

Software for Multiplex Ligation-dependent Probe Amplification (MLPA™)

Megan Kellander, Maurisa Riley, ChangSheng Jonathan Liu

SoftGenetics, LLC, 200 Innovation Blvd. Suite 235, State College, PA 16803 USA

Introduction

MLPA was introduced by the Microbiology Research Center-Holland in January 2002, and has become a rapidly growing technique used in the detection of exon deletions in BRCA1, MSH2, and MLH1 genes associated with breast and colon cancer, as well as trisomies found in Down Syndrome (1,2,3). Copy number changes in genomic DNA as well as mRNA profiling can be achieved through this new technique.

MLPA is a simple method for simultaneous quantification of up to 45 nucleic acid sequences in a single reaction. Amplification products are separated by sequence gel electrophoresis, and MLPA probes are able to distinguish between sequences that differ in only one base pair. Although the technique is efficient, inexpensive, and simple, there is a lack of programs to analyze the amplified data produced from MLPA.

New analysis and reporting functions have been added to our GeneMarker® software to increase MLPA analysis speed and reporting capabilities. GeneMarker is compatible with electrophoresis systems worldwide, including ABI files, MegaBase files, and Beckman files, as well as slab gel output. With GeneMarker's new application, copy number changes and loss of heterozygosity, which are key factors in the study of human cancers, are readily identifiable.

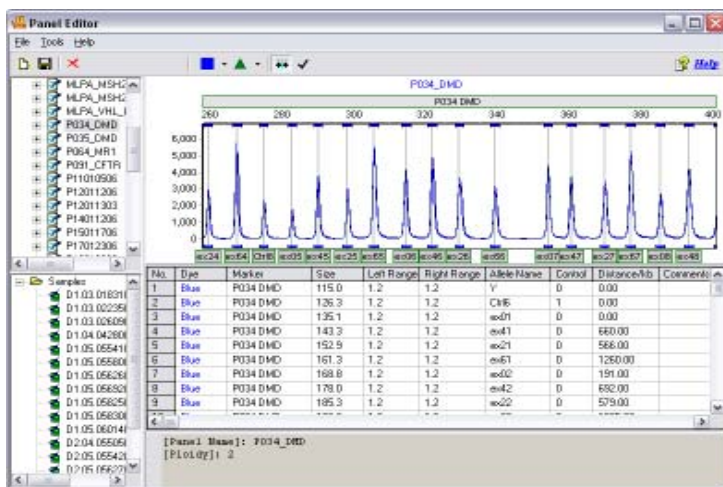


Figure 1. Panel Editor: Identifying Markers and Multiple Internal Control Probes.

Procedure

The recommended MLPA analysis settings to produce few false positives and negatives are listed below. The correct settings for MLPA should be a small global percentage plus a slightly larger or equal local percentage and the user is advised to leave the stutter peak filter on in order for the software to remove stutter peaks within 2.5 bp of each detected allele peak.

Suggested Analysis Parameters

1. **Analysis Type:** MLPA
2. **Peak Detection Threshold:** Intensity > 100; Percentage > 1 Max; Local Region % > 5 Local Max
3. **Stutter Peak Filter (%):** Left: 25 Right: 25
4. **Allele Evaluation: Score:** Reject < 0-1 Check 10 < Pass ; MLPA Normalization Method: Population Normalization.

Panel Editor:

The Panel Editor, found in the tools menu, is designed to give users the ability to indicate and label multiple internal controls as well as markers. The individual markers can be labeled by right clicking on the marker located above the electropherogram. Choose "Edit Marker" and then the user can correctly label the marker as shown below for P034 DMD. The user can also set multiple internal control probes by right clicking in the control column of the appropriate probe. Choose "Set As Control," and a window will open allowing the user to specify the control probe. In Figure 1, the control probes are labeled according to their chromosomal location as shown in the Allele column. Additionally, users may also manually change the left and right allele ranges, and then by a right mouse click select "Set Value" to automatically adjust all alleles to the new boundaries.

