illumina®

GenomeStudio® 2.0 Data Analysis Software

A comprehensive data analysis tool supporting microarray genotyping experiments.

Highlights

- **Broad Compatibility**
  Analyze data from any Illumina array-based genotyping assay
- **Versatile Analysis for Any Genome**
  Genotype both diploid and polyploid organisms
- **Intuitive Project Creation**
  Create and manage genotyping projects with an intuitive interface
- **Customizable Data Visualization**
  Examine data from a genome-wide view to single-base resolution with various display tools

Introduction

Illumina offers a broad spectrum of innovative and powerful Infinium® array-based genotyping assays. GenomeStudio 2.0 Data Analysis Software supports these diverse assays with an integrated DNA-to-data workflow for all Illumina arrays. GenomeStudio Software features performance-optimized tools and a user-friendly graphical interface that enable researchers to convert microarray data into meaningful results quickly and easily.

Versatile Analysis for Any Genome

GenomeStudio Software supports the spectrum of Illumina array-based genotyping applications, enabling analysis of both diploid and polyploid organisms (Figure 1).

Diploid Genotyping

Genotyping data for diploid organisms generated using Infinium assays on the iScan® System are analyzed in the default GenomeStudio Genotyping Module. This module uses trusted algorithms to normalize, cluster, and call genotypes. Data quality is monitored in real time with internal controls and other QC functions. Individual single nucleotide polymorphisms (SNPs) can be viewed as SNP Graphs (Figure 2). Users interested in deeper investigation of individual SNPs, particularly with nonstandard samples (samples with low input or whole-genome amplified samples) can edit SNP Graphs. Genotype summary statistics and results are automatically reported and exportable in a customizable format for use in third-party downstream analysis software.

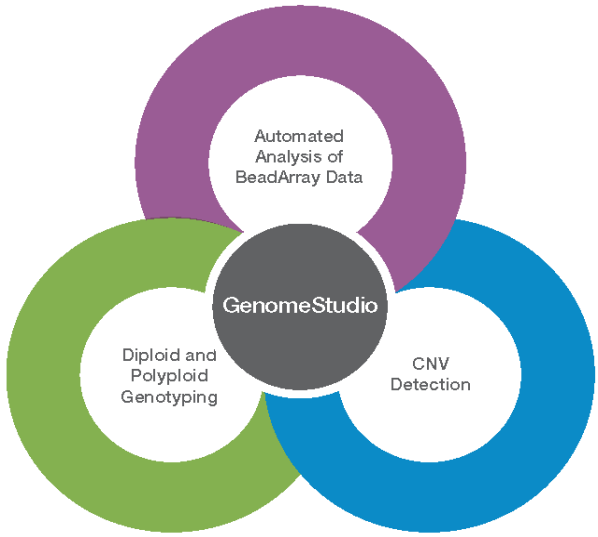

Figure 1: GenomeStudio 2.0 Data Analysis Software—GenomeStudio Software supports automated analysis of data generated using any Illumina array-based genotyping assay and features diploid and polyploid genotyping and CNV detection.

Structural variation is identified using genotyping probes to measure relative intensity shifts to detect breakpoints in copy number variations (CNVs) and loss of heterozygosity (LOH) along chromosomal segments. CNV and LOH are reflected through analysis of a moving window in log R ratio and B allele frequency that is displayed relative to marker position along chromosomal segments.

Polyploid Genotyping

Analyzing data obtained from genotyping polyploid organisms requires special considerations that depend on the level and type of ploidy of the organism. Polyploid genotyping is also dependent on the expectations of segregation in the included samples. For example, in applications involving polyploid plants, genotyping is often performed for lines that are fixed for alternate alleles at a particular SNP locus. Handling of these sample sets is integrated into the GenomeStudio Polyploid Genotyping Module by allowing the investigator to allocate the expected sample sets. By reducing the expectation of seeing heterozygotes, genotypes are called seamlessly. With the Polyploid Genotyping Module, analyzing genotypes of polyploid organisms is performed in a streamlined fashion with automated genotype calling.

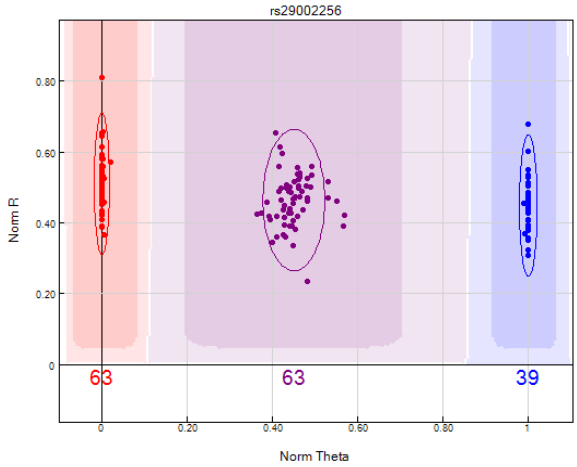

Figure 2: Genotyping Module SNP Graph—The graphical display of results in the GenomeStudio Genotyping Module is a SNP Graph with data points color coded for the call (red = AA, purple = AB, blue = BB). Genotypes are called for each sample (dots) by their signal intensity (Norm R, Y-axis) and Allele Frequency (Norm Theta, X-axis) relative to canonical cluster positions (dark shading) for a given SNP marker.

