Methylation (CH3) Analysis

Agilent Genomic Workbench 5.0

User Guide

Agilent Technologies

Notices

Edition

First Edition, April 2009

Printed in USA

Agilent Technologies, Inc.
3501 Stevens Creek Blvd.
Santa Clara, CA 95051

Software Revision

This guide is valid for 5.0 and later revisions of the Agilent Genomic Workbench software, until superseded.

Warranty

The material contained in this document is provided "as is," and is subject to being changed, without notice, in future editions. Further, to the maximum extent permitted by applicable law, Agilent disclaims all warranties, either express or implied, with regard to this manual and any information contained herein, including but not limited to the implied warranties of merchantability and fitness for a particular purpose. Agilent shall not be liable for errors or for incidental or consequential damages in connection with the furnishing, use, or performance of this document or of any information contained herein. Should Agilent and the user have a separate written agreement with warranty terms covering the material in this document that conflict with these terms, the warranty terms in the separate agreement shall control.

Technology Licenses

The hardware and/or software described in this document are furnished under a license and may be used or copied only in accordance with the terms of such license.

Restricted Rights Legend

U.S. Government Restricted Rights. Software and technical data rights granted to the federal government include only those rights customarily provided to end user customers. Agilent provides this customary commercial license in Software and technical data pursuant to FAR 12.211 (Technical Data) and 12.212 (Computer Software) and, for the Department of Defense, DFARS 252.227-7015 (Technical Data - Commercial Items) and DFARS 227.7202-3 (Rights in Commercial Computer Software or Computer Software Documentation).

Safety Notices

CAUTION

A **CAUTION** notice denotes a hazard. It calls attention to an operating procedure, practice, or the like that, if not correctly performed or adhered to, could result in damage to the product or loss of important data. Do not proceed beyond a **CAUTION** notice until the indicated conditions are fully understood and met.

WARNING

A WARNING notice denotes a hazard. It calls attention to an operating procedure, practice, or the like that, if not correctly performed or adhered to, could result in personal injury or death. Do not proceed beyond a WARNING notice until the indicated conditions are fully understood and met.

In This Guide...

This guide describes how to use the Methylation (CH3) Analysis application of Genomic Workbench to apply an algorithm that helps you assess the regions that are methylated or unmethylated.

1 Importing, Managing, and Exporting CH3 Data and Other Content

This chapter describes how to import, organize, manage, and export CH3 data and other content within the user interface of Agilent Genomic Workbench.

2 Visualizing CH3 Data and Other Content

This chapter shows you how to display log ratio data from imported feature extraction data files and analysis results, as well as gene list and track content, in the Genomic Viewer. It also gives you instructions on how to modify the display to visualize the data and content the way you prefer.

3 Setting Up Methylation (CH3) Analysis

This chapter gives instructions on how to set up the analysis functions for Methylation (CH3) experiments. These include the Preprocessing, Analysis and Reports tabs.

4 Methylation (CH3) Analysis Reference

This chapter describes the tabs, commands, shortcut menus and dialog boxes specific to DNA Analytics methylation (CH3) data analysis. Another section describes the format of the report the program generates.

5 Statistical Algorithms

This chapter provides implementation details of the algorithms used in the CH3 application of Genomic Workbench 5.0. The CH3 application algorithms facilitate the statistical analysis of methylated genomic regions. The first section presents the main methylation detection algorithm. This is followed by a second section which details the visualization options available for methylation analysis.

Contents

1 Importing, Managing, and Exporting CH3 Data and Other Content 11

Importing Files 13

To select a new location for data files 14
To import Agilent FE or GenePix/Axon data files 14
To import a UDF file 16
To import Agilent GEML design files 19
To import GenePix/Axon design files 20
To import a genome build 21
To import tracks 22
To import array attributes 23
To import an experiment file 24
To import a gene list 24

Working with Experiments to Organize Imported Data 25

To view the array designs and data in the program 25
To create a new experiment 26
To add arrays to an experiment 27
To change the order of arrays in an experiment 28
To rename an array in an experiment 29
To remove arrays from an experiment 29
To select or remove calibration array(s) 30
To edit the attributes of an experiment 31
To edit the attribute values of a specific array 31

Managing Content 33

To view a list of the content stored in the program 33
To find specific content items in the Navigator 33
To view the properties of a specific design 34
To update probe annotation in design files 34
To rename an array in the Data pane 35
To remove data or design files from the program 36

To create a gene list 37
To view the genes in a gene list 38
To add one gene list to another 38
To rename a gene list 38
To delete gene list(s) 39
To combine tracks 39
To view the details of a track 40
To rename a track 41
To delete tracks 41

Exporting and Saving Content 42
To export array attributes 42
To export experiments 43
To export a gene list 44
To export tracks 44
To copy what you see in the main window 45
To copy the list of array colors for an experiment 45
To save data and design information from an experiment 46

2 Visualizing CH3 Data and Other Content 47

Activating an Experiment for Visualizing Data 48
To activate an experiment 48
To select or deselect arrays in the experiment 49
To change the display color of an array 51

Visualizing Array Data 53
To display the scatter plots 53
To show significant data points in a scatter plot 53
To change scatter plot appearance 54
To print the scatter plot 55
To locate and view data (or results) within the Views 55

Visualizing Content (Gene Lists/Tracks) 58
To show gene lists in Gene View 58
Select gene list display color 59

View a gene list as a table 59
To change the appearance of genes in Gene View 60
To show tracks in Gene View 60
Change the appearance of tracks 61
Show track information in reports 62
Restrict data to the genomic boundaries of the track 62
Display tracks in UCSC Browser 63
To change the graphical display to a different genome build 63
To copy Views 64

Searching for Probe and Gene Information 65
To search Tab View for specific probe information 65
To search Agilent eArray for probe information 66
To search the Web for information on probes in Tab View 67
To create a custom Web search link 67
To update or delete a custom Web search link 68

3 Setting Up Methylation (CH3) Analysis 69

Working with Methylation Options 70

Changing Preprocessing and Analysis options 71
To combine (fuse) arrays 71
To set up a moving average (log ratio) calculation 73
To apply methylation (CH3) event detection 73
To set up a moving average (ZScore) calculation 74

Visualizing results and generating reports 76
To view results of analysis 76
To save a result 78
To restore a saved result to the display 79
To generate a report 79

4 Methylation (CH3) Analysis Reference 83

Genomic Workbench CH3 Application Main Window 84

Switch Application Menu 85

Command Ribbons 86

- Home command ribbon 86
- View command ribbon 89
- Preprocessing command ribbon 90
- Analysis command ribbon 91
- Reports command ribbon 95
- Help command ribbon 96

Navigator 98

- Search pane 99
- Data pane – icons, special text, and buttons 101
- Data pane – actions and shortcut menus 102
- Experiment pane – icons, special text, and buttons 104
- Experiment pane — actions and shortcut menus 106
- My Entity List pane – Icons, buttons, and special text 111
- My Entity List pane – Actions and shortcut menus 111
- Genome View 114
- Chromosome View 116
- Gene View 118
- The View Cursor 122
- Tab View 123

Status Bar 127

Dialog Boxes 128

- Add Gene List <name> to 128
- Agilent Feature Extraction Importer 129
- Array Set 131
- Choose Gene List Color 132
- Combine Tracks 133
- Confirm overwrite 135
- Create Experiment 136
- Create Gene List 138

Customize Search Link 140
Design Properties 141
Edit Array Color 144
Edit Array Order 145
Experiment Attributes 146
Experiment Properties 147
Export 149
Export Array Attributes – Array 151
Export Array Attributes – Attribute 153
Export Experiments 155
Export Tracks 156
Find in column 157
Gene List 159
Go To Gene/Genomic Location 160
Import 161
Import (experiments) 163
Import GEML design files 164
Import Genome Build 166
Import Track 167
Microarray Properties - Attribute Tab 168
Microarray Properties - FE Features Tab 169
Microarray Properties - FE Headers 170
Preferences – License 171
Preferences – Miscellaneous Tab 173
Preferences – Tracks Tab 174
Preferences – View Tab 176
Probe Methylation Status Setup 178
Probe Methylation Report Dialog 179
Scroll to Column 180
Search probes in eArray 181
Select Color (Edit Array Color) — Swatches Tab 182
Select Color (Edit Array Color) — HSB Tab 183
Select Color (Edit Array Color) — RGB Tab 184

Select data type for experiments 185
Set genome build and species for Axon design files 186
Show/Hide Columns 188
Track 189
UDF Import Summary 191
Universal Data Importer - Map Column Headers 192
Report Format 195

5 Statistical Algorithms 197

Methylation Detection and Visualization Algorithms Overview 198
Methylation Status Detection Algorithm 199
Visualization Algorithms 209
Triangular Smoothing 209
References 213

1 Importing, Managing, and Exporting CH3 Data and Other Content

Importing Files 13
Working with Experiments to Organize Imported Data 25
Managing Content 33
Exporting and Saving Content 42

This chapter describes how to import, organize, manage, and export methylation (CH3) data and other content within the user interface of Agilent Genomic Workbench.

