

ABI PRISM® GeneMapper™ High Performance Genotyping Software for Microsatellite- and SNP-Based Genotyping Applications

Overview

With the advent of automated fluorescent genotyping using dye-label primers and “size- and marker-multiplexing” in a single lane, large-scale genotyping has become a routine laboratory process. ABI PRISM® GeneMapper software is a high-performance genotyping tool designed to provide high throughput and accuracy required by these automated genotyping processes.

Assigning Process Component Based Quality Values (PQV)

Comprehensive data filtering—to substantially minimize data examination time

Because of problems arising from poor sample preparations, presence of PCR artifacts, irregular stutter bands, and incomplete fragment separations, a typical genotyping project requires manual examination of almost all allele calls. We designed GeneMapper software to eliminate this bottleneck by substantially minimizing the need to review data.

GeneMapper software assigns a series of Process Component Based Quality Values (PQV) that enable users to rapidly pinpoint the basis for failed allele calls, as shown in Figure 1. The system then combines these PQVs with the information from the peak determination algorithms to give an overall genotype quality value (GQ).

Using GeneMapper software, called-alleles fall into three categories.

1) High-quality calls. These calls can be accepted without reviewing.

2) Low-quality calls. These calls represent samples with contamination and other failures, which render the resulted data impossible to call. This data can quickly be rejected without reviewing.

3) Intermediate-quality calls. These calls fall between high- and low-quality calls. These calls can be reviewed and edited conveniently in GeneMapper software.

Automatic Quality Assurance Using a Control Concordance PQV

—This particular PQV metric allows users to confirm the correct allele calling of an internal control. As a result, users are assured that unknown samples in the same project are analyzed correctly using all the preset analysis parameters.

High Throughput

GeneMapper software can process over 48,000 genotypes in 6 hours,

increasing productivity six-fold, compared to size and allele calling with GeneScan® and Genotyper® software (Figure 2). Coupled with the use of the PQV system, GeneMapper software substantially minimizes the amount of time users spend analyzing data.

Accurate Allele Calling

GeneMapper software applies a series of novel, patent-pending size calling, allele calling, automatic bin-building and auto-panelizer algorithms to assure allele calls are highly accurate. We have optimized these algorithms using data containing over 50,000 genotypes collected from laboratories worldwide. As a result, GeneMapper software supports a variety of genotyping applications using different microsatellite and SNP markers.

Marker	Sample	Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6	Allele 7	Allele 8	Allele 9	Allele 10	Quality
D1S80	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D2S1328	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D3S1358	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D4S1393	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D5S1455	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D6S1566	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D7S1855	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D8S1179	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D9S1122	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D10S1248	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D11S1418	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D12S1738	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D13S322	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D14S643	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D15S659	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D16S1106	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D17S1196	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D18S81	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D19S2538	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D20S1117	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D21S11	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D22S413	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D23S448	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D24S242	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D25S1678	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D26S1314	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D27S1330	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D28S147	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D29S1934	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D30S1106	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D31S1631	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D32S1328	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D33S1705	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D34S1031	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D35S1330	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D36S1330	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D37S1330	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D38S1330	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D39S1330	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D40S1330	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D41S1330	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D42S1330	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D43S1330	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D44S1330	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D45S1330	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D46S1330	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D47S1330	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D48S1330	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D49S1330	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D50S1330	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D51S1330	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114	116	118	120	122	124	126	128	130	Pass
D52S1330	Sample 1	112	114	116	118	120	122	124	126	128	130	Pass
	Sample 2	112	114									

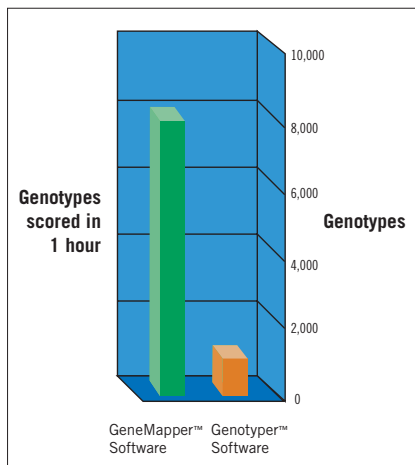


Figure 2. High-throughput GeneMapper software processes 48,000 genotypes in 6 hours.

