

GenomeStudio® Software 2009.2 Release Notes

1.	GenomeStudio Software 2009.2 Framework	1
2.	Illumina Genome Viewer v1.5.....	2
3.	Genotyping Module v1.5.....	4
4.	Gene Expression Module v1.5.....	6
5.	Methylation Module v1.5	6
6.	Sequencing Modules v1.5	7

1. GenomeStudio Software 2009.2 Framework

- 1.1. The ability to import complete dbSNP information (based on Human dbSNP build 129) into the Alleles Table and the Illumina Genome Viewer (IGV) has been added.
 - 1.1.1. You must install *GenomeStudio_dbSNP129.exe* to make the full dbSNP available. The dbSNP SNP Name column in the Allele Table is populated, and dbSNP information is displayed in the IGV.
 - 1.1.2. If the *refseq_path* in the *GenomeStudio.exe.config* file has been changed and you want to populate dbSNP information, copy the dbSNP folder manually to the new location of the human genome folder.
 - 1.1.3. This full dbSNP replaces the curated dbSNP in the previous release.
- 1.2. The ability to hyperlink the content of a table column to a website has been added.
 - 1.2.1. The *LinkToWebsite.txt* file in the GenomeStudio folder specifies the columns in a table to be linked, and the linked websites.
 - 1.2.2. You can define hyperlink content by editing *LinkToWebsite.txt*. For example:

Website Name	GenomeStudio Table	Website URL	Column Name
NextBio	Gene Table	http://www.nextbio.com/b/search/disc/{0}#cat=Compounds	Genes

Note that the website name, the GenomeStudio table, the website address (URL), and the column in the table are tab-delimited. Use {0} to replace the linked content.

- 1.2.3. By default, the Genes column in the Exon Table, Gene Table, and Alleles Table is linked to the NextBio website.
- 1.2.4. *LinkToWebsite.txt* is not overwritten if GenomeStudio is reinstalled.
- 1.2.5. You can add hyperlinks to multiple columns in a table.
- 1.2.6. To open the linked website in a regular web browser, select **Ctrl+N**.
- 1.2.7. To disable hyperlink features in GenomeStudio table columns, remove the content of the *LinkToWebsite.txt* file.

2. Illumina Genome Viewer v1.5

- 2.1. Human Genome Build 36.1 displayed in the IGV has been updated to the UCSC version from May 2009.
- 2.2. The IGV also supports the latest human genome (Build 37).
 - 2.2.1. You can download the latest human genome from the UCSC website, as described in Chapter 5 of the GenomeStudio Framework User Guide.
- 2.3. Performance for Stacked Alignment plots is significantly improved since the last release (GenomeStudio Software 2009.1).
 - 2.3.1. Plotting the entire genome is much faster.
 - 2.3.2. Plotting the entire chromosome in the Illumina Chromosome Browser (ICB) is significantly faster.
 - 2.3.3. Plots for Read Type "All" are faster for DNA Seq projects.
 - 2.3.4. Data plots performance remains the same.
- 2.4. Whole Genome View and the Slide Show View have been restored to the Sequencing modules. This allows you to view sequencing data across all chromosomes in a sample (Whole Genome View) or 1-4 chromosomes at a time (Slide Show View).
- 2.5. The ability to view dbSNP (build 129) information under Gene Annotations in the ICB has been added.
 - 2.5.1. To view dbSNP information, install *GenomeStudio_dbSNP129.exe* as described in section 1.1.1 above.
 - 2.5.2. You can toggle on/off the display of dbSNP from the Gene Annotations dropdown menu in the IGV. This feature is turned off by default.
 - 2.5.3. dbSNP information is visible only at read or base level. You cannot view dbSNP information when zooming out to histogram view.
 - 2.5.4. The Reference SNP ID (rs#), Reference SNP alleles, and strand (+ for forward, - for reverse) are displayed.
- 2.6. Quick search by dbSNP and gene name has been added to the ICB.
 - 2.6.1. To start the search, type the query (such as rs# or GeneSymbol) in the ICB Chr Position field.
 - 2.6.2. If no hit is found in the current chromosome, other chromosomes will be searched until the first hit is found.
 - 2.6.3. A query can be a partial match and/or case insensitive.
 - 2.6.4. The current view is centered on the found locus.
- 2.7. To toggle on/off the Gene table or information control tab displayed in the ICB, select **View | Gene Annotations** or **View | Gene Information Tab**.
- 2.8. You can change the data plot displayed in the IGV by selecting the value from the dropdown menu in the Data Properties panel.
 - 2.8.1. By enabling Immediate Mode in the IGV, you can scroll through all of the data plots by selecting a category, then using the arrow up/down keys.

