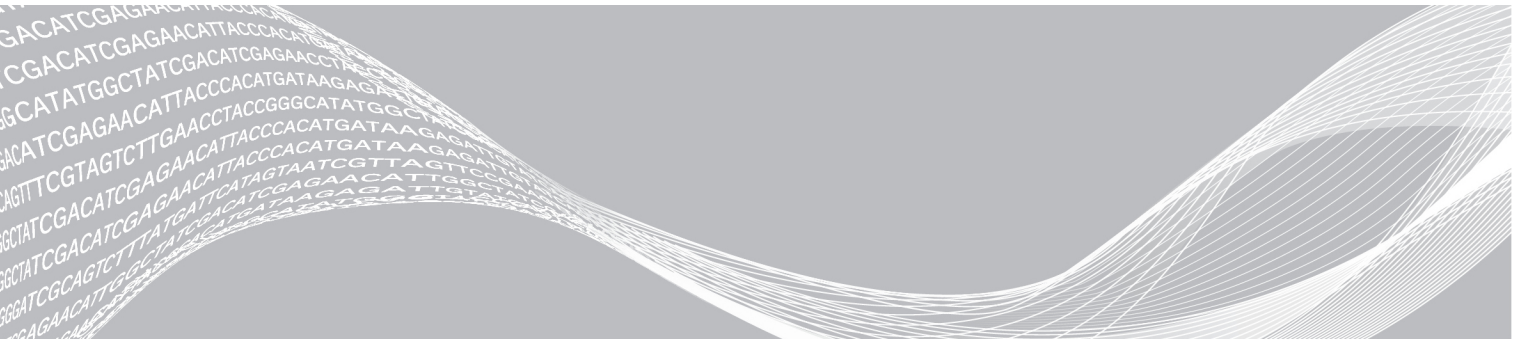
illumina®

Beeline Software 2.0

User Guide



Document # 1000000022181 v02

February 2017

ILLUMINA PROPRIETARY

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

Introduction

The Beeline Software 2.0 provides high-throughput genotyping analysis and lets you quickly import Genotype Call files (*.gtc), filter poorly performing loci and samples, and generate several types of reports. Although Beeline does not support raw intensity (*.idat) files, you can convert *.idat files to *.gtc files using the following software:

- AutoConvert via the Beeline Software.
- AutoCall via the Illumina Laboratory Information Management System (LIMS).
- iScan Control Software (ICS) via iScan System.

Before installing Beeline, you must have the Internet and a MyIllumina account. If you do not have access to the Internet, see *Technical Assistance* on page 14.

Beeline supports the following projects:

- **Sample + SNP Analysis**—Lets you filter your samples and loci data, and create a project in GenomeStudio for modifying cluster files before you generate reports or perform downstream analysis.
- **Sample Validation**—Lets you remove samples from your data before you generate reports or perform downstream analysis. Beeline does not import loci statistics for this project. For best results, use this project in a production environment.

The AutoConvert software installation is included in the Beeline software installation process.

Features

- **Time-Saving Analysis**—The Beeline Software can automatically calculate sample statistics, allele calling, and cluster file generation.
- **Reduced Data Size**—You can select and deselect samples and loci for targeted downstream analysis.
- **Accessible Reporting**—You can generate several types of reports.
- **Seamless Integration**—The software works directly with Infinium data output and creates projects for direct import into the GenomeStudio Software.
- **Polyploid Support**—The software supports polyploid cluster data sets.

Install the Beeline Software

1. Navigate to the Beeline support page, and then select **Downloads**.
2. Select **Beeline Software 2.0**, and then select **Beeline Software 2.0 Installer.**
3. Accept the software license terms and conditions for Beeline, and then click **Install**.
4. Click **Next**.
5. Accept the software license agreement for AutoConvert, and then click **Next**.
6. [Optional] Click **Change** to choose another folder to install AutoConvert v2.0.
   If you change the location, you must change the settings in Beeline. See *Update AutoConvert File Path* on page 11.
7. Click **Next** to accept the folder, and then click **Finish**.

Getting Started

1. From your desktop, click **Start** | **All Programs** | **Illumina** | **Beeline 2.0**.

2 In the Platform Login Form dialog box, enter your email and password using your MyIllumina account.
The first time you log in to Beeline, you are prompted with this authentication dialog box.

3 Click **Log In**.
The Beeline welcome window opens.

The welcome window includes the following features:

- **Open Recent**—Open a recent project.
- **Create New**—Create a project.
- **Training Material**—View the Beeline training video.
- **Help**—View support documentation.
- **Close**—View the main window.
- **Show at start up**—Open the welcome window at startup. To open the software to the main window, deselect the checkbox.

