Evaluating Genetic Analysis Systems: Microsatellite Analysis

Introduction

Microsatellites, also known as short tandem repeats (STRs), are polymorphic DNA loci that contain a 2 to 7 nucleotide repeat sequence. The number of repeat units for a given locus may differ resulting in alleles of varying lengths. Allelic variation, the number of repeats, and allelic frequencies are available for thousands of markers across numerous organisms. The ability for researchers to choose from such a large selection of highly informative markers has made microsatellite analysis a widely accepted tool for linkage studies, association studies, and identification of individual organisms.

Although microsatellite analysis has become routine due to the availability of a large number of highly informative markers for researchers to select from, there are some additional considerations necessary when choosing a capillary electrophoresis system for microsatellite analysis. Optimization of reaction and run conditions, as well as assay design are critical in ensuring high-quality data and minimizing costly reruns. Managing large amounts of microsatellite data requires an efficient and highly automated workflow from data production and analysis, to allele scoring.

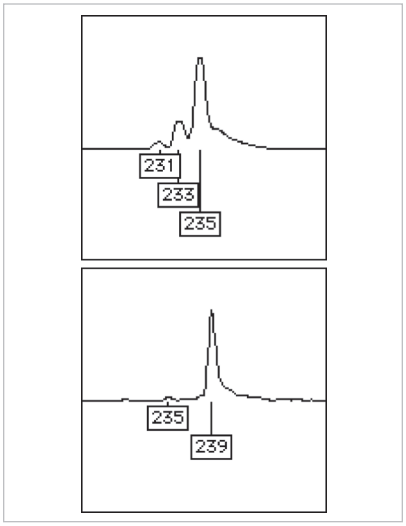

Figure 1. During the PCR amplification of di-, tri-, or tetranucleotide microsatellite loci, minor products that are 1-4 repeat units shorter than the main allele peak are produced. These stutter peaks are more prevalent in dinucleotide repeat loci as longer length repeat units produce less stutter.

Microsatellite Reaction Optimization

One of the most common problems encountered during microsatellite analysis is poor or non-specific amplification. Microsatellite analysis projects often necessitate the interrogation of up to hundreds of loci for a given sample. The potential to manage a large number of reactions requires efficient primer design and robust universal reaction conditions. Efficient primer design and universal reaction conditions minimize the need for costly reruns due to failed PCR reactions. Applied Biosystems microsatellite-based kits are optimized to provide the most robust reaction conditions possible. When performing custom microsatellite applications, these variables must be addressed to ensure the best possible results.

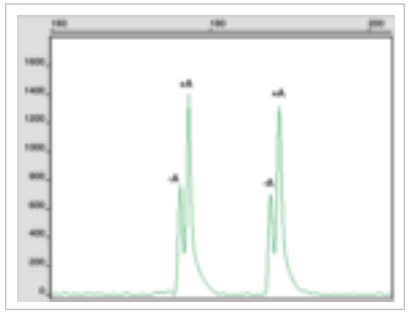

Figure 2. Split peaks resulting from an incomplete 3' A nucleotide addition.

Microsatellite Reaction Artifacts

Most PCR based systems produce artifacts that can complicate data interpretation. Two well-characterized artifacts of PCR amplification observed in microsatellite analysis are stutter, and non-templated 3' A nucleotide additions, sometimes called "Plus A" additions. Chemistry and software can be optimized to differentiate these artifacts from the true allele peaks.

Stutter artifacts are observed as multiple artifact peaks preceding the true allele peak. The number of peaks and intensity are proportional to the length of the repeat and the overall number of repeats in the PCR product. Although stutter is well characterized and reproducible, it can make data analysis difficult. Applied Biosystems

GeneMapper® Software is designed to easily identify and filter out stutter artifacts for accurate scoring of the true alleles (Figure 1).

"Plus A" additions also increase the complexity of the peak pattern making true allele peak recognition difficult (Figure 2). This artifact is locus dependant and greatly impacted by reaction conditions. Applied Biosystems microsatellite-based kits are optimized to promote "Plus A" additions, minimizing the occurrence of split peaks caused by an incomplete A nucleotide addition, and thereby generating a more consistent allele peak pattern. Optimizing reactions to promote "Plus A," and automatically filtering out stutter and "Plus A" peaks in GeneMapper® Software produces more accurate, true allele identification.