The program lets you import many different kinds of files, including array data and design files from Agilent products and other sources, experiment files from DNA Analytics 5.0 and its predecessor DNA Analytics 4.0, genome build files, and other content such as annotation tracks and gene lists.

Once you have imported Feature Extraction files into the program, you can assign identification information and attributes to the array files through the Sample Manager tab. See the *Sample Manager User Guide*.

To view or analyze the data you have imported, you organize the data files into logical units called *experiments*. Experiments are central to the array viewing capabilities of Genomic Workbench. After you create them, and populate them with array data, you can then use the Preprocessing, Analysis, Discovery, and Report tabs of the program to analyze the data in the experiment for aberrations.

For an overview of the Preprocessing and Analysis options and what they do, see the *Agilent Genomic Workbench 5.0 Quick Start Guide*.

For details, see chapters 3 and 4 of this guide.

Experiments can also serve as the basis for automated, unattended CGH analyses in the Workflow tab. The commands under this tab also let you set up image files for automated, unattended Feature Extraction before array analysis. See the *Workflow Guide*.

Importing Files

You can use the Home tab to import many kinds of files into Genomic Workbench. The table below summarizes the kinds of files you can import, and the topics in this section that describe how to import them.

Type of file	Comments	See these topics
Microarray data files	• Agilent Feature Extraction (*.txt) data files • GenePix/Axon (*.gpr) data files • Universal Data Files (UDFs) (*.txt files)	"To import Agilent FE or GenePix/Axon data files" on page 14 "To import a UDF file" on page 16
Microarray design files	• Agilent GEML (*.xml) design files • GenePix/Axon (*.gal) design files	"To import Agilent GEML design files" on page 19 "To import GenePix/Axon design files" on page 20
Genome builds	Agilent-supplied genome information for human, mouse and rat genomes	"To import a genome build" on page 21
Tracks	BED format annotation track files	"To import tracks" on page 22
Array attributes	.txt files that you have created yourself or previously exported from Agilent Genomic Workbench	"To import array attributes" on page 23
Experiments	ZIP format file of experiments exported from DNA Analytics	"To import an experiment file" on page 24
Gene list	*.txt file – Creates a new gene list in the program with the list of imported genes	"To import a gene list" on page 24

NOTE To import Attribute Files, you use the **Sample Manager** tab. See the *Sample Manager Guide*.

To select a new location for data files

By default, the program stores design, data and experiment files in C:\Program Files\Agilent\Genomic Workbench Standard (or Enterprise) Edition 5.0\data. If you like, you can select a different location.

1 In the Home tab, click **User Preferences...**

The Preferences dialog box appears. See "Preferences – Miscellaneous Tab" on page 173.

2 In the **Miscellaneous** tab, under **Data Location,** click **Browse...**

An Open dialog box appears.

3 Select a location, then click **Open.**

The selected location appears in the Preferences dialog box, in Data Location.

4 Click **OK.**

To import Agilent FE or GenePix/Axon data files

You can import several types of microarray data files into Genomic Workbench:

- Agilent Feature Extraction (FE) *.txt data files
- GenePix/Axon (*.gpr) data files
- Universal Data Files (UDFs) (*.txt files) See "To import a UDF file" on page 16 for instructions on how to import this file type.

If you import Agilent Feature Extraction files, the program requires the representative GEML array design files. If you import GenePix/Axon data files, the program requires the representative GenePix/Axon *.gal design files. See "To import Agilent GEML design files" on page 19 or "To import GenePix/Axon design files" on page 20.

1 In the Home tab, do one of the following:

- To import Agilent FE data files, click **Import > Array Files > FE File...**
- To import GenePix/Axon data files, click **Import > Array Files > Axon File...**

A dialog box appears. Only data files of the appropriate type appear. See "Import" on page 161.

2 To select a file for import, click its name. To select additional files, control-click their names.

3 Do one of the following:

- For Agilent FE files, click **Open.**
- For GenePix/Axon files, click **Import.**

In either case, the Agilent Feature Extraction Importer dialog box appears. "Agilent Feature Extraction Importer" on page 129.

4 Set the following, as needed:

Setting	Comments
Name	The names of imported arrays are often cryptic. You can give any array a more meaningful label. **a** Double click the name of the array. **b** Edit the name, as desired. **c** Press **Enter (Return** on a Mac).
Dye Flip	For each array: • Select **Normal** if: ◦ The test samples were labeled with cyanine-5 (red). ◦ The control samples were labeled with cyanine-3 (green). ◦ The imported ration (test/control) should be reported directly. • Select **Flipped** if: ◦ The test samples were labeled with cyanine-3 (green). ◦ The control samples were labeled with cyanine-5 (red). ◦ The imported ratio (control/test) should be reported with the ratio inverted (test/control). The program does not combine dye-flip pairs.
Overwrite arrays with duplicate names	If you mark this option, the program deletes an existing array data file if it has the same name as one you import.

5 Do one of the following:

- To import the file(s) while you wait, click **OK.**
- To import the file(s) in the background, click **Run in Background.** This allows you to work while the program imports the files.

To import a UDF file

UDF files are plain text files that contain array data in tab-delimited format. Files must contain the following six columns of information, in any order:

- Probe name
- Chromosome name
- Start position
- Stop position
- Description
- Signal intensity data (The file can contain additional columns, each with data from an additional array.)

When you import a UDF file, the program creates a new design based on the information you provide during import, and the information in the file itself. This design contains all of the arrays represented in the file. The program also creates and populates a new experiment that contains the arrays.

1 In the Home tab, click **Import > Array Files > UDF File...**

The UDF Files dialog box appears. See "Import" on page 161. Only *.txt files appear in the dialog box.

2 Select the desired UDF file, then click **Open.**

The Select data type for experiments dialog box appears. "Select data type for experiments" on page 185.

3 For each array, set the following, as needed:

Setting	Comments
Experiment Name	By default, the program creates an experiment with the same name as the imported file. To change the name: **a** Double click the name. **b** Edit the name as desired. **c** Press **Enter.**

Setting	Comments
Data type	• Select the mathematical form of the signal intensity data for the array. The options are **ratio, $\log_2$ ratio, $\log_{10}$ ratio,** and **ln ratio.**
Design type	• Select **cgh, expression,** or **CH3.**

4 Click **Continue.**

The Universal Data Importer – Map column headers dialog box appears. The main table in the dialog box contains the first few rows of data from the file. Column headings derived from the first line of the file appear at the top of the table as a guide, but the program does not interpret these headings. See "Universal Data Importer - Map Column Headers" on page 192.

5 Below each column heading, select the label that identifies the content of the column. Use each label exactly once, except for LogRatio, which you can use many times. Alternatively, in **Select Mapping,** select a saved column map.

These options are available:

Label	Description
ProbeName	The column contains names of probes.
ChrName	The column contains names of chromosomes.
Start	The column contains the first chromosomal location to which each probe is designed.
Stop	The column contains the last chromosomal location to which each probe is designed.
Description	The column contains text annotation related to the probe.
LogRatio	The column contains array data values that correspond to each probe. You can use this label more than once.

6 Under **Species Info,** select the **species** and **Genome Build** appropriate to the data in the file.

7 If you expect to import many similar UDFs in the future, follow these steps to save the column map:

   a Under **Mapping Info,** click **Save Mapping As.**
   An Input dialog box appears.

   b Type a name for the column map, then click **OK.**
   The name of the saved map appears in Select Mapping.

   In the future, you can select this mapping and apply it to any UDF file that you import.

8 By default, the program creates a "Virtual Barcode" that becomes the Chip Barcode attribute for the array(s) in the UDF. To substitute a Virtual Barcode of your own choosing, follow these steps:

   a Under **Barcode Info,** clear **Use System Generated Barcode.**

   b Double-click the number in **Virtual Barcode,** then type the desired new Virtual Barcode.

9 Click **Import**.

   The program validates your column mapping. A dialog box appears. If you need to fix the column map, the dialog box details the missing column label(s). If the column map is complete, a message asks if you want to import additional files with the same mapping.

10 Do one of the following:

   - If you want to import additional files with the same column mapping, follow these steps to include these files in the import:

     a Click **Yes.**
     The UDF Files dialog box appears.

     b Click the name of a file to select it for import. Control-click the names of additional files.

     c Click **Open.**

   - If you do not want to include additional file(s) in the import, click **No.**

   The Program imports all requested files, and the UDF Import Summary dialog box appears. This dialog box lists the imported files, the number of lines of data that were imported for each file, and the number of lines that were skipped, if any. If a file name appears in red, the program may not have imported the file. See "UDF Import Summary" on page 191.

11 Click **OK.**

In the Data pane, in the appropriate design type folder within the Data folder, a new design folder appears. The design folder contains the imported array data.

A new experiment appears in the Experiments folder in the Experiment pane, populated with array data. This experiment bears the name of the imported UDF file, unless you changed it during import.