Create a Project From a Directory

1 From your desktop, double-click the Beeline icon.
The application opens.

2 In the Create New pane, select **Sample + SNP Analysis.**

3 Click the **Sample Directory** radio button.

4 At the **Sample Directory Location** field, click **Browse** and navigate to the folder that contains your Genotype Call (*.gtc) files.

5 At the **Manifest** field, click **Browse** and navigate to your SNP manifest (*.bpm) file.

6 At the **Master Cluster File** field, click **Browse** and navigate to your cluster (*.egt) or polyploid cluster (*.egtp) file.

7 [Optional] At the **GenomeStudio Working Folder** field, click **Browse** and navigate to where you want to save your project.
You can specify this folder when you finish creating the project.

8 [Optional] At the **GenCall Score Threshold** field, enter the call score for this project.
We recommend the 0.15 GenCall cutoff score for Infinium products.

9 Click **Finish**.
Your project data opens in the main window.

Create a Project From a Sample Sheet

1 From your desktop, double-click the Beeline icon.
The application opens.

2 In the Create New pane, select **Sample + SNP Analysis.**

3 Click the **Sample Sheet** radio button.

4 At the **Sample Sheet Location** field, click **Browse** and select your sample sheet.

5 [Optional] At the **Override Path from Sample Sheet** field, click **Browse** and navigate to the to the folder that contains your Genotype Call (*.gtc) files.

6 At the **Manifest** field, click **Browse** and navigate to your SNP manifest (*.bpm) file.

7 At the **Master Cluster File** field, click **Browse** and navigate to your cluster (*.egt) or polyploid cluster (*.egtp) file.

8 [Optional] At the **GenomeStudio Working Folder** field, click **Browse** and navigate to where you want to save your project.
You can specify this folder when you finish creating the project.

9 [Optional] At the **GenCall Score Threshold** field, enter the call score for this project.
We recommend the 0.15 GenCall cutoff score for Infinium products.

10 Click **Finish**.
Your project data opens in the main window.

Controls and Statistics

Table Controls Toolbar

Use the controls on the toolbar to change the data that appears in the samples or loci statistics table. The controls do not modify the data in your project.

Icon	Tool Name	Description
	Calculate	Updates all statistics in the table. The tool is active when statistics are out of sync.
	Select all	Highlights all the rows in the data table.
	Copy selected rows to the clipboard	Copies selected rows or columns to the clipboard.
	Export displayed data to a file	Exports selected rows or columns to a file.
	Import columns into the table	Imports sample or loci data from a file you specify.
	Sort column (Ascending)	Sorts the data table in ascending order by the selected column.
	Sort column (Descending)	Sorts the data table in descending order by the selected column.
	Sort by multiple columns	Sorts the data table by more than one column.
	Line plot	Displays a line plot of the sample or loci statistics.
	Scatter plot	Displays a scatter plot of the sample or loci statistics.
	Histogram plot	Displays a histogram of the sample or loci statistics.
	Box plot	Displays a box plot of the sample or loci statistics.
	Frequency plot	Displays a frequency plot of the sample or loci statistics.
	Pie chart	Displays a pie chart of the sample or loci statistics.
	Calculate new column	Calculates data for creating a new column.

Icon	Tool Name	Description
	Column chooser	Adds or removes columns from the data tables and rearrange the column order.
	Filter rows	Hides or shows rows of data based on the criteria you specify. This feature does not remove samples or loci from your data set. It is for viewing information in the data table only. If you want to make the data set smaller and easier to work with, use the threshold or manual filtering features described in Filtering Data.
	Clear filter	Removes filter settings. This feature does not change threshold criteria set by following the instructions in Filtering Data. It affects the way information is displayed in the data table only.
	Lock selected columns	Locks the selected column so it cannot scroll with other columns.
	Unlock selected columns	Unlocks the selected column so that it can scroll with other columns.

Samples Table

The Samples table shows the statistics for each sample.

Table 1 Samples Table Columns

Column Heading	Description
Index	The sample index row.
Sample ID	The sample ID.
Sample Name	The sample name.
GTC File Name	The Genotype Call file name.
Passed Threshold	Indicates whether samples pass all filters. [1] — Pass [0] — No pass
Manual Override	Indicates whether the sample is manually included or excluded. [1] — Include [0] — Exclude
Used	The value is equivalent to the value in Passed Threshold column. If users make a manual override, then the value is equivalent to the value in the Manual Override column.
Call Rate	Number of calls / (Number of calls + Number of no calls)
Calls	The total number of calls.
No Calls	The total number of no calls.
LogRDev	The standard deviation of the LogR ratio values for the sample.
p10 GC	The 10th percentile of GenCall values.
p50 GC	The 50th percentile of GenCall values.
p05 Green	The 5th percentile of the raw intensity in the green channel.
p50 Green	The 50th percentile of the raw intensity in the green channel.
p95 Green	The 95th percentile of the raw intensity in the green channel.
p05 Red	The 5th percentile of the raw intensity in the red channel.
p50 Red	The 50th percentile of the raw intensity in the red channel.