Multiplexing

The time and cost of microsatellite analysis projects can be minimized through co-electrophoresis of multiple markers in each capillary. The ability to multiplex is dependant on run condition optimization, differences in the relative sizes of fragments, and the number of dye labels compatible with the capillary electrophoresis system.

Fluorescent labeling enables the analysis of many independent loci in the same capillary injection using color and size to distinguish between fragments. When considering a dye combination, the number of dyes available for fragment labeling and the selection of more spectrally resolved dyes are important in avoiding analysis complications due to spectral overlap. While most electrophoresis systems have the capability to analyze four dye chemistries, Applied Biosystems five dye chemistry system is designed to increase multiplexing capacity, improve spectral resolution, and decrease project costs by providing a 33% increase in throughput over standard four dye chemistry (Figure 3).

Data Production

Another important component in the microsatellite project workflow is data production. A project's time and cost are impacted by the system's ability to rapidly produce reliable, high-quality results in an automated fashion. The ABI PRISM® 3100 and 3100-*Avant* Genetic Analyzers allow users to easily schedule and prioritize runs for maximum flexibility, Data Collection Software promotes quick run set-up, while multiple pre-optimized instrument configurations deliver the highest resolution and precision for data scoring with minimal manual intervention.

Precision and resolution are a function of fragment size, dye chemistry, sizing algorithm enhancements, and optimization of electrophoretic separation conditions. Parameters to consider when optimizing separation conditions include temperature regulation, polymer type, capillary array length, and the quality and ease-of-use of consumables designed to work with the instrument.

A common misconception surrounding DNA fragment sizing is that the calculated size of a DNA fragment is equivalent

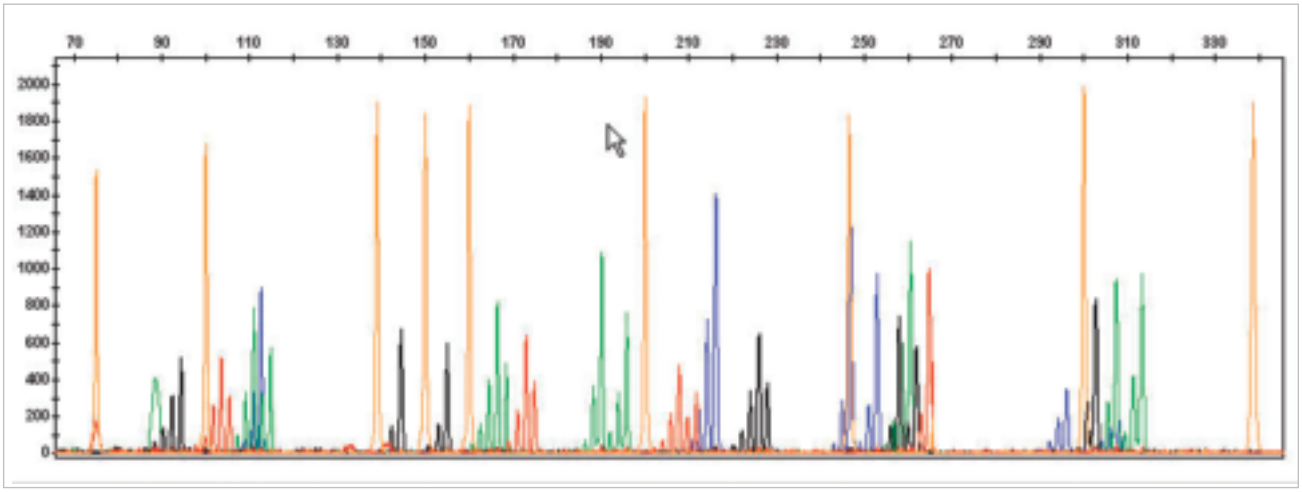

Figure 3. Five dye chemistry increases the number of markers that can be run together in a single capillary, thus maximizing throughput. Above is an example of eighteen microsatellite loci co-electrophoresed in a single capillary.