To import Agilent GEML design files

You must import Agilent GEML (*.xml) microarray design files into Genomic Workbench that match the Agilent Feature Extraction data files. Your imported GEML files contain array-specific information such as probe names, annotations, and chromosomal locations, and are associated with a specific genome build.

1 In the Home tab, click **Import > Design Files > GEML File...**

The Import Design Files dialog box appears. See "Import" on page 161. The dialog box shows only *.xml files.

2 To select a file for import, click its name. To select additional files, control-click their names.

3 Click **Open.**

The program validates the selected file(s), and the Import GEML Design Files dialog box appears. See "Import GEML design files" on page 164.

If a design file fails validation, **Corrupt** appears in the Status column beside it, and the program will not import the file. To remove the corrupt array from the list, click its **Remove** button  .

4 Click **Start Import.**

The program imports the file(s). The files appear as new design folders in the Data folder of the Navigator with the genome build as a node within the folder.

You can import two design files with the same name, but associated with different genome builds. If you do, the program creates a single design folder with two nodes, one for each genome build.

To import GenePix/Axon design files

You can import GenePix/Axon (*.gal) microarray design files into Genomic Workbench. The program requires the GenePix/Axon design files that match all GenePix/Axon array data files you import.

1 In the Home tab, click **Import > Design Files > Axon File...**

The Import Axon Design Files dialog box appears. See "Import" on page 161. The dialog box shows only *.gal files.

2 To select a file for import, click its name. To select additional files, control-click their names.

3 Click **Import.**

The program validates the selected file(s), and the Set genome build and species for Axon design files dialog box appears. See "Set genome build and species for Axon design files" on page 186.

If a design file fails validation, **Corrupt** appears in the Status column beside it, and the program will not import the file. To remove the corrupt array from the list, click its **Remove** button  .

4 For each design file, select the appropriate **Species** and **Genome Build.**

5 Click **Start Import.**

The program imports the file(s). The files appear as new design folders in the Data folder of the Data pane, organized by application (CGH, ChIP, or methylation, for example).

To import a genome build

In general, the program uses the genome build specified in the array design file, and protects it from changes. If a genome build is not available in the program, you can import one.

NOTE Use arrays from a single genome build in an experiment.

1 In the Home tab, click **Import > Genome Build...**

The Import Genome Build dialog box appears. See "Import Genome Build" on page 166.

2 Set the following. All are required.

Setting	Instructions
Species	• Type the genome's species of origin, as you would like it to appear within the program.
Build Name	• Type the name of the genome build you want to import, as you would like it to appear within the program.
RefSeq File	This file contains information on gene locations for Gene View. **a** Click **Browse...** A dialog box appears. **b** Select the desired file, then click **Open.**
Cyto-band File	This file contains the graphic information on the cyto-bands for Genome and Chromosome Views. **a** Click **Browse...** A dialog box appears. **b** Select the desired file, then click **Open.**

3 Click **OK.**

To import tracks

You can import BED format track files into DNA Analytics. Track files contain specific features correlated with chromosomal locations, and apply to a specific genome build of a given species.

1 In the Home tab, click **Import > Track...**

The Import Track dialog box appears. See "Import Track" on page 167.

2 Set the following. All are required.

Setting	Instructions
Species	• Select the species to which the track applies.
Build Name	• Select the specific genome build of the species to which the track applies.
Track Name	• Type a name for the track. This name identifies the track within the program, including the name that appears if you include the track in Gene View.
Track File	**a** Click **Browse...** A dialog box appears. **b** Select the name of the track (*.bed) file that you want to import. **c** Click **Open.** The location of the file appears in Track File.

3 Click **OK.**

The program imports the track. To view the track in Gene View, and to manage tracks, see "To show tracks in Gene View" on page 60.

To import array attributes

An array attributes file is a tab-delimited *.txt file that contains a list of arrays by barcode, and values for specific array attributes. Attributes are pieces of array-specific information, such as the hybridization temperature and the name of an array set that contains the array.

Although you can import array attributes with this function, the Sample Manager application lets you import and assign array attributes much more easily. See the *Sample Manager User Guide*. This menu item will be eliminated from the next version of the program, and the only way to import array attributes will be through the Sample Manager application.

To import an array attributes file

1 Click **File > Import > ArrayAttributes.**

The Import microarray attributes dialog box appears. See "Import" on page 161.

2 Select the desired microarray attributes file, then click **Import.**

The program imports the file. If the barcodes in the file do not match the barcodes of arrays in the program, a dialog box appears. The dialog box lists the barcodes in the file that do not match. Click **No** to stop the import process, or click **Yes** to continue anyway.

Array attributes files must follow these guidelines:

- The first line of the file contains the names of array attributes, separated by tabs.
- The first attribute must be `Chip Barcode`.
- The rest of the lines of the file list the values of each attribute, one line per array. The order of attribute values must follow the order of attributes in the first line of the file.

Here is an example.

Chip Barcode	ArraySet	Hyb'd By
251270010402	E986	Maurice R.
251270010423	E986	Maurice R.
251270019455	E986	Maurice R.

To import an experiment file

In Genomic Workbench, an experiment is a set of links to microarray data and design files, and any associated results. A Genomic Workbench experiment file is a single ZIP file that contains the design and data files associated with one or many experiments. You can import experiment files created in Genomic Workbench on another computer, as well as DNAx 4.0 experiments.

1 In the Home tab, click **Import > Experiments.**

The Import Experiments dialog box appears. See "Import" on page 161.

2 Select the ZIP file that contains the experiment(s) you want to import, then click **OK.**

The program imports the experiment file. Designs appear as new folders in the Data pane, in the applicable design type folder. Array data appears within the applicable design folder, organized by genome build. In addition, the experiment(s) appear, populated with the appropriate arrays, in the Experiment pane.

NOTE

Genomic Workbench experiment files contain all of the design and array data files associated with an experiment, but do not include any analysis parameter settings, array selections, or analysis results. To export the data and design files from one or more experiments, see "To export experiments" on page 43.

To import a gene list

A gene list file is a plain text (*.txt) file that contains one gene name per line. When you import a gene list into Genomic Workbench, it appears in the Gene List folder in the My Entities List pane. You can use the gene list to highlight specific genes, or to restrict the appearance of genes and data, in Gene and Chromosome Views. See "To show gene lists in Gene View" on page 58.

1 In the **My Entities List** pane, double click the **Entities** folder it.

2 Right-click the **Gene List** folder, then click **Import Gene List.**

3 Select the desired gene list file.

4 Click **OK.**

Working with Experiments to Organize Imported Data

This section describes how to arrange imported array data and designs into organizational units called *experiments*. Experiments, found in the Experiment pane of the Navigator, contain links to specific array data and design files in the Data pane. After you set up an experiment, you can then analyze selected array data within the experiment. Because experiments only contain *links* to the actual data and design files, any number of experiments can use a given set of files. In the data analysis applications (CGH, ChIP, or methylation, for example), experiments also contain saved experiment results.

To view the array designs and data in the program

- To view the directory of data in the program, use the Data pane (Figure 1). Double-click a folder to expand or collapse it, or click the ⊞ and ⊟ buttons.

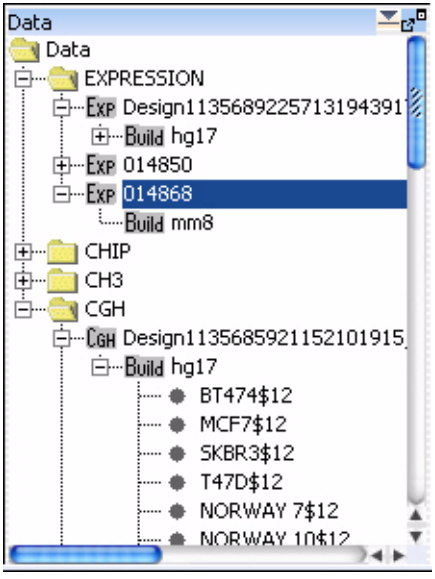


Figure 1 Data pane of the Navigator

In the Data pane, the program organizes design files by the application (CGH, ChIP, or methylation, for example) to which they apply. It organizes array data files by genome build under the design with which they are associated.

You can right-click many elements of the Data pane to open shortcut menus. For details, see "Data pane – actions and shortcut menus" on page 102.

Many icons can appear in the Data pane. See "Data pane – icons, special text, and buttons" on page 101 for a complete list.

The Search pane can help you find specific data files or other content. See "To find specific content items in the Navigator" on page 33.

To create a new experiment

In Genomic Workbench, *experiments* are organizational units that contain links to data and design files. To view or analyze data, you must first create an experiment and associate the desired data files with it. Because experiments only contain *links* to the actual data and design files, any number of experiments can use a given set of files. In data analysis applications (CGH, ChIP, or methylation, for example), experiments can also contain saved experiment results.

1 In the Home tab, click **Create Experiment...**

  The Create Experiment dialog box appears. See "Create Experiment" on page 136.