The Polyploid Genotyping Module implements the density-based spatial clustering of applications with noise (DBSCAN) algorithm and the new PolyGenTrain algorithm to call as many clusters as expected for up to octoploid organisms. After generating clusters, genotypes can be called in a fully automated fashion. Similar to the Genotyping Module for diploid organisms, individual SNPs can be graphically displayed (Figure 3), and genotype summary reports are exportable for further downstream analysis.

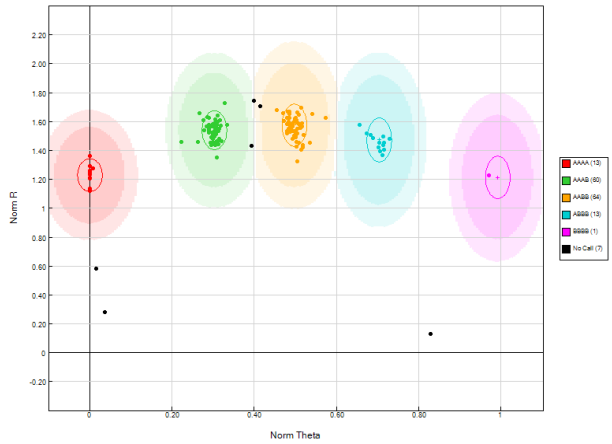

Figure 3: Polyploid Clustering Module SNP Graph—The graphical display of results in the GenomeStudio Polyploid Module is a SNP Graph with data points color coded per cluster for the call. The number of clusters expected depends on user-defined parameters for the level of ploidy (eg, tetraploidy, hexaploidy, octoploidy) and type of ploidy (eg, alloploidy vs. autoploidy).

Intuitive Project Creation Interface

GenomeStudio Software makes it easy to create projects and identify input data locations with an intuitive wizard interface. After a project is created, users can view and display all data associated with each experiment type easily.

Users with the optional Illumina Laboratory Information Management System (LIMS), sample tracking, and robotic automation control can take advantage of GenomeStudio Software integration for increased efficiency in overall project management for all Infinium genotyping assays. These platforms, custom designed for Illumina workflows, allow labs to maximize their throughput.

Simple Graphical User Interface

GenomeStudio Software provides a set of intuitive graphical user interfaces (GUI) and data visualization features for the control and display of genotyping and clustering results (Figure 4). GenomeStudio Software enables thorough understanding of the large data sets generated by Illumina genotyping assays by providing multimodal examination from a broad, genome-wide view down to a fine-grained single-base view. GenomeStudio Software displays results at all scales to enable researches to examine high-resolution genome-wide data effectively.

Versatile Data Visualization

Displaying data in the Illumina Genome Viewer (IGV) quickly reveals genome-wide orientation and broad trends. Chromosome or region-level trends of array data are visualized in the Illumina Chromosome Browser (ICB). To find trends across samples, markers, or different assays, the GenomeStudio framework provides a wide range of graphical plotting and display tools (Table 1).

When trends or interesting regions are identified with graphical analysis tools, looking at individual data points becomes essential. GenomeStudio Software supports this single-base level of analysis of individual SNP genotypes with table displays (Table 1). Table views are customizable for sorting and to show or hide various data categories. Table data is exportable in formats compatible with other downstream analysis tools.