- 2.9. You can now change plot height by dragging the bottom edge of the plot.
 - 2.9.1. Mouse over the bottom edge of the plot. When the arrow appears, drag the line to a desirable position.
 - 2.9.2. You cannot adjust plot height if *Select Min and Max* reads is the active selection.
- 2.10. The *Mark* function in the IGV context menu has been renamed *Mark Data in Table*.
 - 2.10.1. The Sequences table is brought to the front with the marked read in the current view.
 - 2.10.2. There is a colored line above the marked read in the IGV.
 - 2.10.3. When marking a paired end read from the IGV, both reads are marked in the IGV and in the Sequences table.
 - 2.10.4. Marking a spliced read also marks the other splice site read in the Sequences table.
- 2.11. The ability to view gapped alignment has been added to the IGV.
 - 2.11.1. Display of gapped alignment works with data from CASAVA v1.6.
 - 2.11.2. The insertion site is represented by the two bases surrounding the insertion site appearing in **bold text**.
 - 2.11.2.1. You can use the Jump to Sequences table function to view the inserted base(s) in the read.
 - 2.11.3. Each deleted base is represented by a hyphen (-).
- 2.12. A tool tip for the read is now available in the IGV for all sequencing modules
 - 2.12.1. The tool tip contains the Read 1 chromosome and its position (and Read 2 chromosome and its position for a PE read), Read type, and Seq Table Row for Read 1 (and Read 2 for a PE read).
- 2.13. The reverse strand color for SE read has been changed to **orange** for better visualization of bases that do not match the reference sequence.
- 2.14. In previous releases of GenomeStudio software, when scrolling along a chromosome in the IGV with multiple plots, the IGV reverted back to the bottom plot. This issue has been resolved in the current release.
- 2.15. Copying an image of the stacked alignment now includes a scale for the base position.
- 2.16. Four Zoom buttons have been added to the ICB:
 - 2.16.1. Zoom to Base Level
 - 2.16.1.1. Replaces menu item Zoom To-->Base Level in the last release
 - 2.16.1.2. Bases in the center of the current view are displayed.
 - 2.16.1.3. Only enabled if current view is not at the base level
 - 2.16.2. Zoom to Gene Level
 - 2.16.2.1. Only enabled when viewing data at the base level
 - 2.16.2.2. Centers the view on the first gene in the current view. Centers on the current view if no genes are found in a 20000 genomic coordinate window.

- 2.16.3. Zoom to Chromosome Level
 - 2.16.3.1. Replaces menu item Zoom to-->Chromosome level in the last release
- 2.16.4. Zoom to Genome Level
 - 2.16.4.1. Opens the IGV to show the genome level view.
- 2.17. Base level view can now be displayed up to 10 pixels, compared with only 8 pixels in previous releases.
- 2.18. The List view is replaced by a tree view to toggle on/off sequence plot visibility.
 - 2.18.1. Each checkbox toggles the visibility of the plots on/off.
 - 2.18.2. An advanced PE read type section has been added to show/hide the complexity of the plots.
- 2.19. **Known Issue:** In sequencing projects in the IGV, when you change the active genome by selecting **Edit | Preferences | Genome**, the display window appears empty. To display the ICB, select **View | Chromosome Browser**.
- 2.20. **Known Issue:** Data plots created using the Partner Offset value in the Sequences table are empty. This issue should be resolved in the next release.
- 2.21. **Known Issue:** When you jump to IGV from the reads in the last chromosome in the Alleles table, you may see an error message if the dbSNP display is on and you are navigating to the end of chromosome Y, or from another chromosome to chromosome Y. This issue will be resolved in next release. The workaround is to hide dbSNP information. You can turn dnSNP back on when you navigate away from the tail of chromosome Y.
- 2.22. **Known Issue:** You must click on a GenomeStudio data track to activate its color gradient.