2 Type a **Name** and an optional **Description** for the experiment.

3 Do one of the following:

  - To create an empty experiment, and add data to it later, click **OK.** The program creates the experiment. To add arrays to the experiment later, see "To add arrays to an experiment" on page 27.

- To create an experiment and populate it with data now, follow these steps: (You can add or remove data from the experiment later, as well.)

  a Click **Properties.**
  The Experiment Properties dialog box appears. See "Experiment Properties" on page 147.

  b Under **Select Design,** select the design and genome build associated with the desired array data.
  The applicable arrays appear in Array List.

  c In **Array List,** click the name of an array that you want in your experiment. Control-click the names of additional arrays.

  d Click  .
  The program transfers the selected arrays to the Selected Array List.
  The dialog box also gives you other options for adding arrays. See "Experiment Properties" on page 147 for details.

  e Click **OK.**
  The program creates the new experiment, and populates it with the selected arrays.

In both cases, a folder with the name of the new experiment appears in the Experiment pane.

To add arrays to an experiment

After you create an experiment, or import one, you can add arrays to it. When you add arrays to an experiment, you create links between the experiment and the array data and design files. Because the program does not move the actual files, any number of experiments can share the same arrays.

1 In the **Experiment** pane, double-click the **Experiments** folder to expand it.

2 Right-click the name of the desired experiment, then click **Show Properties.**

The Experiment Properties dialog box appears. See "Experiment Properties" on page 147.

3 Under **Select Design,** select the design file and genome build associated with the arrays you wish to add.

The arrays associated with the selected design file and genome build appear in Array List.

4 In **Array List,** select the arrays you wish to add to the experiment. To select a single array, click its name. To select additional arrays, control-click their names.

5 Click > .

The program transfers the selected arrays to the Selected Array List.

The dialog box also gives you other options for adding arrays. See "Experiment Properties" on page 147 for details.

6 Click **OK.**

If needed, the program adds appropriate design and genome build folders to your experiment folder in the Experiment pane. It places the arrays you selected in the appropriate genome build folder.

To change the order of arrays in an experiment

You can change the order in which arrays appear in an experiment in tables in Tab View. If you choose to display separate scatter plots in Gene View and Chromosome View for each array, the array order also determines the order in which these plots appear. You can use this feature to organize your arrays more logically, or to make it more convenient to view certain arrays. It is especially useful if you have many arrays.

1 In the Experiment pane, right-click the name of the desired experiment, then click **Edit Array Order...**

The Edit Array Order dialog box appears. See "Edit Array Order" on page 145.

2 In **Design,** select the design that contains the arrays whose order you want to change.

The arrays from the selected design appear in Array Name.

3 Do any of the following:

- To move an array up in the list, click its name, then click ▲.
- To move an array down in the list, click its name, then click ▼.

- To sort the list based on a specific microarray attribute, select the desired attribute in **Order by.**

4 Click **OK.**

To rename an array in an experiment

When you rename an array in an experiment, you change the array's name only within the context of the given experiment. The name of the array remains unchanged in the Data pane, and in other experiments.

1 Expand the folders in the **Experiment** pane until you can see the array you wish to rename.

2 Right-click the name of the desired array, then click **Rename.**

An Input dialog box appears.

3 Type the new name for the array, then click **OK.**

To remove arrays from an experiment

When you remove arrays from an experiment, you only remove the links between the experiment and the data files. The files remain available in the program for use in other experiments. To completely remove files from the program, see "To remove data or design files from the program" on page 36.

1 In the **Experiment** pane, expand folders until you can see the desired experiment, and the array(s) that you want to remove from it.

2 In the **Arrays** or **Calibration Arrays** folder of the desired experiment, click the name of an array to select if for removal. Control-click the names of additional arrays.

3 Right click one of the selected array names, then click **Delete.**

A Confirm dialog box appears.

4 Click **Yes.**

The program removes the links between the experiment and the selected array data files. If the removal of arrays leaves a design folder in the experiment empty, the program removes this folder as well.

To select or remove calibration array(s)

After you add an array to an experiment, you can select it as a calibration array. The program lists calibration arrays within the Calibration Arrays folder of the experiment with a special icon  . You can also remove the calibration designation from an array.

To select an array as a calibration array

1. Expand the folders of the **Experiment** pane until you can see the array that you want to designate as a calibration array.
2. Right-click the name of the desired array, then click **Select for Calibration.** To select all of the arrays of a given design in the experiment as calibration arrays, right-click the genome build folder of the desired design, then click Set for Calibration.

   The program designates the array as a calibration array. In the Calibration Arrays folder of the applicable genome build and design within the experiment, the array appears with a special icon  .

To remove the calibration array designation from an array

1. Expand the folders of the **Experiment** pane until you can see the desired array. The program lists calibration arrays in the Calibration Arrays folder(s) of the applicable genome build(s) and design(s) within each experiment.
2. Right click the name of the desired array, then click **Deselect from Calibration.**

   To remove the calibration array designation from multiple arrays at once, select all of the desired arrays. Right-click one of the arrays, then click **Deselect from Calibration.** (To select multiple arrays, click the name of one array, then control-click the names of additional ones. To select a contiguous block of arrays, click the name of the first array, then shift-click the name of the last one.)

   The program removes the calibration designation from the array(s), and moves the arrays to the Arrays folder of the applicable genome build and design within the experiment. The icons of the arrays change to the standard (non-calibration) array icon  .

To edit the attributes of an experiment

Array attributes are pieces of array-specific information, such as chip barcode or hybridization temperature. You can show or hide attributes for the arrays in the experiment with the Experiment Attributes dialog box. See "Experiment Attributes" on page 146.

1 Right-click the experiment whose attributes you want to show or hide, or to change.

   You see the array attributes and their values that were set up in the Sample Manager table. See the *Sample Manager Guide*.

2 Click **Edit Attributes**.

3 Click **Show/Hide Attributes**.

   The Show/Hide Columns dialog box appears. See "Show/Hide Columns" on page 188.

4 Mark the check boxes whose attribute columns you want to show, or clear the check boxes whose attribute columns you want to remove.

5 Double-click the cell whose array attribute value you want to change.

6 Click **Save Changes**.

   If you do not want to save the changes, click **Cancel Changes**.

7 Click **Close**.

NOTE You cannot create new attributes using this dialog box. To do this, you must use the Sample Management tab. See the *Sample Manager Guide*.

To edit the attribute values of a specific array

Array attributes are pieces of array-specific information such as chip barcode or hybridization temperature. You can view a list of attributes for each array that is available in the program.

1 Expand the folders of the **Data** pane or the **Experiment** pane until you can see the array of interest.

2 Right-click the name of the array, then click **Show Properties.**

The Microarray Properties dialog box appears, with a list of array attributes. See "Microarray Properties - Attribute Tab" on page 168. You can also edit the attributes of a specific array from this dialog box. In addition, if the array is an Agilent array, you can view header and feature information passed through from the Agilent Feature Extraction program.

3 When you are finished, click **Close.**

NOTE You use the Sample Manager tab to organize, create, import, and export array attributes. See the *Sample Manager User Guide*.

Managing Content

This section describes how to create, find, rename, update, combine, and/or remove several types of content stored in Genomic Workbench. To display the data, gene list and track content, see Chapter 2, "Visualizing CH3 Data and Other Content".

To view a list of the content stored in the program

The Data and My Entity List panes of the Navigator show the content stored in Genomic Workbench.

Data pane – Shows all of the array data files stored in the program, organized by application (CGH, ChIP, or methylation, for example), then by array design, then by applicable genome build. For more information, see "To view the array designs and data in the program" on page 25.

My Entity List pane – Shows the gene lists and tracks stored in the program. To view the names of gene lists or tracks available in the program, double-click the names of folders to expand or collapse them, or click the ⊞ or ⊟ buttons.

NOTE Content that is available exclusively in the Design tab, such as the probes and probe groups that you use to create custom microarray designs, is covered in a separate section. See the *eArray$_{XD}$ User Guide*.

To find specific content items in the Navigator

At the top of the Navigator is a search pane that can help you find specific content items. See "Search pane" on page 99.

1 Type a search term in the box at the top of the Navigator. The search term is not case-sensitive, but it must reflect the entire name of the

content item that you want to find. You can use asterisks (*) as wildcards to represent a group of unspecified characters.

2 By default, the program searches all panes of the Navigator. To restrict your search to a specific pane, click . In the list that appears, select the desired pane.

3 Click .

The program searches the selected pane(s). If it finds item(s) that match your search term, it expands folders so that the items are visible, and highlights them in red. You may need to scroll down to see retrieved items.

4 To clear the results of a search, click .

To view the properties of a specific design

Design properties include general information about a design, such as its name, application type, and associated species. They also include a list of the names and chromosomal locations of probes.

1 Expand the folders of the **Data** pane until you can see the genome build folder(s) within the desired design folder.

2 Right-click the desired genome build folder, then click **Show Properties.**

The Design Properties dialog box appears. See "Design Properties" on page 141.

