

Automated Contamination Check Tool in GeneMarker®HID Software

September 2015

Adam Dillman, Xianchen Zhang, Teresa Snyder-Leiby, Jonathan C.S. Liu

Introduction

Improvements to the sensitivity of the polymerase chain reaction (PCR) have necessitated increased diligence in detecting possible contaminants. Contamination may jeopardize data analysis and can originate from many common sources, particularly from the skin or hair of a staff member.

GeneMarkerHID software solves this problem by enabling the user to create a database of staff profiles as an elimination database. This database can then be compared against profiles in the current project to search for any contamination. Profiles of the current project may also be compared to one another to detect possible profile-to-profile contamination.

Procedure

1. Load raw data and process it using a preloaded template.
2. Review Size and Allele Calls.
3. Navigate to the Contamination Check tool (Project → Contamination Check Tool)
4. Select analysis parameters and comparison database (See figure 1).
5. Click the refresh icon.
6. Review results, save and export comparison table.

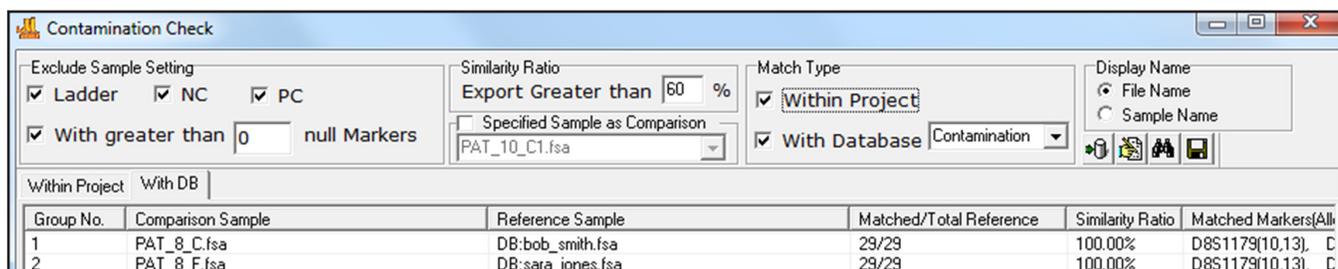
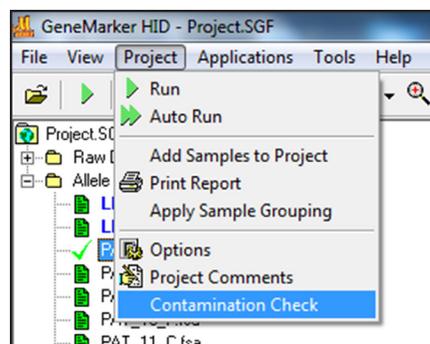


Figure 1: The contamination check tool provides several options for customization. The user can choose to exclude different sample types, including negative controls, pos. controls, ladders, and samples with a user-specified number of null alleles. The user may also choose to perform the comparison between all samples, or a specific sample. Finally, the user can choose to compare profiles to profiles within the project, to samples within a database, or both.

The user may construct the lab's contamination database using the "Submit to Database" option. This is a convenient location to save staff and technician profiles, or profiles for anyone who may otherwise be in contact with the data. Simply choose which files to submit, and then select "Submit". Samples can later be edited using the database editor (See figure 2).

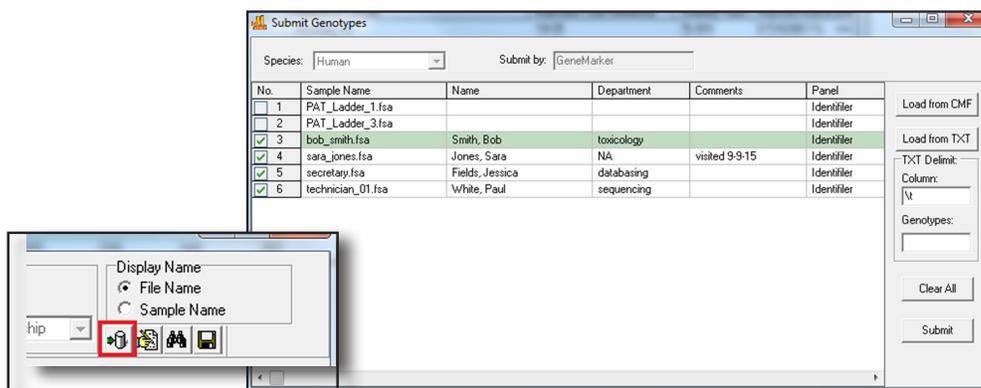


Figure 2: Profiles can be submitted to a Contamination Database using the “Submit to Database” icon highlighted in red. Select which genotypes to add to the database and then click Submit. Profiles in future projects can be checked against the contamination database profiles to determine the likelihood of contamination.

Results

After applying your settings, click the refresh icon (binocular icon in figures 1 and 2) to begin the comparison. Results are shown in the summary table, and can be sorted by clicking the column headers. Information displayed for each sample includes the Sample name, the name of the sample against which it was compared (the reference sample), the number of matching alleles compared to the number of reference alleles, and the similarity ratio, determined by the equation below. See figure 3.

$$\text{Similarity Ratio} = \frac{\text{Number of alleles shared by Comparison and Reference Samples}}{\text{Total number of alleles in Reference Sample}} \times 100\%$$

Group No.	Comparison Sample	Reference Sample	Matched/Total Reference	Similarity Ratio	Matched Markers
1	PAT_8_C.fsa	DB:bob_smith.fsa	29/29	100.00%	D8S1179(10,13), C
2	PAT_8_F.fsa	DB:sara_jones.fsa	29/29	100.00%	D8S1179(10,13), C
3	PAT_8_M.fsa	DB:secretary.fsa	25/25	100.00%	D8S1179(13), D21
4	PAT_9_C.fsa	DB:technician_01.fsa	30/30	100.00%	D8S1179(14,15), C
5	PAT_8_C.fsa	DB:secretary.fsa	19/25	76.00%	D8S1179(13), D21
6	PAT_9_F.fsa	DB:technician_01.fsa	20/30	66.67%	D8S1179(15), D21
7	PAT_8_M.fsa	DB:bob_smith.fsa	19/29	65.52%	D8S1179(13), D21
8	PAT_9_M.fsa	DB:technician_01.fsa	19/30	63.33%	D8S1179(14), D21
9	PAT_8_C.fsa	DB:sara_jones.fsa	18/29	62.07%	D8S1179(10,13), C
10	PAT_8_F.fsa	DB:bob_smith.fsa	18/29	62.07%	D8S1179(10,13), C

Figure 3: Comparison Results. For each sample we are shown the Reference sample, the number of matching alleles, and the similarity ratio. Note that in this particular example we used the “Export Greater than” option to filter out comparisons with similarity ratios lower than 60%.

Discussion

The Contamination Check tool rapidly and effectively calculates the potential risk of contamination for each profile, thereby reducing analysis time and improving confidence in the resulting findings. GeneMarkerHID software is compatible with major CE and Rapid Instrument data files, Windows® 7 – 10 OS and all major commercial human identity chemistries.

Visit <http://www.softgenetics.com/analysisCorner2.html> to view a webinar on the contamination check tool in GeneMarkerHID software

Acknowledgements

Special thanks go to the Connecticut State Police Department and the Netherlands Forensic Institute for helpful discussions on the design of this tool.