to the actual length of the fragment in base pairs. Since the electrophoretic mobility of DNA is sequence dependent, DNA fragments of the same length can have different mobilities based upon their individual sequences and can therefore vary in their calculated sizes. Hence, sizes are calculated based on the mobility of the fragment and not specifically its length in base pairs. Reproducibility and precision of results are much more important than accuracy in microsatellite-based projects. Microsatellite analysis project results are dependent on the ability to get the same answer no matter how many times a sample is run. The 3100 and 3100-*Avant* systems are optimized to provide researchers with the most reproducible data possible for any microsatellite application (Figure 5).

Data Management

Microsatellite analysis projects can involve data from hundreds to thousands of samples depending on whether few individuals are screened with many loci, or many individuals are screened with only a few loci. To economically manage the wide range of microsatellite project sizes, software data management workflow must be optimized. Effective automated genotyping software should offer flexibility, high precision for fragment size and allele scoring, automation, and advanced data management features. GeneMapper® Software employs different modes of operation to manage a wide variety of project sizes. Advanced algorithms are optimized to recognize and filter amplification chemistry artifacts, such as "Plus A" and stutter peaks, and differentiate between microsatellite repeats of varying lengths. GeneMapper® Software also flags questionable sample files for manual review via Genotyping quality (GQ) scores, and produces sortable tables in an easy to interpret format. These software benefits allow scientist to review large amounts of microsatellite data accurately and rapidly (Figure 4).