To update probe annotation in design files

Agilent regularly updates probe annotations on its eArray Web portal. If you have imported Agilent array designs into Genomic Workbench, and you are a registered eArray user, you can update those design files from within Genomic Workbench. For more information about eArray, go to earray.chem.agilent.com and click **Help.**

1 In the Home tab, click **User Preferences.**

The Preferences dialog box appears.

2 In the Miscellaneous tab, under **eArray User Details,** type your eArray **Username** and **Password.** See "Preferences – Miscellaneous Tab" on page 173.

3 Click **OK.**

4 Expand the folders of the **Data** pane until you can see the design that you want to update.

5 Right-click the desired design, then click **Update from eArray.** This option appears only for Agilent designs.

A confirmation dialog box appears.

6 Click **Yes.**

The program downloads an updated design, if one is available.

To rename an array in the Data pane

This topic describes how to rename an array in the Data pane. If you rename an array in this way, and subsequently add the array to an experiment, the array appears in the experiment with the new name. However, the name of the array is unaffected in any experiment to which it is already linked. To rename an array only within the context of a specific experiment, see "To rename an array in an experiment" on page 29.

1 Expand the folders of the **Data** pane until you can see the array you want to rename.

2 Right-click the name of the array, then click **Rename.**

An Input dialog box appears.

3 Type a new name for the array, then click **OK.**

The program renames the array.

To remove data or design files from the program

You can delete array design and data files from the program when you are finished with them.

1. If an array that you want to delete is associated with an experiment, first delete it from the experiment. See "To remove arrays from an experiment" on page 29.
2. In the **Data** pane, expand folders until you can see the design folder or array that you want to delete.
3. Do one of the following:
   - For array data files, click the name of the first array, then control-click the names of additional arrays within the same design.
   - For array design folders, click the name of the first design folder, then control-click the names of additional ones. This selects the designs and all array data files within them for deletion.
4. Right-click the name of a selected design folder or array data file, then click **Delete.**

   A confirmation dialog box appears.
5. Click **Yes.**

   The program deletes the selected files.

CAUTION When you delete files, you permanently remove them from Genomic Workbench. To restore deleted files, you must import them again.

To create a gene list

When you create a gene list, you create a list of the genes in a contiguous chromosomal region that you define. To create a list of genes in multiple regions, create multiple gene lists, and combine them. See "To add one gene list to another" on page 38.

1 Follow these steps to define a chromosomal region for your gene list. If you know the exact start and end locations of the desired chromosomal region, skip to step 2.

   a In Genome View, select the desired chromosome.
   The selected chromosome appears in Chromosome View.

   b In Chromosome View, in the plotting area to the right of the chromosome, drag the pointer over the approximate desired chromosomal region.
   The program encloses the region in a blue box, and displays the region in greater detail in Gene View.

   c In Gene View, adjust the view so only the genes of interest appear. For a description of the adjustment commands available in Gene View, see "Gene View" on page 118.

2 Right-click anywhere within the log ratio plotting area in Gene View, then click **Create Gene List...**

   The Create Gene List dialog box appears. See "Create Gene List" on page 138.

3 In the dialog box set the Name, Description and Color.

4 In the dialog box select the desired chromosomal region for the new gene list.

5 Click **OK.**

   The new gene list appears in the My Entity List pane of the Navigator in the Gene List folder.

To view the genes in a gene list

You can view the genes in a gene list as a table.

1 Expand the folders in the **My Entity List** pane until you can see the desired gene list.

2 Right-click the gene list, then click **View In Table.**

The Gene List dialog box appears, with a table that contains the names of the genes in the gene list. You can also use this dialog box to edit the description of the gene list and its display color. See "Gene List" on page 159.

You can also create gene lists. For details, see "To create a gene list" on page 37.

To add one gene list to another

You can add one gene list (a source gene list) to another (the target gene list). The program appends the source gene list to the end of the target gene list, and leaves the source gene list unchanged.

1 Expand the folders in the **My Entity List** pane until you can see the gene lists that you want to combine.

2 Right-click the desired source gene list, then click **Add to Gene List.**

A dialog box appears. For details, see "Add Gene List <name> to" on page 128.

3 In **Select target gene list,** select the desired target gene list.

4 Click **OK.**

To rename a gene list

The name of a gene list identifies it within the Gene List folder of the My Entity List pane. You can rename gene lists.

1 Expand the folders of the **My Entity List** pane until you can see the gene list that you want to rename.

2 Right-click the desired gene list, then click **Rename.**

An Input dialog box appears.

3 Type a new name for the gene list, then click **OK.**

To delete gene list(s)

1 In the My Entity List pane of the Navigator, expand the **Gene List** folder.

2 Click the name of a gene list that you want to delete. Control-click the names of additional gene lists.

This selects the lists.

3 Right-click one of the selected gene lists, then click **Delete.**

A confirmation dialog box appears.

4 Click **Yes.**

To combine tracks

You can create a track that contains elements from two or more existing tracks. The existing tracks must be available in Genomic Workbench, and they must be associated with the same genome build.

1 In the **My Entities List** pane, double-click the **Entities** folder to expand it, if necessary.

2 Right-click the **Tracks** folder, then click **Combine Tracks.**

The Combine Tracks dialog box appears. See "Combine Tracks" on page 133.

3 In **Name,** type a name for the combined track. The program uses this name to identify the track in the Tracks folder, and to label the track if it appears in Gene View.

4 Click **New Condition.**

A new row appears in the Track/Operator list.

5 Under **Track,** select the first track that you want to combine.

6 Under **Operator,** select one of the following:

Operator	Comments
AND	Creates a combined track out of 2 tracks consisting of elements that appear in both tracks
OR	Creates a combined track out of 2 tracks consisting of elements that appear in either of the tracks
MINUS	Removes the elements of the second track from the first track.

7 Click **New Condition,** then select another Track/Operator pair. You can set up as many Track/Operator pairs as you like, but you must set up at least two.

To remove the bottom row from the list, click **Delete Condition.** To delete all rows from the list, and erase any entry in Name, click **Reset.**

8 Click **Save.**

Your combined track appears in the Tracks folder of the My Entities List pane. The Combine Tracks dialog box remains open for you to create another combined track.

9 Click **Close.**

To view the details of a track

The table that you bring up contains the values for a list of track attributes.

1 In **My Entity List** pane, expand the Tracks folder to see the track.

2 Right-click the name of the track, then click **View Details.**

Data describing the track appear in a Track table. See "Track" on page 189.

To rename a track

The name of a track identifies it both within the Tracks folder of the My Entity List pane, and in Gene View when you select **Show in UI** for the track. You can rename tracks.

1 Expand the folders of the **My Entity List** pane until you can see the track that you want to rename.

2 Right-click the desired track, then click **Rename.**

An Input dialog box appears.

3 Type a new name for the track, then click **OK.**

To delete tracks

1 In the My Entity List pane of the Navigator, expand the **Tracks** folder.

2 Click the name of a track that you want to delete. Control-click the names of additional tracks.

This selects the tracks.

3 Right-click one of the selected tracks, then click **Delete.**

A confirmation dialog box appears.

4 Click **Yes.**

Exporting and Saving Content

This section describes how to export several kinds of files from the program. You can also export many additional kinds of files from the separate data analysis applications and other parts of the program. See the appropriate User Guides.

To export array attributes

You can export selected array attributes for any imported arrays that you choose. You first select the arrays and then the attributes that you want exported for your array selection. You can access this capability from the Home tab or the short-cut menu for an experiment.

1 Click **Home > Export > Array Attributes...**

OR

In the Experiment pane of the Navigator, right-click an experiment of interest, and click **Export Attributes**.

The Export Array Attributes dialog box appears with the Array tab displayed. See "Export Array Attributes – Array" on page 151.

If you opened this dialog box by right-clicking an experiment, only those arrays associated with the experiment appear in the Selected Array List. You can add or subtract from the list.

2 Under **Select Design,** select the design file and genome build associated with the arrays you wish to add.

The arrays associated with the selected design file and genome build appear in Array List.

3 In **Array List,** select the arrays whose attributes you intend to export. To select a single array, click its name. To select additional arrays, control-click their names.

4 Click 

.

The program transfers the selected arrays to the Selected Array List.

5 Click **Next** to choose attributes for the selected arrays.

The Export Array Attributes dialog box appears with the Attribute tab displayed. See "Export Array Attributes – Attribute" on page 153.

All of the attributes for the arrays are already located in the Selected Attribute List. If you don't want all the attributes exported, then you must transfer those to the Available List.

6 In the Selected Attributes List, highlight those attributes you do not intend to export.

7 Click  .

8 Click **OK**.

The Export dialog box appears. See "Export" on page 149.

9 Select the folder in which to locate the attributes, and click **Export**.

The attributes will be saved to the selected folder as a .txt file.

To export experiments

You can export experiments as a ZIP file to transfer them to another computer. Exported experiments contain the associated design and array data files, only. The program does not export information about array selections, or any analysis parameters or results.

1 In the Home tab, click **Export > Experiments...**

The Export Experiments dialog box appears. See "Export Experiments" on page 155.