Pre-Defined MRC Panels:

The software features pre-defined MLPA panels based on the MRC Holland probe sets. These pre-made panels can be imported into the Panel Editor and **will need minor adjustments to fit the panel to your experimental data**. To access these panels, open the Panel Editor, select import Pre-defined panels from the File menu, and select the panel of interest from those available.

Data Normalization

Due to the variations of PCR efficiencies from small to large DNA fragments or from sample to sample, two selectable normalization methods are provided. The first normalization method is the traditional method based upon the control probes as described in reference 4.

The second method, unique to GeneMarker, normalizes peak intensities based upon the statically most probable median intensities. In order to correct for the peak intensity variation over size, an exponential function $a \cdot e^{-bz}$ is used to fit to the square root of peak intensities, where z is size, and a and b are fitting constants.

The normalization using the control probes is shown in figure 2. This correction removes the trend of dropping intensities as the DNA fragment size increases, and sets the height ratios of control probes to approximately 1. However the trend of peak intensities vary greatly from one sample to another with the internal control probes. We have found that the use of fewer control probes often results in large errors in the intensity normalization.

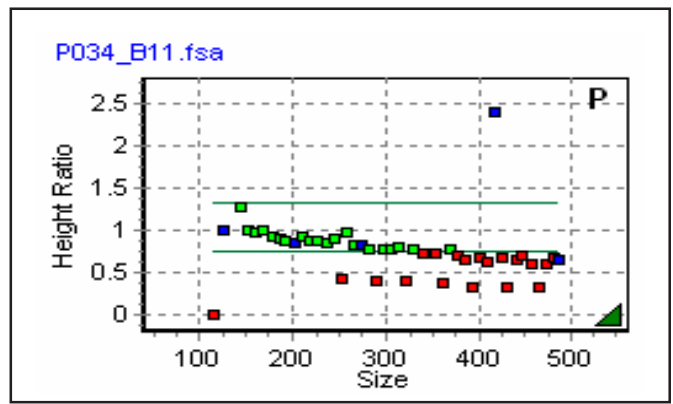


Figure 2. Traditional Control Probe Normalization.

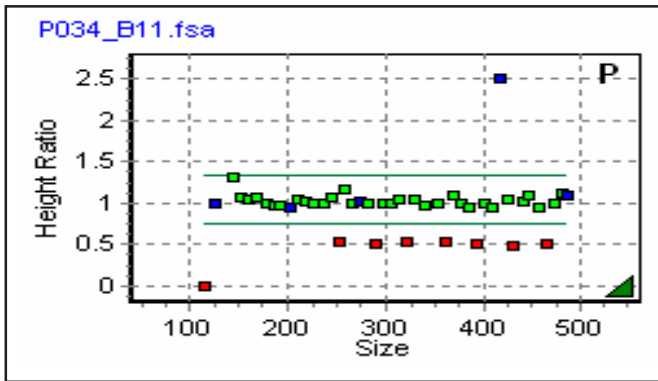


Figure 3. SoftGenetics' Population Normalization.

SoftGenetics' "**Population Normalization**" addresses these problems (Figure 3). Median peak intensities are derived from the first nine data points, then sliding to data points 2-10, 3-11, etc. to ascertain the local median intensities. Outliers are rejected after applying a median filter. All of the probes (control and test) with the median intensities are then used to fit the exponential function. **This methodology creates higher accuracy and lower false positive rates.**

Results

After intensity normalization, the data is plotted in two formats: ratio and regression type. First, the intensity ratios of identical probes from the patient sample can be compared to that of the control samples. Deletions and duplications appear as outliers from the data set, as points outside of the threshold lines. (Figure 4)

The second regression plot method provides the peak intensity deviations of the patient compared to that of the control sample, as defined by the user. The square root of peak intensity is used to form the best fit of the peak intensity deviation exponential function. To calculate the regression and reject outliers: 80% of peaks with smaller deviations are retained in the regression line calculation, and 20% are rejected initially. GeneMarker then iterates multiple times to reject and retain peaks, with a confidence of 99%. The 99% confidence level of the regression is shown within the green lines; duplications > 1.33 and deletions < 0.75 of DNA copy number are shown as red dots outside of the red lines. The points with normal DNA copy number are green dots, and the control probes are blue dots. (Figure 5)