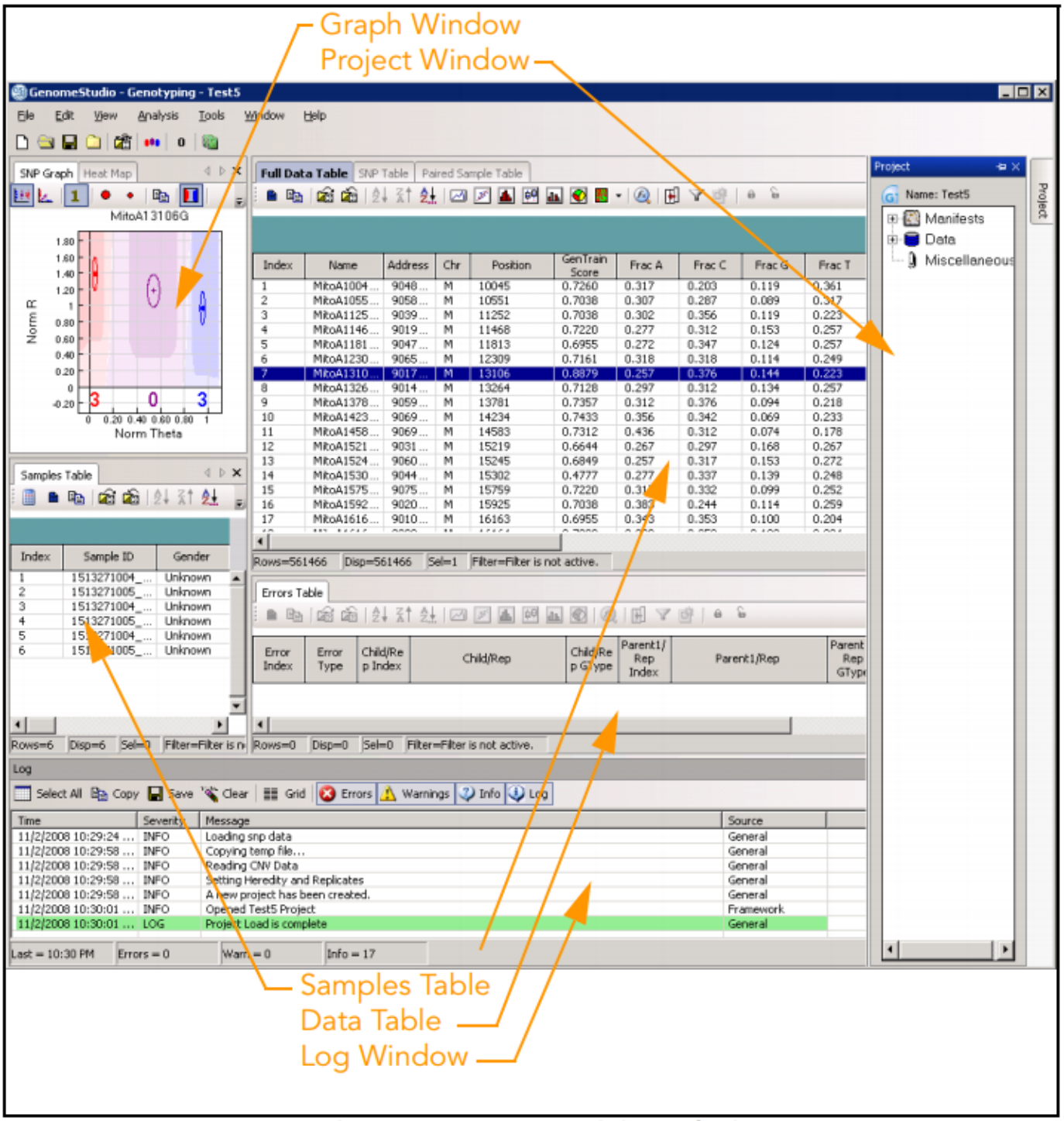

Figure 4: GenomeStudio Software Interface—GenomeStudio Software features a user-friendly graphical interface for the control and display of genotyping and clustering results.

Table 1: GenomeStudio Software Display Options

Global Visualization
Illumina Genome Viewer (IGV)
Illumina Chromosome Browser (ICB)
Graphs
Heat Maps
Scatter Plots
Histograms
Line Graphs
Box Plots
Frequency Plots
Pie Charts
Tables
Samples Table
SNP Table
Full Data Table

Integrated Controls Dashboard

Illumina array-based assays contain internal sample-dependent and sample-independent controls so researchers can have confidence that they are producing high-quality data. Researchers can use the Controls Dashboard, integrated into GenomeStudio Software, to monitor the performance of all controls easily.

Table 2: Minimum GenomeStudio Software System Requirements

Parameter	Microarray Data Analysis
CPU Speed	2.0 GHz or greater
Processor	64-bit
Memory	8 Gb or more
Hard Drive	250 Gb or larger
Operating System	Windows 7 or higher

Summary

GenomeStudio Software provides an integrated platform for data analysis of all Illumina array-based genotyping assays. The graphical display of results generated from data analysis supports high-level and in-depth views of whole-genome variation. GenomeStudio Software enables analysis of genotyping data from both diploid and polyploid organisms to drive biological insights more efficiently.

Ordering Information

GenomeStudio Software is available at no charge and can be downloaded from the GenomeStudio Support Page on the Illumina website.

Learn More

To learn more about GenomeStudio 2.0 Software and third-party analysis tools, visit www.illumina.com/genomestudio.

To learn more about analysis workflows for microarray data, read the Microarray Data Analysis Workflows technical note.

Illumina, Inc. • 1.800.809.4566 toll-free (US) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com
For Research Use Only. Not for use in diagnostic procedures.

illumina®