2 Mark the experiments that you want to export. To export all experiments, click **Select All.**

3 In **Export Format,** select one of these options:

- **5.0 Format** – Exports the experiment(s) in a format that you can import into Genomic Workbench. This is the most current experiment format, but it is not compatible with previous versions of the program.
- **3.0 Format** – Exports the experiment(s) in a format that you can import into Agilent CGH Analytics 3.0 or later. This is a "legacy" format that you can use to maintain compatibility with earlier versions of the program.

4 Click **OK.**

An Export dialog box appears. See "Export" on page 149.

5 Select a location and type a name for the exported ZIP file.

6 Click **Export.**

The program exports all selected experiment(s) together as a single ZIP file.

To export a gene list

You can export a gene list as a text file that contains one gene per line.

1 In the **My Entity List** pane, in the **Gene List** folder, right-click the gene list that you want to export, then click **Save As.**

A Save As dialog box appears.

2 Select a location and type a name for the file.

3 Click **Save.**

A success message appears.

4 Click **OK.**

To export tracks

You can export selected tracks as a BED format track file. You can then import this file into Genomic Workbench on another computer, or into a genome browser that accepts BED format files.

1 In the **Home** tab, click **Export > Tracks...**

The Export Tracks dialog box appears. See "Export Tracks" on page 156.

2 Mark the tracks that you want to export. To select all tracks for export, click **Select All.**

3 Click **OK.**

An Export dialog box appears.

4 Select a location and type a name for the exported track file, then click **Export.**

The program exports the track(s) as a single BED format track file.

To copy what you see in the main window

You can copy panes of the main window to the clipboard as images, and then paste them into a new document in another program (such as Word, or PowerPoint). The images contain only what actually appears on your screen—regions to which you must scroll are not included.

1 In the **View** tab, click **Copy.**

2 In the shortcut menu that appears, click the name of the pane that you want to copy. You can copy any view, or the Navigator. To copy all of the panes, click **All.**

The program copies the selected pane(s) to the clipboard.

3 Open a document in a program that accepts images. In that program, click **Edit > Paste,** or the appropriate paste command.

To copy the list of array colors for an experiment

You can copy the list of arrays in an experiment, and the colors assigned to them, to the clipboard as an image. You then paste the image into a document in another program such as Word or PowerPoint.

1 In the **Experiment** pane, expand the **Experiments** folder.

2 Right-click the name of the desired experiment, then click **Edit Array Color...**

The Edit Array Color dialog box appears. See "Edit Array Color" on page 144.

3 In the dialog box, click **Edit > Copy.**

The program copies the names of the arrays and their colors to the clipboard as an image.

4 Open a program that accepts images. Click **Edit > Paste,** or the appropriate paste command for the specific program.

To save data and design information from an experiment

You can save the data and design information from a single design in an experiment as a tab-delimited text file.

1. In the **Experiment** pane, expand the **Experiments** folder until you see the genome build(s) associated with the design you want to export.
2. Right-click the name of the desired genome build, then click **Save As Text File...**

   A dialog box appears.
3. Select a location and type a name for the saved file, then click **Save.**

2
Visualizing CH3 Data and Other Content

Activating an Experiment for Visualizing Data 48
Visualizing Array Data 53
Visualizing Content (Gene Lists/Tracks) 58
Searching for Probe and Gene Information 65

This chapter shows you how to display log ratio data from imported feature extraction data files, as well as gene list and track content, in the Genomic Viewer. It also gives you instructions on how to modify the display to visualize the data and content the way you prefer.

To find out how to import, organize, manage and export data and other content, Chapter 1, “Importing, Managing, and Exporting CH3 Data and Other Content”.

To learn about the options for the main window and the dialog boxes for visualizing data, see Chapter 4, “Methylation (CH3) Analysis Reference”.

See the *Agilent Genomic Workbench 5.0 Quick Start Guide* for an overview of all the applications you can use with this software, both those that require a license and those that are free capabilities.

Activating an Experiment for Visualizing Data

An experiment is a set of links to microarray data and design files, and any associated results.

When you activate an experiment and the Preprocessing and Analysis options have not been turned on or set to apply, the program shows the log ratio data of selected arrays in the active experiment, if certain options have been set.

When you activate an experiment and Preprocessing and Analysis options have been turned on or set to apply, the program automatically begins the analysis of the selected array data with current settings and displays its results, if certain options have been set.

This section describes how to activate an experiment, select or deselect arrays for further analysis, and analyze arrays one at a time.

To activate an experiment

When you make an experiment active, the program begins the analysis with the current settings. You can either set the Preprocessing and Analysis parameters that you want before you activate the experiment or change the settings one at a time after the first analysis and re-analyze. Every time you change a Preprocessing or Analysis setting for an active experiment, the program recalculates results.

1 Create a new experiment and populate it with data. See "To create a new experiment" on page 26.

2 In the Navigator, double-click the name of the experiment.

The Experiment Selection dialog box appears.

3 Click **Yes.**

In the Experiment pane of the Navigator, the name of the experiment turns blue. The name also appears in the title bar of the main window. Tables of data and design information appear in Tab View. For details on the available tabs, see "Tab View" on page 123.

If you have chosen to show the results of an algorithm calculation, then moving average or aberration results appear for the first array when you activate the experiment, if you have not selected any other arrays.

You can select or deselect arrays in the experiment both before and after you activate it. Every time you select or deselect an array in an *active* experiment or change a setting, the program re-analyzes the new data set with the changed settings. See "To select or deselect arrays in the experiment" on page 49.

When you reactivate or select the experiment after deselecting it or selecting another one, the experiment is simply restored if the settings in the UI haven't changed. If they have changed, the program re-analyzes all of the arrays selected when the experiment was last activated.

To select or deselect arrays in the experiment

To include arrays for view and analysis, you select them from among the arrays available either in an inactive experiment or the active one. When you first create an experiment, the program automatically sets the first array in the experiment for analysis. If you do not select any more arrays for analysis, only the first one will be analyzed when the experiment is activated.

To select the arrays for analysis before experiment activation:

1. Hold down the **Shift** key to highlight contiguous arrays or hold down the **Ctrl** key to highlight noncontiguous arrays.
2. Right-click the highlighted arrays, and click **Select**.

   Even though the selected arrays do not change color, they will change color after activation.

In the Navigator, an array's icon has two appearances after experiment activation:


Array not selected.


Array selected. The specific color matches the color of the column headings for the array in Tab View. In addition, the program displays aberration results and moving averages related to this array in this color.

To configure a custom color for the array, see "To change the display color of an array" on page 51.

To select or deselect arrays in an *active* experiment:

1. In the Navigator, expand the folders of the active experiment.
2. Click the name of an array you want to include in the analysis.

   To include additional arrays, control-click their names. To include a contiguous block of arrays, click the name of the first array in the block, then shift-click the name of the last one.
3. Right-click the name of one of the highlighted arrays, then click **Select.**

   After you select the arrays, the program re-analyzes the data set within the experiment and posts the data in Genome, Chromosome, and Gene Views. You can see the data and results for just the selected arrays in the Selected Arrays tab in Tab View.

   To show analysis results if they do not appear, see "To view results of analysis" on page 76.

   To customize the appearance of the results in Genome, Chromosome, and Gene Views, see the *Data Viewing Guide*.

You can also use the headings of columns in Tab View that contain array data to select and deselect arrays.

- Click a column heading to select that array only.
- Control-click a column heading to select or deselect an array without affecting the status of other arrays.
- Right-click a column heading to open a shortcut menu with options that allow you to select or deselect that array, or all arrays.

For more details on Tab View, see the *Data Viewing Guide*.

To change the display color of an array

The color assigned to an array affects the color of its icon when you select the array within an experiment. It also affects the colored square in the array's column heading in Tab View.

1 In the Experiment pane, in the **Experiments** folder, expand the folder of an experiment until you can see the array whose color you want to edit.

2 Right-click the desired array, then click **Edit Array Color...**

The Select Color dialog box appears. The dialog box offers three different ways to choose the desired color. "Select Color (Edit Array Color) — Swatches Tab" on page 182.

3 Select the desired color in one of the following ways:

Dialog box tab	Instructions
Swatches	• Click the desired color swatch.
HSB (Hue/Saturation/Brightness)	Type or adjust the values in H (Hue), S (Saturation), and B (Brightness), or alternately, follow these steps: **a** Select H, then drag the slider to select a hue based on the color strip to its right. **b** Click an appropriate location in the large color box to the left of the slider to set the saturation and brightness levels of the color. Both the HSB and equivalent RGB values of the color appear in the dialog box. Note these values—they will be useful if you need to duplicate this color in the future.
RGB (Red/Green/Blue)	Do any of the following. Note the final RGB Values—they will be useful if you need to duplicate this color in the future. • Drag the Red, Green, and Blue sliders. • Type or adjust values in the boxes to the right of the sliders.

Samples of the color in different contexts appear under Preview. The upper half of the right-most color sample shows the original color for comparison.

