SoftGenetics Application Note May 2007

# STR Locus Specific Bracket (LSB) Analysis with JelMarker™ /GeneMarker® HID

*Tamela Serensits, Zhang Xuanhe, Yiqiang Jacie Wu, Changsheng Jonathan Liu*

## Introduction

Microsatellites, or Simple Sequence Repeats (SSRs), are polymorphic loci present in nuclear DNA that consist of repeating units 1-4 base pairs in length. They are typically neutral, co-dominant and are used as molecular markers which have wide-ranging applications in the field of genetics, including forensic identification, kinship and population studies (**1, 2**).

Short Tandem Repeat (STR) analysis is a relatively new technology in the field of forensics, having come into popularity in the mid-to-late 1990s (**3, 4**). The STRs in use today for forensic analysis are mostly tetra- or penta-nucleotide repeats (4 or 5bp repeat units). Four and five basepair repeat units result in a high degree of error-free data while being robust enough to survive degradation in non-ideal conditions.

The FBI Laboratory's Combined DNA Index System (CODIS) has identified 13 core STR loci for human identification. Several companies have developed STR testing kits based on these core loci, including ABI's Profiler/Cofiler, Identifiler, and SGM Plus kits and Promega's PowerPlex kits. In addition to these autosomal core loci, there have been significant advancements in Y-STR genotyping. Some Y-STR kits include ABI's Yfiler and Promega's PowerPlexY.

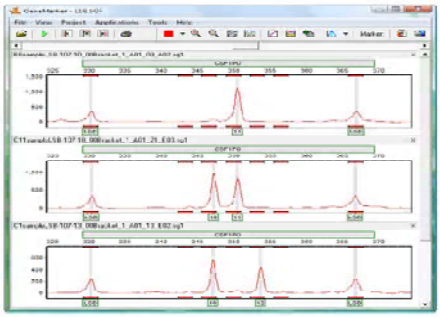

***Figure 1*** LSB Data in GeneMarkerHID

One thing in common with all of these STR analysis kits is their requirement of an Internal Lane Standard (ILS) for sizing and an allelic ladder for repeat unit identification. The ILS is a set of known size fragments labeled with an individual dye color which is multiplexed with each DNA sample (**5, 6**). An allelic ladder is a separate sample of DNA fragments which contains all reported alleles for each given loci. Typically a forensic laboratory will run between two and four allelic ladders with each 96-well data set.

Recently, a new technique has been developed to size DNA fragments based on co-migrating locus marker fragments called Locus Specific Brackets (LSBs) (**7**). There are several advantages to LSBs. First, since standard size fragments co-migrate with the sample DNA, a separate dye-color dedicated to the Internal Lane Standard (ILS) is not necessary. Secondly, run-to-run sizing precision is improved because sample DNA and LSB fragments are subject to the same environmental conditions in the capillary. Lastly, electrophoretic mobility over the 100-400bp range of typical forensic STR markers is nearly linear; therefore allelic ladder samples are not necessary for repeat unit identification. One disadvantage for LSBs, is the lack of commercial software designed to handle data that does not contain an ILS (**8**).

JelMarker is a software package designed to read image or trace files from gel electrophoresis and convert lane and band information into raw trace data for analysis with genotyping software such as GeneMarker HID. JelMarker can successfully align LSB fragments from many samples, generate a synthetic ILS based on these LSBs, and export the results for each sample in a five-color trace file format (SG1). Once the LSBs are aligned within a dataset, genetic profile analysis using GeneMarker HID can be performed.

## Procedure

1. Launch **JelMarker** software
2. Select **File → Open Traces**
3. **Figure 2. Modifying Automatic Alignment in JelMarker**
   Add LSB trace files (.fsa) to the *Open Files* box
4. Once the traces are uploaded, select **Process → Run Traces**
5. The automatically aligned data will be displayed
6. Click the **Alignment** icon to toggle between raw data and aligned data
7. When the data contains instrument errors, the user can manually adjust automatic fragment selection in the raw data window by selecting the **Edit Point** icon and using the right-click menu **Add/Delete Band** options in the gel image (***Figure 2***)
8. Once manual updates are complete and fragments are aligned correctly, go to **File → Save as SG1**
9. Choose a directory folder to save the *Original* and *Aligned* data fragments in the five-color data format (SG1).