Ease of Use

GeneMapper software provides interactive tools to

- Generate new and custom markers
- Edit existing markers
- Set up analysis methods and custom table display profiles
- Edit allele calls

Once analysis methods and marker definitions are established, the data analysis process is fully “transparent” to users. We have streamlined these decision-making steps so that repetitive size and allele calling is simply a single mouse-click operation.

Fast Editing

GeneMapper software includes a complete set of editing tools that let you edit called peaks, markers, and bin definitions. These are carefully designed graphical user interfaces allowing rapid editing of allele calls and bin definitions. In addition, a novel SNP panel manager allows flexible viewing of SNP traces, genotype assignments and bins in multiple colors for on-the-fly bin editing.

Data Sharing

GeneMapper software allows users to access data from other local GeneMapper databases for convenient

data sharing without using client-server infrastructure and incurring the associated IT support cost. Data access is controlled by passwords set by local GeneMapper software users.

One Software Tool for Multiple Applications

Many SNP-based genotyping users also perform microsatellite-based genotyping. GeneMapper software is designed to allow these users to perform both linkage analysis and SNP validations using a single software, without needing to learn and validate multiple software tools.

Application-Specific Features

Specific capabilities of microsatellite-based genotyping —

Supports Linkage Mapping of Diploid and Polyploid Organisms

Supports allele calling of organisms with up to 50 alleles. GeneMapper software users can genotype both diploid (such as human) and polyploid organisms (many plants and animals).

Supports Di-, Tri-, and Tetranucleotide Markers

GeneMapper software allows users to call alleles from sample files containing highly multiplexed genotypes. Di-, tri- and tetranucleotides microsatellite markers are supported. Tri and tetra

nucleotide-based analysis parameters are user-adjustable to allow maximum analysis flexibility.

Features Automatic Bin Builder (ABB) Algorithm

An Automatic Bin Builder (ABB) algorithm is designed to enhance the allele calling process for linkage mapping applications. Using the ABB, GeneMapper software creates relevant bins based on the chosen panel information, and successive allele calls from the sample file collection. The ABB continues to optimize the distances between bins and the precise positions of bin centers. As more sample files are processed, the uncertainty of the bin boundaries and positions of the bin centers is minimized. Finally, the collection of called-alleles will be assigned to respective bins, and the appropriate bin centers and bin width will be defined by the ABB, automatically (Figure 3).

Employs Unique Microsatellite-Based PQV Parameters

Specific PQVs are available for analysis using di-, tri-, and tetranucleotides for optimal performance. Up to 20 PQVs can be utilized to monitor the genotyping process.

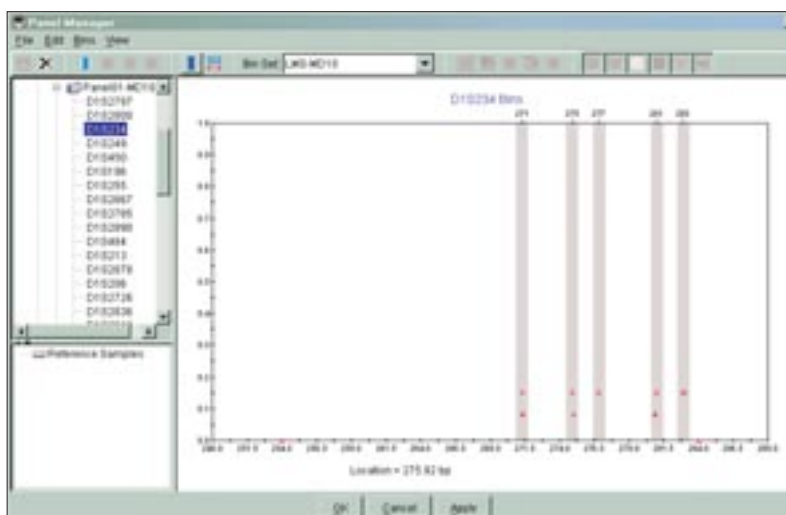


Figure 3. The Automatic Bin Builder (ABB) algorithm automatically creates precise bins for high-throughput microsatellite-based genotyping applications.

Specific capabilities of SNP-based genotyping –

Displays Electropherograms with Multiplexed Markers and Bin Overlays

GeneMapper software displays electropherogram traces in multiple colors. GeneMapper software also displays electropherograms with multiplexed markers and bin overlays. Marker boxes are displayed immediately below peak clusters to facilitate quick visualization of allele calls from highly multiplexed samples (Figure 4).