Column Heading	Description
p95 Red	The 95th percentile of the raw intensity in the red channel.
Gender	The specified gender.
Gender Est.	The estimated gender.
Sample Group	Your sample group name.
Sample Plate	Your sample plate identifier.
Sample Well	The well within your sample plate.
Comment	Your-defined comment.
Cluster File	The cluster file used to create the Genotype Call (*.gtc) file.
Date	The imaging date.
Status	The clustering status of the sample.
Sentrix ID	The barcode number of the Universal Array Product to which this sample was hybridized.
Scanner Name	The name of the scanner that scanned the sample.

Loci Statistics

The Loci table shows the statistics for each locus.

Table 2 Loci Table Columns

Column Heading	Description
Index	The locus index row.
Name	The locus name.
Chr	The locus chromosome.
SNP	The locus variation.
Passed Threshold	Indicates whether loci pass all filters. [1] — Pass [0] — No pass
Manual Override	Indicates whether the locus is manually included or excluded. [1] — Include [0] — Exclude
Used	The value is equivalent to the value in Passed Threshold column. If users make a manual override, then the value is equivalent to the value in the Manual Override column.
# Calls	The total number of calls.
# No Calls	The total number of no calls.
Call Freq	Number of calls / (Number of calls + Number of no calls)
Minor Freq	The frequency of the minor allele.
Het Excess	If fAB is the heterozygote frequency observed at a locus, and p and q are the major and minor allele frequencies, then het excess calculation is (fAB - 2pq) / (2pq + epsilon).
ChiTest100	The Hardy-Weinberg p-value estimate calculated using genotype frequency. The value is calculated with 1 degree of freedom and is normalized to 100 individuals.
AA Freq	The frequency of homozygote allele A calls.
AB Freq	The frequency of heterozygote calls.
BB Freq	The frequency of homozygote allele B calls.
Cluster Step	The cluster separation score.

Column Heading	Description
AB R Mean	The normalized r-values mean for AB genotypes.
AB T Mean	The normalized theta angles mean for AB genotypes.
Plus/Minus Strand	Indicates whether the assay is designed with plus or minus strand.
Comment	The user-defined comment.

Filtering Data

Filter data to eliminate poorly performing samples or loci to produce cleaner genotyping data in your reports. Beeline can automatically filter samples or loci based on thresholds you define.

Set Thresholds to Filter Samples or Loci Automatically

1 In the main window, select the **Samples** tab or **Loci** tab, and then click **Set Thresholds**.
The Sample or Loci Thresholding dialog box opens.

2 In the Y Axis pane, select the column for filtering your data.
In the line plot, the red line represents the samples or loci threshold value and the blue line represents the samples or loci.

3 To change the threshold value, drag the red line up or down on the graph or enter the value in the **Samples Threshold** or **Loci Threshold** field.
Beeline includes samples or loci above the threshold value (red line) and excludes samples or loci below the threshold value.

4 Click **Add/Update** to add the parameter.

5 Repeat steps 2 through 4 to add additional columns to filter the samples or loci in your data set.

6 When you are finished, click **OK**.
Beeline filters the samples and updates the Samples or Loci table. Beeline includes only the samples with the value of (**1**) in the Passed Threshold and Used columns. If the value is (**0**), then the samples are not included in a report or GenomeStudio project.

Include or Exclude Samples or Loci

1 In the main window, select the **Samples** or **Loci** tab, and then select the rows you want to include or exclude from your report or project.
Press the **Ctrl** key to select multiple rows.

2 Right-click on the selected rows, and then select **Manual override out** to exclude or **Manual override in** to include the samples or loci.
If you set a value for the Manual Override column, then the value of the Used column is equal to the value of the Manual Override column. Otherwise, the value of the Used column is equal to the value of the Passed Threshold column.

3 Repeat steps 1 and 2 to select all the samples or loci you want to include or exclude from your report or project.

Update SNP Statistics or Call Samples

After filtering your samples or loci, some columns in the Sample or Loci table are highlighted in red to indicate that values are out of sync. Update the samples or loci statistics to resync the data.