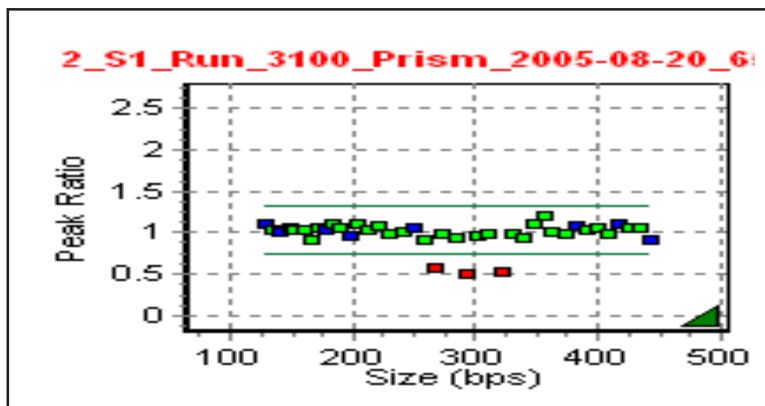


Figure 4. The height ratio plot of a patient sample compared to that of a control samples

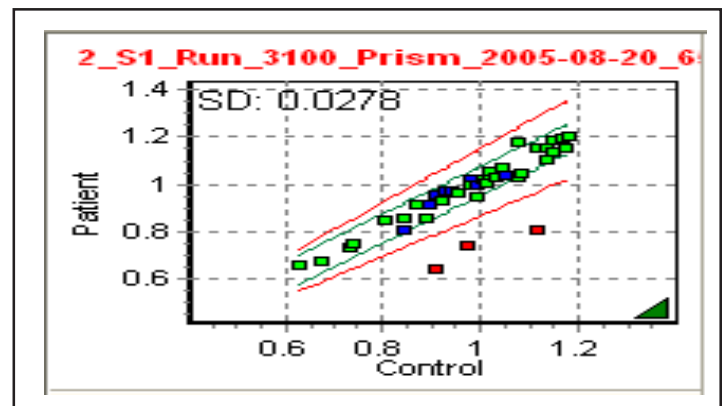


Figure 5. The regression plot of the square root of peak intensity deviation of a sample compared to a control.

Patient Report: GeneMarker provides a patient report that can be printed or saved as a PDF file:

