: Limit of Detection Validation Tool

- Determining the limit of detection (LOD) or analytical threshold (AT) for a lab's data is an important but very time-consuming step in the validation process
- Manually calculating LOD generally involves using genotyping software to analyze validation data so that every single peak is called. Then, it is necessary to manually delete the calls for true peaks, stutter, and artifacts in order to export the peak heights of all of the noise peaks. The average of the noise peak heights is determined for each dye and then a certain number of standard deviations is added to determine the analytical threshold. This process was performed for five samples and took 38 minutes to complete.
- The Limit of Detection tool in GeneMarkerHID automates this process so that it can be performed in less than five minutes for any number of samples. A) Electropherogram for visual confirmation of deleted alleles, stutter, and other artifacts (red box). B) Noise table displaying peak heights used in the LOD determination (Noise Peak tab in purple box). C) LOD Table (aqua box) showing results per sample, per dye, including average peak heights, standard deviation, and LOD per number of standard deviations specified at the top (green box).



Stutter Ratio									
		TH01				vWA			
No.	Sample Name	N-x	N+x	N-0.5x	N+0.5x	N-x	N+x	N-0.5x	N+0.5x
1	Fusion_TrioB_AF_500pg_1E02	0.0282	**	**	**	0.0493	**	**	**
		**	**	**	**	**	0.0194	**	**
2	Fusion_TrioB_C_500pg_1C02	0.0183	**	**	**	0.0660	**	**	**
		**	**	**	**	**	**	**	**
3	Fusion_TrioB_M_500pg_1A02	0.0149	**	0.0149	**	0.0120	**	0.0118	**
		0.0173	0.0125	**	**	0.0769	0.0220	**	**
4	Fusion_TrioE_C_500pg_2D05	0.0136	**	0.0078	**	0.0708	**	**	**
		**	**	**	**	**	**	**	**
5	Fusion_TrioE_M_500pg_2F05	0.0156	**	0.0066	**	0.0598	**	**	**
		**	**	**	**	**	**	**	**
	Count	6	1	3	0	6	2	1	0
	Mean	0.0180	0.0125	0.0098	0.0000	0.0558	0.0207	0.0118	0.0000
	SD	0.0048	0.0000	0.0037	0.0000	0.0214	0.0013	0.0000	0.0000
	Max	0.0282	0.0125	0.0149	0.0000	0.0769	0.0220	0.0118	0.0000

Figure 5: Stutter Ratio Validation Tool

- Labs often want to estimate the stutter expected when following their SOP in order to optimize their stutter filter settings, which can be a time-consuming process.
- It took 20 minutes to analyze these five samples without any stutter filters, then manually remove artifacts and shared stutter peaks in preparation for calculating average stutter in a spreadsheet. Generally, labs have many more samples than this in their stutter validation study so the time required can be significant.
- The Stutter Ratio Tool in GenemarkerHID can provide a chart like the one above in minutes for any number of samples.