3. Genotyping Module v1.5

- 3.1. The term "Lot" or "Version", used in previous releases to indicate different sets of manifests, has been renamed "Subset". Following is a list of locations where this change has been made:
 - 3.1.1. **Sample Table:** The hidden column "Version" has been renamed "Subset".
 - 3.1.2. **Full Data Table and SNP Table:** "Cluster all SNPs in the Current Lot" has been changed to "Cluster all SNPs in the Current Subset"; "Cluster Selected SNPs in the Current Lot" has been changed to "Cluster Selected SNPs in the Current Subset".
 - 3.1.3. **SNP Graph:** The tooltip for "Cluster This Lot" has been changed to "Cluster This Subset".
 - 3.1.4. **Export Cluster Positions:** "Is this cluster for Lot A, or Lot B..." has been changed to "Is this cluster for Subset A, or Subset B..."
 - 3.1.5. **Sample Sheet:** The "Version" column has been renamed "Subset".
- 3.2. GenomeStudio software is now compatible with data and projects for the GoldenGate Indexing™ (GGI) assay, created using the Illumina LIMS workflow (version 4.1 and greater).
 - 3.2.1. A GGI project can be created **only** through LIMS.

- 3.2.2. Use a super manifest file (*.xtn) to create a GGI project.
 - 3.2.2.1. The super manifest file contains information from each subset manifest.
- 3.2.3. The GGI project has a super cluster file (*.xtc) associated with it.
 - 3.2.3.1. The super cluster file consolidates regular cluster files (*.egt) for every subset manifest.
 - 3.2.3.2. The super cluster file can be generated only in GenomeStudio v2009.2 and greater.
 - 3.2.3.3. To export the super cluster file to LIMS, select **File | Export Cluster Position to LIMS**.
 - 3.2.3.4. The default name of the super cluster file is the name of the super manifest. You can rename the super cluster file, or overwrite the existing cluster file in LIMS.
 - 3.2.3.5. To link the cluster file with the product, you must activate the super cluster file through the LIMS Project Manager. Please see Illumina LIMS documentation for details.
- 3.2.4. A new data table called the Full SNP Table is generated for GGI projects.
 - 3.2.4.1. To display the Full SNP Table, which is hidden by default, select **Window | Full SNP Table**.
 - 3.2.4.2. The Full SNP Table contains SNP information from all subsets.
 - 3.2.4.3. To zero the selected SNP in the current subset, right-click and select **Zero Selected SNP** from the Full SNP Table context menu.
 - 3.2.4.4. To zero a SNP across all subsets, select **Zero Selected SNPs** in the Full Data Table or SNP Table.
 - 3.2.4.5. Row selection in the Full SNP Table is synchronized with other data tables and the SNP graph.
 - 3.2.4.6. Selecting a SNP in the SNP table selects the SNPs from all subsets in the Full SNP table.
 - 3.2.4.7. When you select a SNP in the Full SNP table, the Error Table displays error(s) for this SNP from all subsets, with the error(s) for this SNP in the current subset highlighted.
 - 3.2.4.8. **Known Issue:** For SNPs that are not part of the first subset, selections in the Sample Graph and Embedded Heat Map are not synchronized with row selections in the Full SNP Table.
- 3.2.5. The *Send Requeue to LIMS* option in the Samples Table is not available for GGI projects.
- 3.2.6. Two types of controls are available in the controls dashboard for GGI projects: Second Hybridization controls and user-defined controls.
 - 3.2.6.1. A user-defined control is specified in manifest file (*.xtn), and contains the following fields: **code_name**, **dashboard_message**, **description**, **control_name**, **control_Desc**, and **control_color**.