1 In the main window, complete 1 of the following options:

- To update SNP statistics, select the **Loci** tab, and then click **Calculate SNP Stats**.
- To call samples, select the **Sample** tab, and then click **Call Samples**.

2 When the samples and loci statistics are in sync, you can do the following:
- Generate a report or chart.
- Create a GenomeStudio project.

Modifying the Cluster File

If you want to modify the cluster files, you can create a project containing your filtered data in GenomeStudio, modify the cluster files, and then merge the changes back into your Beeline project.

Create a Project for GenomeStudio

Create a GenomeStudio project from your filtered samples and loci data.

1 In the main window, select 1 of the following options:
- **Create Project from Used**—To create a project from samples and loci with a value of 1 in the Used column.
- **Create Project from Selected**—To create a project that contains specific samples and loci statistics. Before you select this option, highlight the samples and loci you want to include in the project.

The Save Project dialog box opens.

2 At the **Projects Repository** field, browse to the location you want to save the project.

3 At the **Project Name** field, enter the project name, and then click **OK**.
The project opens in GenomeStudio.

Merge Cluster File Changes in GenomeStudio

1 In GenomeStudio, click **File | Export Cluster Positions | For All SNPs**.
GenomeStudio opens the folder with the subClusterFile.egt file for this project.

2 Select **subClusterFile.egt** for the file name, and then click **Save**.

3 Click **Yes** to overwrite the cluster file.

4 Return to the Beeline main window.

5 In Beeline, click **Calculate SNP Stats** and **Call Samples** before you filter samples or generate reports.

Export Cluster File and Merge the Changes in Beeline

After modifying the cluster file in GenomeStudio, you can export the cluster file and merge the changes in Beeline.

1 In GenomeStudio, click **File | Export Cluster Positions | For All SNPs**.
GenomeStudio opens the folder with the subClusterFile.egt file for this project.

2 In the **File Name** field, enter a new name for the cluster file, and then click **Save**.

3 Return to the Beeline main window.

4 In Beeline, click **Project | Merge Cluster Subset**.

5 Navigate to the modified cluster file, select the file, and click **Open**.

6 Click **Calculate SNP Stats** and **Call Samples** before you filter samples or generate reports.

Generating Reports

You can generate the following reports in Beeline:

- **DNA Report**—Provides a snapshot of the samples in the Samples statistics table.
- **Summary Report**—Provides a high-level overview of the project, including number of samples and loci, manifest and cluster file path, and the report processing date.
- **Locus Summary Report**—Provides a snapshot of the loci in the Loci statistics table.
- **Final Report**—Provides allele calls of each locus for all samples.

You can specify the number of samples to include per file.

- **All** samples in one file.
- **One** sample per file.
- **Custom** number of samples per file.

Generate a Final Report

1. In the main window, click **Reports** from the toolbar.
   The Beeline Reports dialog box opens.
2. Click the **Final Report** checkbox.
3. At the **Samples/file** field, enter the number of samples per file.
4. [Optional] You can customize the data you want to include or exclude from the final report.
   For example, highlight the data you want to exclude in the Included Columns pane, and then click  . The Excluded Columns pane lists the data to exclude. You can save this as a **Favorite Format**, which also stores the **Samples/file** and the **Delimiter**.
5. At the **Output path** field, browse to the location you want to save the report.
6. At the **Name** field, enter the report name, and then click **OK**.
   The software generates the final report.

Generating Charts

You can create custom charts or use the available chart templates in Beeline.

In the main window, click **Charts** in the toolbar. The templates are located in the Available Charts pane and include the following:

- Call Rate vs. Index
- P10 GC vs. Index
- Controls: Staining Red or Green
- Controls: Extension Red or Green
- Controls: Target Removal Red or Green
- Controls: Hybridization Red or Green
- Controls: Stringency Red or Green
- Controls: Non-Specific Binding Red or Green
- Controls: No-Polymorphic Red or Green

- Controls: Restoration Red or Green

Generate a Custom Chart

1 In the main window, click **Charts** in the toolbar.

2 In the Reports pane, click to create a custom report.

3 Double-click **New Report** to rename the report.

4 In the Available Charts pane, click to create a chart.
The Create New Chart dialog box opens.

5 At the **Table** field, select **Sample** or **Loci** to set the source data for the chart.

6 At the **Chart Type** field, select the type of chart you want to create.

7 Click **OK**.

8 In the Available Charts pane, double-click the chart template and rename the chart.

9 On the right pane, customize your chart template by selecting the statistics:
- Left axis, right axis, and X axis (line plot only)
- Data and labels (scatter plot only)

10 Highlight the custom chart template, and click in the **Charts in Selected Report** pane.
The Charts in Selected Report pane shows the custom chart template. Continue to add chart templates to your custom report.