Employs unique SNP-Based PQV Parameters

Parameters

Unique PQV parameters are available to filter out problematic data and automatically examine the “rules” of SNP genotyping data analysis. For example, a Double Peak (DP) PQV automatically flags data containing two peaks with the same color within a bin. With the additional PQV parameters, GeneMapper software users can quickly filter highly multiplexed data generated from several 3100 or 3700 systems (Figure 5).

Available Auto-Panelizer Features

The available Auto-Panelizer can generate optimized markers multiplexed for a specific panel to maximize the number of markers in a panel while minimizing dye-color overlaps. Proper functioning of the Auto-Panelizer features requires the use of a new ABI PRISM® chemistry kit to be released in the near future. When this chemistry is available, GeneMapper software users can utilize the Auto-Panelizer features already available in GeneMapper software v 2.0.

Complete System Solution

GeneMapper software contains a universal size-calling component that is compatible with the ABI PRISM® 310, 377, 3100, and 3700 Genetic Analyzers. Windows NT-based sample files obtained from these instruments can be analyzed directly by GeneMapper software.

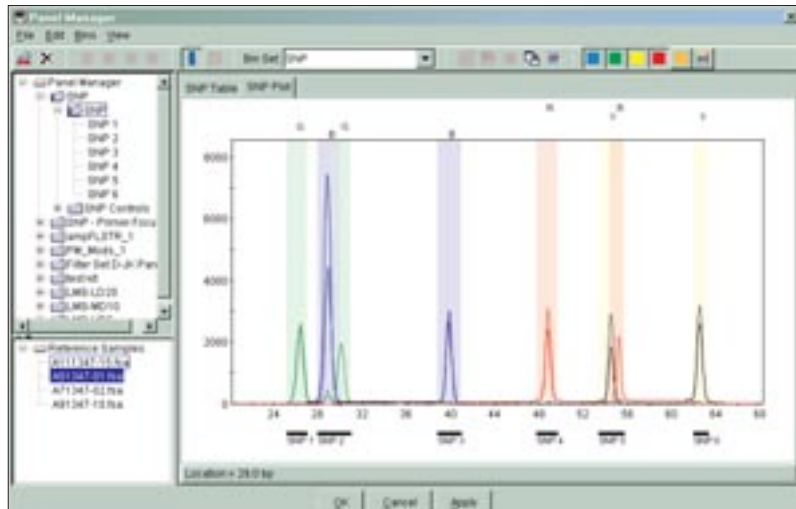


Figure 4. GeneMapper software allows simultaneous viewing of multicolor SNP traces overlaid with bins and with genotype assignments under corresponding bins for fast bin editing.

Macintosh® platform-based sample files need to be converted to Windows NT® platform-based files, using a conversion utility from GeneMapper software, prior to the beginning of the data analysis process.

GeneMapper software is fully compatible with ABI PRISM® GeneScan™ Size Standards, Linkage Mapping Sets, and SNaPshot® Multiplex kits, Linkage Mapping Sets and associated GeneScan™ size standard. GeneMapper Database is preloaded with the size standards, markers, and panel information from the Linkage Mapping Sets.

Analysis Accuracy and Concordance Comparing GeneMapper software to GeneScan and Genotyper Software

- Files containing over 8,000 Genotypes were obtained from ABI PRISM® 377 DNA Sequencers in batches of 96 samples.
- The accuracy of the allele calls was checked by comparing GeneMapper software allele calls with those made by an independent user who previously determined the allele calls using Genotyper® software v 3.6 and GeneScan® software v 3.5.

Results

- GeneMapper software is 99.99% accurate when comparing high-quality calls with those made manually.
- In addition, 99.8% of calls made by GeneMapper software agreed with the results obtained by the second-user calls using GeneScan and Genotyper software. The 0.2% discordance was due to manual scoring errors.

Conclusions

- With GeneMapper software, a single operator can score almost 200,000 genotypes in a 24-hour period.
- One operator can easily keep pace with the output from several 3100 or 3700 systems.
- The allele-calling algorithms in GeneMapper software assure highly accurate data analysis for SNP-based and microsatellite-based genotyping.
- GeneMapper software's ease of use, timesaving features, and PQV system makes it also beneficial for medium-throughput users of 377 DNA Sequencers and 310 Genetic Analyzers.