- 3.3. A new clustering algorithm (GenTrain 2.0) has been added to this GenomeStudio software release. You can choose to use the original clustering algorithm (GenTrain 1.0), or the new one (GenTrain 2.0).
 - 3.3.1. GenTrain 1.0, the clustering algorithm used in previous GenomeStudio and BeadStudio software releases, is the default clustering algorithm.
 - 3.3.2. GenTrain 2.0 can be applied only to Infinium® assay data. A Technical Note describing this algorithm will be available soon.
 - 3.3.3. To select the clustering algorithm you want to use, go to **Tools | Options | Module | Genotyping**.
- 3.4. In the Samples Table, the column header for the manifest information has been changed from the manifest name to Array Info.
- 3.5. The Export Cluster Position to LIMS function for the Infinium and GoldenGate assays is enabled in this release.
- 3.6. Illumina LIMS now is selected by default in the LIMS login window in GenomeStudio.
- 3.7. An issue found when the number of SNPs in a normalization bin was less than 400 has been resolved. As a result, for SNPs in small normalization bins, the normalized intensity value in GenomeStudio Software 2009.2 may be different from that in previous versions of GenomeStudio software.
- 3.8. The name of the cluster file used in the project has been added in GenomeStudio log.
 - 3.8.1. If you use a GenomeStudio clustering algorithm (GenTrain 1.0 or GenTrain 2.0) to cluster, the name of algorithm is added in the log.
- 3.9. In GenomeStudio Software 2009.1, if samples were not available for all manifests, the software retrieved incorrect normalization IDs from the lookup table, and incorrect normalized intensity values were displayed for those samples. This issue has been resolved in GenomeStudio Software 2009.2.

4. Gene Expression Module v1.5

- 4.1. An issue with the BEAD_STDEV value changing if *Export to GeneSpring* was selected multiple times has been addressed in this release.
- 4.2. You can now view BeadChip images from the Samples Table in *.jpg and *.tiff formats.

5. Methylation Module v1.5

- 5.1. Background and Average normalization for GoldenGate assay have been restored in this release.
- 5.2. A clustering issue with heat map data has been fixed in this release.
- 5.3. An issue with the exported control file has been resolved.
- 5.4. An issue with the Save Project Copy As function has been resolved in this release. The assay type is now saved along with the project. When you open a saved project, GenomeStudio no longer asks you to select an assay type.

6. Sequencing Modules v1.5

- 6.1. Any chromosome name (such as c1, c1.fa, or E. coli) from CASAVA, except NMNM, can be loaded in GenomeStudio software.
- 6.2. An Indels Table has been added to the DNA Sequencing Module.
 - 6.2.1. To display the Indels Table, select **Window | Indels Table**.
 - 6.2.2. The Indels Table displays the content of the *indels.txt* file created by IndelFinder in CASAVA.
 - 6.2.3. The Indels Table contains the sample name (Sample), chromosome (Chr), position (Pos), CIGAR format (CIGAR), and other columns imported from the *indel.txt* file.
- 6.3. The Sequences table includes the following new columns:
 - 6.3.1. Partner Chr - partner read chromosome
 - 6.3.2. Partner Strand - partner read strand
 - 6.3.3. Partner Offset - partner bases offset
 - 6.3.4. In the Sequences table, the new Insertion or Deletion column indicates insertion or deletion as detected following gapped alignment.
 - 6.3.4.1. The letter "Y" in the Insertion or Deletion column indicates that there is insertion or deletion in the read.
 - 6.3.4.2. A deleted base is represented by a hyphen (-) in the read.
 - 6.3.4.3. Bases at the end of a read with a deletion event are pushed out beyond the maximum read length of the run. These are designated **overhanging bases** in GenomeStudio.
 - 6.3.4.4. The number of overhanging bases visible in the Sequences table can be configured as a percentage of maximum read length through the *GenomeStudio.exe.config* file.