11 When finished, click **Export** in the Reports pane.

12 At the **File name** field, enter the name for the custom report and click **Save**.
Beeline generates the report in PDF format.

Convert IDAT Files to GTC Files

Although Beeline does not support raw intensity (*.idat) files, you can convert *.idat files to *.gtc files using AutoConvert via Beeline .

1 In the main window, click **Covert Idat to Gtc** on the toolbar.
The Convert to GTC dialog box opens.

2 At the **IDat Folder** field, click **Browse** and navigate to the folder that contains your IDAT files.

3 Select 1 of the following options:
- **Existing IDAT folder**—Save the GTC files in the same folder that contains your IDAT files.
- **Alternate folder**—Save the GTC files in an alternate folder. (Recommended)

4 At the **Manifest** field, click **Browse** and navigate to your SNP manifest (*.bpm) file.

5 At the **Cluster File** field, click **Browse** and navigate to your cluster (*.egt) or polyploid cluster (*.egtp) file.

6 Click **Next**.
Beeline coverts your IDAT files to GTC files and saves them in a folder.

Update AutoConvert File Path

During the Beeline software installation, if you selected another location to install AutoConvert, you must update the AutoConvert file path for Beeline to locate the AutoConvert Software.

1 In the main window, click the **Tools** menu, and then select **Beeline Settings**.
The Beeline Settings dialog box opens.

2 At the **AutoConvert Path** field, click **Browse** and navigate to the folder you selected at installation.

3 At the **GenomeStudio Path** field, keep the default file path.

4 Click **OK**.
Beeline updates the AutoConvert file path.

System Requirements

Table 3

System	Requirements
Processor	64-bit
Memory Size	8 GB or more
Hard Drive	100 GB or larger
Video Display	1280 x 1024
Operating System	Windows 7 or later
Programming Framework	Microsoft .NET Framework 3.5
Network Connection	1 GbE or faster
Other Required Software	iScan Control Software 3.2 or later GenomeStudio 2.0

Resources

- Beeline
- GenomeStudio
- Array Software Support
- *GenomeStudio Framework User Guide* (Document # 11318815)
- *GenomeStudio Genotyping Module User Guide* (Document # 11319113)
- *iScan System User Guide* (Document #11313539)
- *Improved Cluster Generation with GenTrain2* (Technical Note)

Revision History

Document	Date	Description of Change
Document # 1000000022181 v02	February 2017	• The v01 revision mistakenly omitted the Revision History section. v02 corrects the omission.
Document # 1000000022181 v01	December 2016	• Added new content on creating projects from a sample sheet • Minor changes due to updates in user interface
Document # 1000000022181 v00	August 2016	• Supported Beeline v2.0 • Updated format to current style standards • Added new content on generating charts and reports, creating projects, sample data reference
Document # 15016340 Rev. A	July 2010	• Initial release

Technical Assistance

For technical assistance, contact Illumina Technical Support.

Website: www.illumina.com
Email: techsupport@illumina.com

Illumina Customer Support Telephone Numbers

North America 1.800.809.4566	**Germany** 0800.180.8994	**Singapore** 1.800.579.2745
Australia 1.800.775.688	**Hong Kong** 800960230	**Spain** 900.812168
Austria 0800.296575	**Ireland** 1.800.812949	**Sweden** 020790181
Belgium 0800.81102	**Italy** 800.874909	**Switzerland** 0800.563118
China 400.635.9898	**Japan** 0800.111.5011	**Taiwan** 00806651752
Denmark 80882346	**Netherlands** 0800.0223859	**United Kingdom** 0800.917.0041
Finland 0800.918363	**New Zealand** 0800.451.650	**Other countries** +44.1799.534000
France 0800.911850	**Norway** 800.16836	

Safety data sheets (SDSs)—Available on the Illumina website at support.illumina.com/sds.html.

Product documentation—Available for download in PDF from the Illumina website. Go to support.illumina.com, select a product, then select **Documentation & Literature**.

Document # 1000000022181 v02

Illumina 5200 Illumina Way, San Diego, California 92122 US
1.800.809.4566 toll-free (US) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com
Illumina Cambridge Limited, Chesterford Research Park, Little Chesterford, Saffron Walden, CB10 1XL, UNITED KINGDOM
Australian Sponsor: Illumina Australia, 1 International Court, Scoresby, Victoria, 3179 Australia

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

illumina®