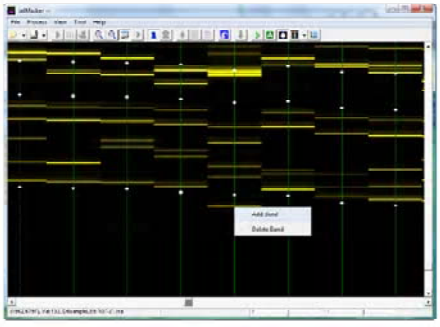

***Figure 2*** Modifying Automatic Alignment in JelMarker

## Results

JelMarker uses a special technique to align LSB data. First, we smooth the traces using a Fast Fourier Transform (FFT) calculation and then we calculate the Signal-to-Noise Ratio to determine the background. Once we've determined the noise level of the background, we identify the peak bands by intensity. Finally, we apply a linear transfer alignment to all dye colors based on the position of the last peak. The first dye color's LSBs (typically blue) are then used as the basis for the synthetic ILS.

We use the last peak as the basis of alignment because we can guarantee, with a degree of certainty, that it is an LSB and not a true allele. For troublesome data, JelMarker is flexible to allow for correction of common problems. For example, in some cases, the last peak in an electropherogram is not an LSB.

In ***Figure 3***, the fragment circled in yellow is a spike caused by instrument error. JelMarker's automatic alignment feature identified this spike as the last fragment in that lane, as indicated by the white calibration dot.

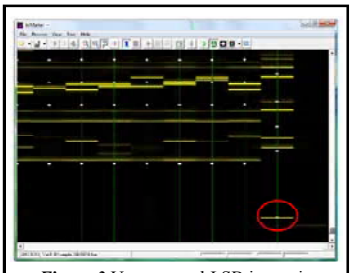

***Figure 3*** Unprocessed LSB image in JelMarker

To correct for this error, the user simply right-clicks the incorrect fragment in the gel image and selects **Delete Band**. JelMarker's smart technology will then automatically identify the next fragment up as the last fragment and realign the data with this new choice. Likewise, if the last fragment is very faint (low intensity) in the gel image, JelMarker may not identify it as the last fragment. In that case, the user simply has to locate the faint last fragment in the sample and select **Add Band** from the right-click menu. Once the data is correctly aligned, it can be exported as a five-color SG1 file. ***Figure 4*** is an example of a correctly aligned LSB image.

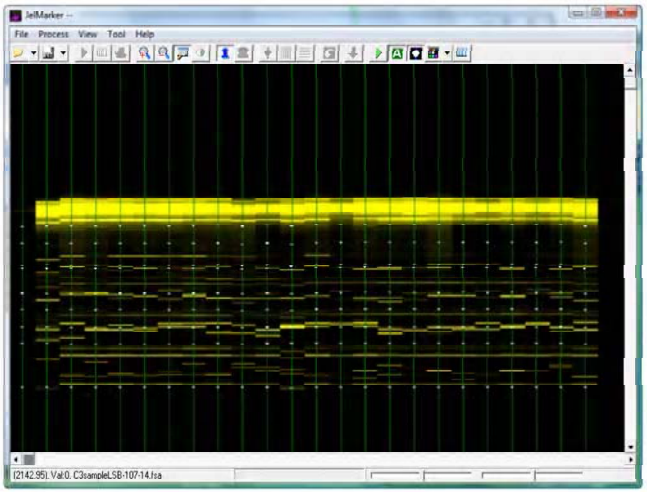

**Figure 4.** Correctly Aligned LSB Image

Two groups of SG1 files are exported from JelMarker. The first group is the *Original Data* and the second group is the *Aligned Data*. Either dataset can be uploaded to GeneMarker HID and compared to a panel. In addition to the original four-color LSB data aligned in JelMarker, the resulting SG1 files contain a synthetically produced ILS in the fifth dye color. Using this synthetic ILS as the basis for size calling, the samples can be successfully uploaded into GeneMarker HID for further analysis (***Figure 5***).

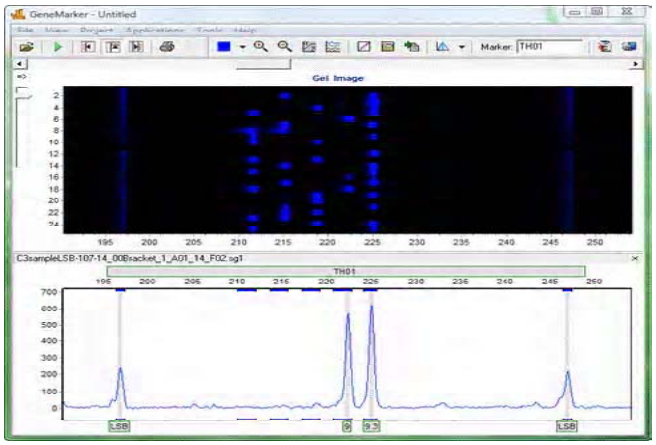

***Figure 5*** Single Locus in GeneMarkerHID

**Discussion**

JelMarker's successful alignment of LSB fragments and the subsequent STR analysis in GeneMarker HID prove to be very promising new techniques for the forensic community. Not only will the well space taken up by allelic ladders become available but additional dye colors for more loci positions could be implemented.