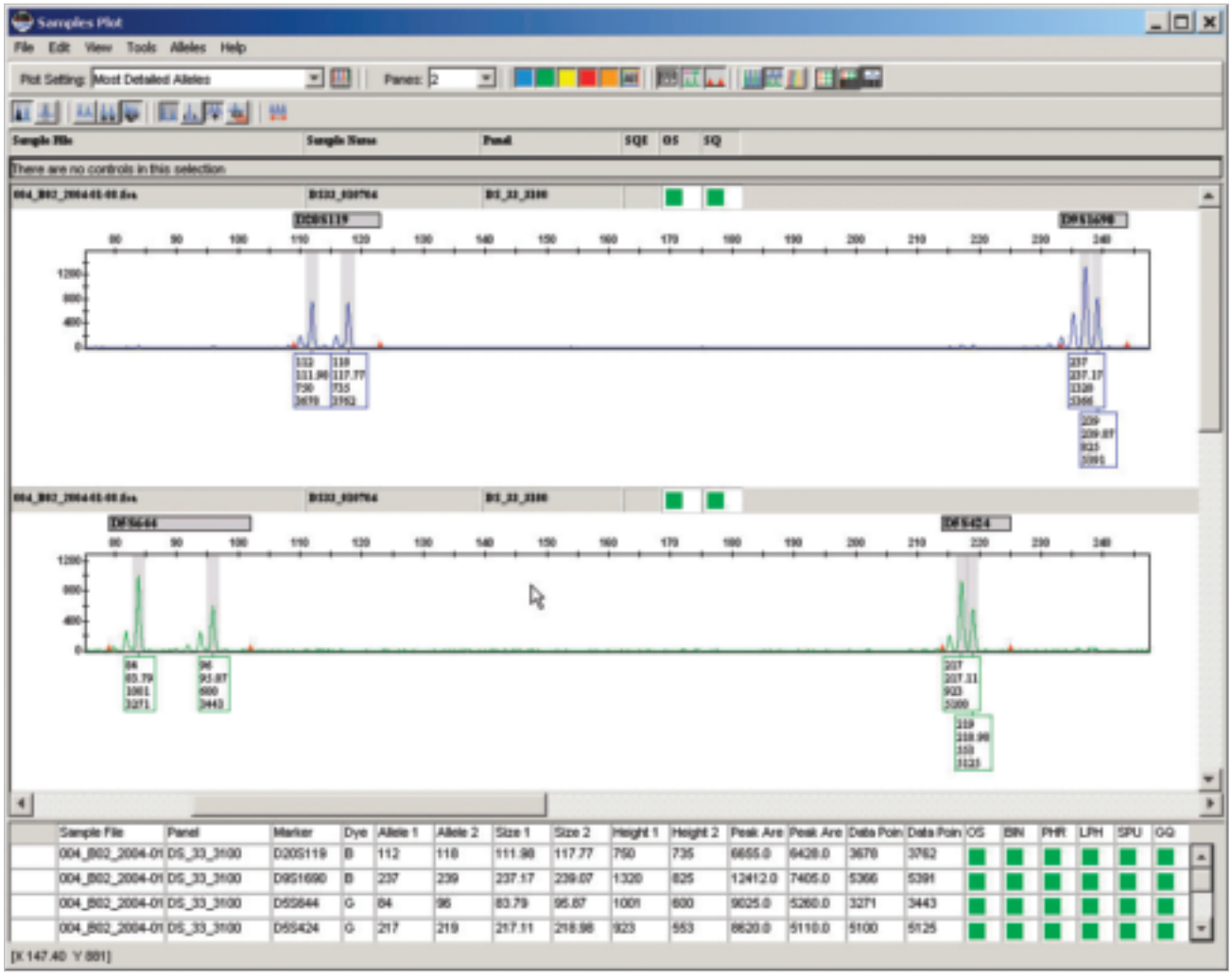

Figure 4. Advanced data management software produces intuitive data tables for fast allele scoring and data quality review.

Conclusion

Features such as five color chemistry, pre-optimized run conditions, and intuitive software with flexible and intelligent data management options make the ABI PRISM® 3100 and 3100-*Avant* Genetic Analyzers great systems for both small and large scale microsatellite studies. Integrated with GeneMapper® Software, the 3100 series instruments allow one-button operation for data collection, fragment size calling, and allele scoring in both graphical and tabular form. These features enable researchers to generate large amounts of high-quality data with minimal hands-on time and manual data review, thus dramatically reducing the time and cost of a microsatellite project.

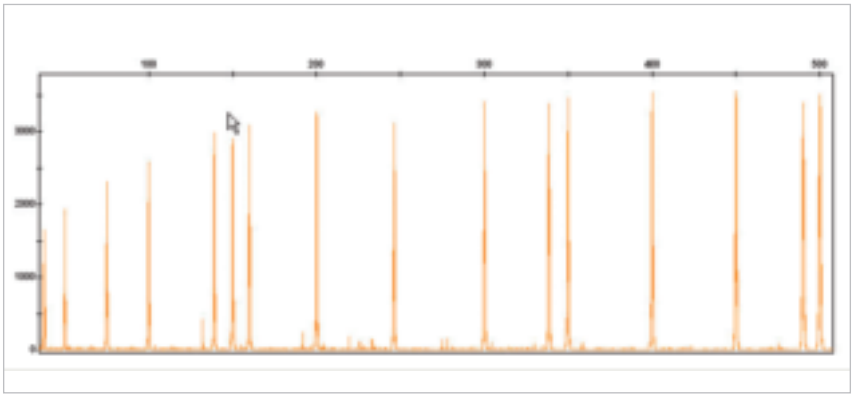

Figure 5. Overlay of sixteen size standard electropherograms displaying system sizing precision.

iScience™

iScience. To better understand the complex interaction of biological systems, life scientists are developing revolutionary approaches to discovery that unite technology, informatics, and traditional laboratory research. In partnership with our customers, Applied Biosystems provides the innovative products, services, and knowledge resources that make this new, **Integrated Science** possible.

Worldwide Sales Offices
Applied Biosystems vast distribution and service network, composed of highly trained support and applications personnel, reaches 150 countries on six continents. For international office locations, please call the division headquarters or refer to our Web site at **www.appliedbiosystems.com**

Applera is committed to providing the world's leading technology and information for life scientists. Applera Corporation consists of the Applied Biosystems and Celera Genomics businesses.

Headquarters
850 Lincoln Centre Drive
Foster City, CA 94404 USA
Phone: 650.638.5800
Toll Free: 800.345.5224
Fax: 650.638.5884

For Research Use Only.
Not for use in diagnostic procedures.

Printed in the USA, 05/04
P+s Publication 106MI51-01

Applied Biosystems