4 Adjust the color as desired, then click **OK.**

You can also manage all of the colors for all of the arrays in an experiment. Right-click the desired experiment, then click **Edit Array Color...** For details on the dialog box that appears, see "Edit Array Color" on page 144.

Visualizing Array Data

After you activate an experiment, you can change how data appear within the Views or change the appearance of the Views that contain the data (or results).

To display the scatter plots

1 In the Gene View, click **Scatter Plot,** and do any of the following:

To do this	Follow these steps
Show or hide all log ratio data points	• To show all data points – Mark all three **Log Ratio** check boxes • To hide all data points – Clear all three **Log Ratio** check boxes.
Show or hide significant or insignificant data points	See "To show significant data points in a scatter plot" on page 53.
Change the size of data points	• In **Point Size,** select a size for the data points.

2 Click X to close the Scatter Plot window.

To show significant data points in a scatter plot

Although you must use the data analysis algorithms to perform statistical calculations on your array data, the CH3 application can classify log ratio data points as significant, or not, based on a simple cutoff value.

You select the cutoff value, and the program displays data points whose log ratios are above, below, or within the range of the cutoff value on the scatter plots in Gene and Chromosome Views, in three different colors.

1 In the Gene View, click **Scatter Plot**.

2 In **Cutoff,** select the desired cutoff value. The cutoff is an absolute value, and it defines a range. For example, if you select a value of 1.25,

the program uses a cutoff range of –1.25 to +1.25 to classify data points.

If you select **None,** the program classifies all data points as significant.

3 Mark the kinds of points you want to appear in the scatter plot:

 – (Red) Points with log ratios above the selected cutoff range

 – (Green) Points with log ratios below the selected cutoff range

 – (Black) Points with log ratios within the selected cutoff range

4 Select the **Point Size** (in pixels) for the points in the scatter plot.

5 Click X to close the Scatter Plot window.

To change scatter plot appearance

You use the Preferences dialog box to change the appearance of the scatter plots in Chromosome and Gene views.

1 Right-click any part of a scatter plot, then click **Preferences.**

The Preferences dialog box appears. "Preferences – View Tab" on page 176.

2 Do any of the following:

To do this	Follow these steps
Show or hide the scatter plot	**a** In the View tab under **Data Visibility,** in **View,** select **All Views.** **b** Do one of the following: To show the scatter plot, mark **Scatter Plot.** To hide the scatter plot, clear **Scatter plot.** **c** Click **OK.**
Change the symbol that appears for data points	You can select the symbol separately for each design type. **a** In the View tab, under **Rendering Patterns,** select the desired **Design type.** **b** Under **Styles,** in **Scatter Plot,** select the desired symbol. **c** Click **Apply.**

To do this	Follow these steps
Show a separate scatter plot in Gene and Chromosome Views for each selected array	**a** In the View tab, under **View Alignment,** under **Rendering Style,** select **Stacked.** **b** Click **Apply.**
Show one scatter plot that contains data for selected arrays	**a** In the View tab, under **View Alignment,** under **Rendering Style,** select **Overlaid.** **b** Click **Apply.**
Enable tool tips for the scatter plot in Gene View	Tool tips show information about an individual data point when you place the pointer over it. **a** Click the **View** tab. **b** Under **Data Visibility,** in **View,** select **Gene View.** **c** Mark **Scatter Tool Tip.** **d** Click **Apply.**

3 Click **OK.**

To print the scatter plot

You can print the scatter plot as it appears in Genome, Chromosome, and Gene views. Each view selected in the analysis is printed on separate pages. Chromosomes and genes appear on the printed pages, but tracks do not.

1 Click **File > Print.**

The Print dialog box appears.

2 Set print options, as desired, then click **OK.**

To locate and view data (or results) within the Views

- To navigate through the data of the selected arrays, do any of the following. In general, all views are synchronized–if you select a location or region in one view, the other views move there as well.

To do this	Follow these steps
Select a specific chromosome to view	• In Genome View, click the desired chromosome. All other views switch to the selected chromosome.
View data in a region of the selected chromosome	• In Chromosome View, in the scatter plot, drag the pointer over the desired region. Gene View expands (or shrinks) to show only the selected region. Tab View scrolls to the new cursor location.
Zoom in and out in Gene View	• Click  to zoom in. • Click  to zoom out.
Scroll through the selected chromosome	• Click  to scroll up. • Click  to scroll down.
Re-center Gene View or Chromosome view	Click anywhere in Chromosome View, or anywhere within the scatter plot in Gene View. The location you click becomes the new cursor location.
Move all views to a specific genomic location	**a** Click **Home > Go To Gene/Genomic location...** A dialog box appears. **b** Under **Genomic Location,** select a **Chromosome,** and type a **Base Position.** **c** Click **Go.** All views move to the selected location.
Center all views on the location of a specific gene	**a** Click **Home > Go To Gene/Genomic location...** A dialog box appears. **b** Under **RefSeq by Symbol,** either select the desired gene (if available) or type the name of the gene. **c** Click **Go.** All views move to the location of the selected gene.
Center Chromosome and Gene views based on data in Tab View	• In Tab View, click any entry in any table, except a column heading. Chromosome and Gene views become centered on the genomic location corresponding to the selected entry.

To do this	Follow these steps
Scroll to a specific column in Tab View	**a** In Tab View, right-click any column heading, then click **Scroll To Column...** A dialog box appears **b** In **Select Column,** select the desired column. **c** Click **OK.**
View the exact chromosomal location of the cursor	At the bottom of the main window, look at the first cell of the Status bar. The location appears as the chromosome followed by the base position. For more information on the status bar, see "Status Bar" on page 127.

Visualizing Content (Gene Lists/Tracks)

To show gene lists in Gene View

A gene list defines a set of genes of interest. Within the program, you can highlight the genes in the gene list in Gene View, or restrict the display of data, genes, and tracks to the regions defined by a gene list.

You can import gene lists into DNA Analytics, and you can also create them in the program and export them. See "To import a gene list" on page 24, and "To export a gene list" on page 44.

In Gene View, the names of all genes normally appear in gray. When you apply a gene list, the program highlights the listed genes in their defined display color. You can also restrict the genes and/or data that appear in Gene View and Chromosome View to only the listed genes.

1 In the My Entity List pane of the Navigator, expand the **Gene List** folder. If the desired gene list does not appear, create or import it. See "To create a gene list" on page 37, or "To import a gene list" on page 24.

2 Right-click the desired gene list, then do one of the following to apply it:

- To show all genes and all data, and highlight the listed genes in their display color, click **Highlight.**
- To show only the listed genes and only the data associated with those genes, click **Show only.**

Gene and Chromosome views change accordingly. In the My Entity List pane of the Navigator, the name of the gene list appears in italics.

To remove the effects of a gene list, right-click the active gene list in the Navigator, then click **Show All.**

Select gene list display color

In Gene View, the names of all genes normally appear in gray. When you apply a gene list, the program highlights the listed genes in their defined display color. You can customize this color.

1 In the My Entity List pane of the Navigator, expand the **Gene List** folder.

2 Right-click the name of the gene list whose color you want to change, then click **View in Table.**

The Gene List dialog box appears.

3 Under **Color,** click **Color...**

A dialog box appears.

4 Select the desired color.

The dialog box offers three different ways to choose the desired color. See "Select gene list display color" on page 59.

5 Adjust the color as desired, then click **OK.**

6 In the Gene List dialog box, click **OK.**

View a gene list as a table

You can view the description of a gene list and the names of the genes in it.

1 In the My Entity List pane of the Navigator, in the **Gene List** folder, right-click the desired gene list, then click **View in Table.**

The Gene List dialog box appears. See "Gene List" on page 159. The names of the genes appear in Gene Names. You can also use this dialog box to edit the description of the gene list, or to change its display color. To change the display color, see "Select gene list display color" on page 59.

2 When you are finished viewing the list, click **OK.**

You can also export a gene list. See "To export a gene list" on page 44.

To change the appearance of genes in Gene View

You use the Preferences dialog box to change the appearance of the genes in Chromosome and Gene views.

1 Right-click any part of the Gene View, then click **Preferences.**

The Preferences dialog box appears.

2 Click **Tracks**.

"Preferences – Tracks Tab" on page 174.

3 Do any of the following:

To do this	Follow these steps
Show or hide genes in Gene View	**a** Under **Visualization Parameters**: To show genes – Under **Genes,** mark **Show Gene Symbols.** To hide genes – Under **Genes,** clear **Show Gene Symbols.** **b** Click **Apply.**
Change the display font for genes (and track annotations) in Gene View	**a** In the Gene Symbols tab, under **Font,** select a new **Font, Font Style,** and **Font Size.** **b** Click **Apply**
Change the display angle for genes (and track annotations) in Gene View	**a** Under **Visualization Parameters,** under **Genes,** in **Orientation (Degrees),** type a new orientation in degrees. 0° is horizontal. **b** Click **Apply.**

4 Click **OK.**

To show tracks in Gene View

Tracks contain information indexed to specific genomic locations. A multitude of tracks from diverse sources is available for many species. You can display tracks alongside genes and microarray data in Gene View.