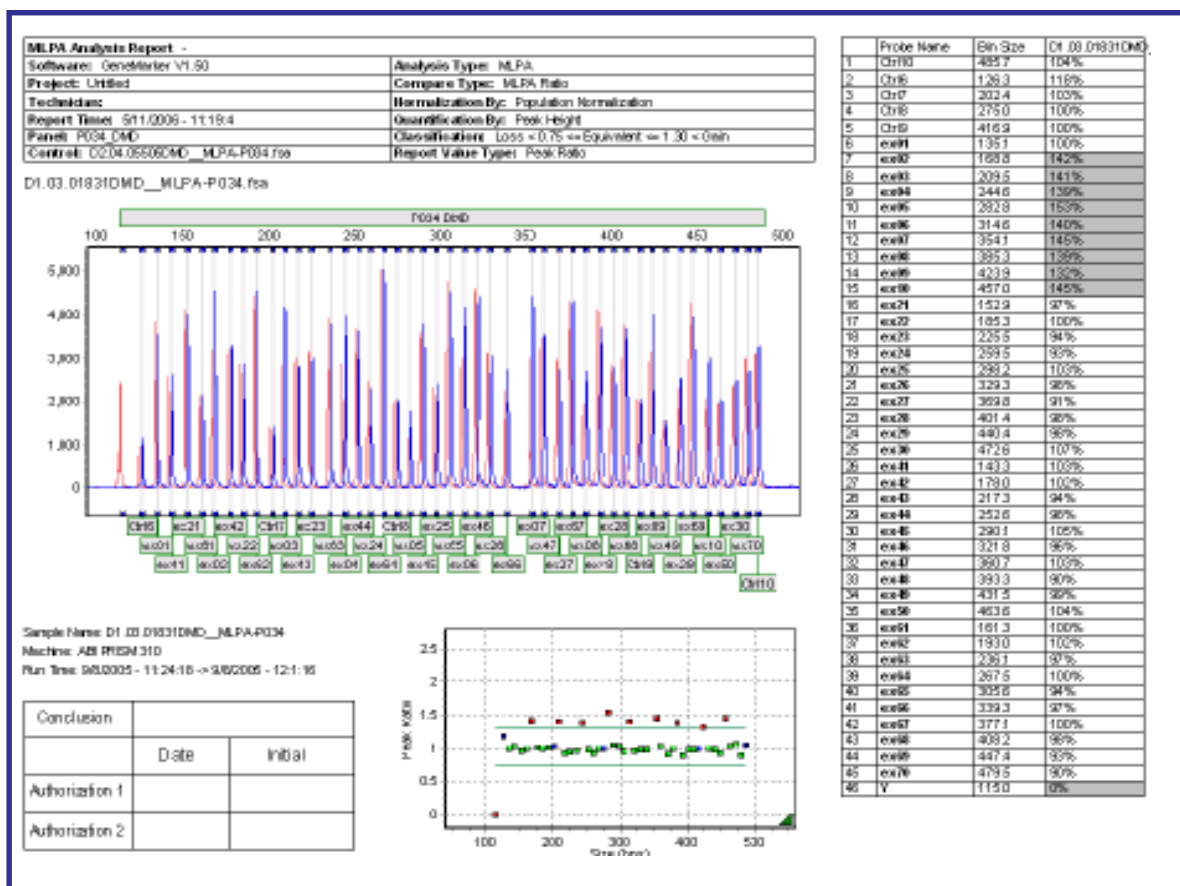


Figure 6: The patient report includes sample ID, analysis parameters, report, graph and electropherogram by sample.

Discussion

Various techniques including DGGE (Denaturing Gradient Gel Electrophoresis), DHPLC (Denaturing High Performance Liquid Chromatography), and SSCA (Single Strand Confirmation Analysis) effectively identify SNPs and small insertions and deletions. MLPA, however, is one of the only accurate, time efficient techniques to detect genomic deletions and insertions (one or more entire exons), which are frequent causes of cancers such as hereditary non-polyposis colorectal cancer(HNPCC), breast, and ovarian cancer. MLPA can successfully and easily determine the relative copy number of all exons within a gene simultaneously with high sensitivity and GeneMarker has the tools for necessary quantification and reporting of the data.

Acknowledgment

We would like to thank Resie Vervenne, Gerard Pals and colleagues from the VU Medical Center in Amsterdam for their collaboration on the development of the GeneMarker's MLPA analysis.

References

- Genomic deletions of MSH2 and MLH1 in colorectal cancer families detected by a novel mutation detection approach.** JJP Gille, FBL Hogervorst, G Pals, JTh Wijnen, RJ van Schooten, CJ Dommering, GA Meijer, ME Craanen, PM Nederlof, D de Jong, CJ McElgunn, JP Schouten and FH Menko. *British Journal of Cancer*. 2002. **82**(892-897).
- Large Genomic Deletions and Duplications in the BRCA1 Gene Identified by a Novel Quantitative Method.** Frans B.L. Hogervorst et. al. *Cancer Research*. 2003. **63** (1449-1453).
- Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification.** Jan P. Schouten, Cathal J. McElgunn, Raymond Waaijer, Danny Zwijnenburg, Filip Diepvens and Gerard Pals. 2002. *Nucleic Acids Research*. Vol. 30, No.12 **e57**.
- Regression-enhanced examination of MLPA(REX_MLPA): towards automated and objective scoring.** Lampros A. Mavrogiannia, David J. Cockburn, Helen Snowden, Graham R. Taylor. Yorkshire Regional DNA Laboratory, St. James's Univeristy Hospital, Leeds.

MLPA is a registered trademark of MRC, Holland

GeneMarker is a registered trademark of SoftGenetics LLC

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