The result is a more discriminatory test for forensic profile analyses.

As this new genotyping technique advances, LSB design will improve allowing more loci to fit within a dye color and reducing the danger of masking true alleles (**8**). Large loci such as FGA will require extra LSBs to take into account a slight drop in the linearity of electrophoretic mobility as the fragment sizes increase.

From these preliminary findings, JelMarker and GeneMarker HID, used in combination, can simplify analysis of LSB data and allow the user to focus on obtaining accurate profiles. The Allele Report option in GeneMarker HID (***Figure 6***) displays all the necessary information to make accurate profile analysis simple!

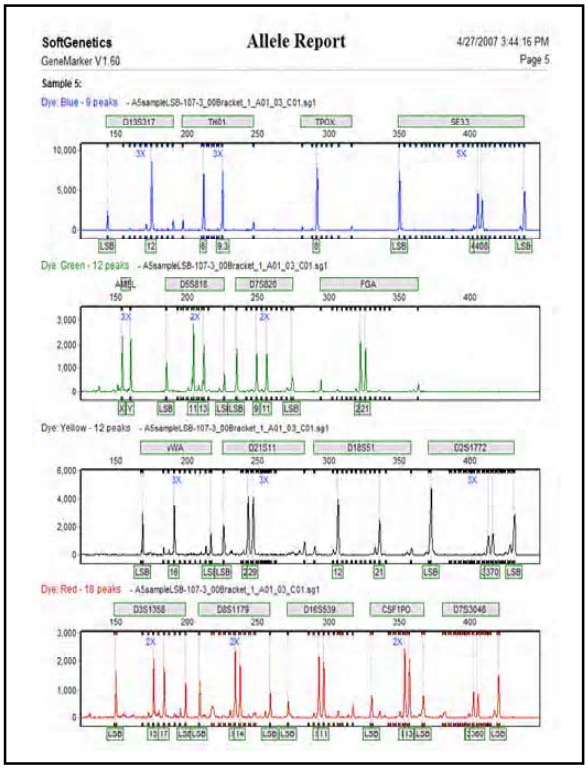


**Figure 6.** Allele Report in GeneMarker HID

**Acknowledgements**

We would like to thank Christine Picard, M.S. and Jeffrey Wells, Ph.D. of West Virginia University for providing LSB data and their continuing collaboration with SoftGenetics.

**References**

1. **Genetics and genomics of core short tandem repeat loci used in human identity testing.** JM Butler. *Journal of Forensic Science.* 2006. **51** (253-265).
2. **Inferring ethnic origin by means of an STR profile.** AL Lowe, A Urquhart, LA Foreman, IW Evett. *Forensic Science International.* 2001. **119** (17-22).
3. **Automated DNA profiling employing multiplex amplification of short tandem repeat loci**. CP Kimpton, P Gill, A Walton, A Urquhart, ES Millican, M Adams. *PCR Methods and Applications.* 1993. **3** (13-22).
4. **Evaluation of an automated DNA profiling system employing multiplex amplification of four tetrameric STR loci.** C Kimpton, D Fisher, S Watson, M Adams, A Urquhart, J Lygo, P Gill. *International Journal of Legal Medicine.* 1994. **106** (302-311).
5. **Application of dual internal standards for precise sizing of polymerase chain reaction products using capillary electrophoresis**. JM Butler, BR McCord, JM Jung, JA Lee, B Budlowe, RO Allen. *Electrophoresis.* 1995. **16** (974-980).
6. **Genotyping of forensic short tandem repeat (STR) systems based on sizing precision in a capillary electrophoresis instrument.** K Lazaruk, PS Walsh, F Oaks, D Gilbert, BB Rosenblum, S Menchen, D Schelbler, HM Wenz, C Holt, J Wallin. *Electrophoresis.* 1998. **19** (86-93).
7 **DNA bracketing locus compatible standards for electrophoresis**. PC Dau, D Liu. US Patent 6,013,444.
8. **Locus-specific brackets for reliable typing of Y-chromosome short tandem repeat markers.** JM Butler, JE Appleby, DL Duewer. *Electrophoresis.* 2005. **26** (2583-2590).

***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