For example, if run's maximum read length is 100 bases, and a certain read has 13 deleted bases, 13 bases are overhanging. You can configure the software to show only 5% of the maximum read length as overhanging. In this case, the Sequences table would show $100 + 100 * .05 = 105$ bases. The Remaining 8 bases would be curtailed.
 - 6.3.4.5. In the *GenomeStudio.exe* config file, the default percentage of allowable overhanging bases is set to 10. You can change this value prior to project creation.


```
<add key = "perct_increase_read_length" value = "0.1" />
```
 - 6.3.4.6. You can optimize GenomeStudio project creation performance by adjusting the *perct_increase_read_length* value. For maximum speed, set the percentage value to 0.
 - 6.3.4.7. The inserted base is represented by â, ã, ĝ, ¤ for low quality base a, c, g, t, and by Â, Ã, Ĝ, ¤ for high quality base A, C, G, T.
- 6.4. In the IGV, the Jump to Sequences table and Mark Data in Table functions now work correctly with the dynamic filter in the Sequences table.

- 6.5. The Jump to PE Mate function in the Sequences table now works with the dynamic filter.
 - 6.5.1. Jump to PE Mate doesn't work for Chimera reads with the dynamic filter.
- 6.6. **Known Issue:** The Jump to PE Mate function works only for reads in the first sample. However, if you use the dynamic filter in the Sequences table to choose a sample, you can use the Jump to PE Mate function with no issues.
- 6.7. The Mark function in Sequences table has been improved for better synchronization with the IGV.
 - 6.7.1. Marking a read in the IGV marks the read with the same color in the Sequences table.
 - 6.7.2. Marking a read in the Sequences table marks the same read in the IGV.
- 6.8. **Known Issue:** When you mark a read in the Sequences table while the dynamic filter is on, the mark cannot be saved, and will disappear when you change the dynamic filter.
- 6.9. The **dbSNP SNP Name** column in the Alleles table now uses the full dbSNP build 129 instead of curated dbSNP information, which only contains dbSNP in the exon regions.
 - 6.9.1. You must install *GenomeStudio_dbSNP129.exe* in order for dbSNP information to be populated in the Alleles table.
- 6.10. A Genes column has been added to the Alleles table in the DNA Sequencing and RNA Sequencing modules.
 - 6.10.1. The information displayed in the Genes column is dynamic. If you change the Gene Labels in the IGV, the content in the Genes column changes to the corresponding gene label.
- 6.11. The splice junction reads from CASAVA v1.6 output in the RNA Sequencing Module now have the correct Chr ID in the Sequences Table, instead of using the splice-site as the Chr ID as in previous releases.
 - 6.11.1. CASAVA v1.6 no longer outputs splice reads in a separate file. As a result, the two shorter reads in a splice junction site are now mapped to the genome, instead of mapping them to a splice junction as in previous releases.
- 6.12. The default maximum number of rows that can be displayed in the Sequences table is 18 million. This allows optimal performance on a computer with 3.5 GB of RAM.
 - 6.12.1. In *GenomeStudio.exe.config*, you can configure the maximum number of rows displayed to between 1 million and 134 million.

Example: `<add key = "TableMaxRowCountForRefresh" value="18000000"/>`
 - 6.12.2. Illumina recommends setting the maximum allowable row count to 18 million rows or fewer.
 - 6.12.3. By adjusting the dynamic filter in the Sequences table, you can shift the 18 million row window (or a window of any other defined value) to a different region.
 - 6.12.4. All reads in a project (even beyond 134 million) can be displayed in the IGV.

For Research Use Only

© 2009 Illumina, Inc. All rights reserved.

Illumina, illuminaDx, Solexa, Making Sense Out of Life, Oligator, Sentrix, GoldenGate, GoldenGate Indexing, DASL, BeadArray, Array of Arrays, Infinium, BeadXpress, VeraCode, IntelliHyb, iSelect, CSPro, GenomeStudio, and Genetic Energy are registered trademarks or trademarks of Illumina, Inc. All other brands and names contained herein are the property of their respective owners.