1 Select and show microarray data. See "To activate an experiment" on page 48.

2 In the My Entity List pane, open the Tracks folder.

3 Right-click the track you want to display, and click **Show in UI**.

Or, you can do this:

1 In Gene View, right-click anywhere within the scatter plot, then click **Preferences.**

The Preferences dialog box appears. See "Preferences – Tracks Tab" on page 174.

2 Click **Tracks**.

3 Mark the **Show in UI** check box of each desired track.

4 Click **OK.**

The program displays the selected tracks in Gene View.

Change the appearance of tracks

Within the Preferences – Tracks dialog box, you can modify the appearance of tracks in several additional ways. See the table below.

To do this	Follow these steps
Include track information in reports	**a** In the list of tracks, in the **Show in Report** column, mark the check boxes of the desired tracks. **b** Click **Apply.** Doing this adds a column with the hits from the track file. For each aberrant interval, it reports the entries from the track file that falls under that interval in that separate column.
Show or hide annotations in all tracks	• To show annotations in all tracks: under **Tracks,** mark **Show Annotations.** • To hide annotations in all tracks: under **Tracks,** clear **Show Annotations.**
Display all selected tracks as a single track	• Under **Tracks,** mark **Show Overlaid.** The program combines the annotations of all selected tracks into a single track named **Overlaid Track.** • To show tracks individually again, clear **Show Overlaid.**
View the parameters and the list of annotations of a track	• In the list of tracks, next to the desired track, click **Details...**

To do this	Follow these steps
Change the display font for track annotations (and genes)	**a** Under **Font,** select a new **Font, Font Style,** and **Font Size** for track annotations. **b** Click **Apply.** The program changes the display font of track annotations and genes in Gene View.
Change the order in which tracks appear in Gene View.	The order of tracks in the Gene Symbols tab controls the left-to-right order of tracks in Gene View. **a** Click the name of the track you want to move. **b** Do any of the following: • To move the track up in the list of tracks (and farther left in Gene View), click its name, then click **Up.** • To move the track down in the list of tracks (and farther right in Gene View), click its name, then click **Down.** **c** Click **Apply.**
Change the display angle of track annotations (and genes)	• Under **Genes,** in **Orientation,** type a new orientation (in degrees). 0° is horizontal. The program changes the display angle of track annotations and genes in Gene View.

Show track information in reports

1 In the list of tracks, in the **Show in Report** column, mark the check boxes of the desired tracks.

2 Click **Apply.**

Doing this adds a column with the hits from the track file. For each aberrant interval, it reports the entries from the track file that falls under that interval in that separate column.

Restrict data to the genomic boundaries of the track

1. In the list of tracks in My Entity List, right-click the track whose boundaries you want to use to restrict the display of the data.

3 Mark **Genomic Boundaries**.

You can remove the boundaries by clearing the check box.

Display tracks in UCSC Browser

1 Right-click Gene View, and click **Show in UCSC**.

The UCSC Browser appears if you are connected to the Internet.

2 Follow the instructions on the web site for what you want to do.

To change the graphical display to a different genome build

The default graphical display for Genome, Chromosome and Gene Views represents human genome build 18.

- To change the graphical display to a different genome build, activate an experiment whose data are based on a design file of a different genome build.

The display automatically changes when you activate an experiment containing a design file with a different genome build, such as human genome build 17, or a mouse or rat genome build.

If a genome build is not available for the design file you import, you must import the genome build first. See "To import a genome build" on page 21.

The program will not let you add arrays belonging to one genome build to an experiment containing arrays of a different genome build.

See also "To create a new experiment" on page 26, "To add arrays to an experiment" on page 27, and "To activate an experiment" on page 48.

To copy Views

You can copy panes of the main window to the clipboard as images, and then paste them into a new document in another program. The images contain only what actually appears on your screen—regions to which you must scroll are not included.

1 Click **View > Copy.**

A menu of Views appears.

2 Click the View you want to copy to the clipboard. To copy all available views as a single image, click **All.**

3 Open a document in another program that accepts images, such as a word processor or graphics program.

4 In the other program, click **Edit > Paste.**

The copied image appears.

Searching for Probe and Gene Information

To search Tab View for specific probe information

You can find a specific entry in a column of a data table in Tab View. For more information on Tab View, see "Tab View" on page 123.

1 In Tab View, right-click anywhere in the column you want to search, then click **Find in column.**

The Find in column dialog box appears. The column to be searched also appears in the title bar of the dialog box.

2 Set the search parameters, as described below.

Parameter	Comments/Instructions
Find in column	• Type the text you want to find (the *search term*). This can be an entire entry, or part of one.
Direction	• Select one of these options: ◦ **Up** – Search the column upwards from the current cursor location (the highlighted row of the table). ◦ **Down** – Search the column downwards from the current cursor location (the highlighted row of the table). Tip: Click a row in Tab View to highlight it.
Conditions	• Mark any of these, as desired: ◦ **Match Case** – Return entries that match upper and lower case characters in the search term. ◦ **Match whole word** – Return an entry only if the entire entry matches the search term.

3 Click **Find Next.**

If the program finds a match, it highlights the row that contains the matching entry, and resets the cursor to the corresponding position. You can click **Find Next** as many times as you like, and the program continues to search for additional matching entries in the column. If it finds no match, **String not found.** appears in black at the bottom of the dialog box.

4 When you finish your search, click **Cancel.**

To search Agilent eArray for probe information

You can use the chromosomal region that appears in Gene View, or another chromosomal region as the basis for a probe search on the Agilent eArray Web site. eArray is a powerful microarray design system for CGH, ChIP and gene expression applications. It contains a massive database of validated, annotated probes, and a full complement of tools for custom microarray design.

Before you can search for probes in eArray, you must be a registered eArray user. For more information, go to eArray.chem.agilent.com. You must also provide your eArray user name and password in the Miscellaneous tab of the Preferences dialog box. See "Preferences – Miscellaneous Tab" on page 173.

1 In Gene View, right-click anywhere in the plotting area, then click **Search probes in eArray...**

   The Search probes in eArray dialog box appears. "Search probes in eArray" on page 181.

2 Do one of the following to define the chromosomal region for your search:

   - To set the region to the one that currently appears in Gene View, select **For complete gene view.**
   - To set the region numerically, select **User Defined,** then select a **Chromosome** and type **Start** and **Stop** locations for the desired region.

3 Click **OK.**

   The eArray Web portal opens in your internet browser.

To search the Web for information on probes in Tab View

You can use any entry in a table in Tab View as the basis for a Web search.

1 In Tab View, right-click any data table entry other than a column heading.

2 Click one of the available sites.

If the site you want does not appear in the shortcut menu, you can create a custom search link. See "To create a custom Web search link" below.

The selected site opens in your Internet browser. The program passes the table entry to the site as a search string.

To create a custom Web search link

If you need to search a different database or site based on data table entries, you can create your own custom search link. When you right-click a table entry in Tab View, a shortcut menu opens, and your custom link appears in it. If you select this link, DNA Analytics opens the site in your Web browser and passes the table entry to the site as a search string.

1 Right-click any data table entry in Tab View, except a column heading, then click **Customize Link...**

The Customize Search link dialog box appears. See "Customize Search Link" on page 140.

2 Click **New.**

3 In the Input dialog box, in **URL name,** type a name for the link.

This name will appear in the shortcut menu that opens when you right-click a data table entry.

4 Click **OK.**

5 In **URL,** type the complete URL needed to pass a search string to the site. Use `<target>` as the query string value.

For example, this URL passes selected table entries to Google.com:
`http://www.google.com/search?hl=eng&q=<target>`

6 Click **Update**, then click **Yes.**

To update or delete a custom Web search link

1 Right-click any data table entry in Tab View other than a column heading, then click **Customize Link...**

The Customize Search link dialog box appears.

2 In **URL Name,** select the custom search link you want to update or delete.

3 Do one of the following:

To do this	Follow these steps
Update a Web search link	**a** Edit the **URL name** and the **URL** as needed. **b** Click **Update.** A Confirm dialog box appears. **c** Click **Yes.**
Delete a Web search link	• Click **Delete.**

4 Click **Close.**

3
Setting Up Methylation (CH3) Analysis

Working with Methylation Options 70
Changing Preprocessing and Analysis options 71
Visualizing results and generating reports 76

This chapter gives instructions on how to set up the analysis functions for Personal Forensics and Toxicology Database experiments. These include the Preprocessing, Analysis and Reports tabs.

Chapter 2 presents the descriptions of the tabs, commands, shortcut menus, window options, dialog box options and icons you use in the Methylation (CH3) Analysis to set up analysis for Personal Forensics and Toxicology Database data.

Chapter 3 shows how the analysis algorithms work.

For an overview of the Methylation (CH3) Analysis and a quick run-through of how you can use it to view and analyze Personal Forensics and Toxicology Database data, from importing data to generating reports, see the *Agilent Genomic Workbench Quick Start Guide*.

If you are reading a hard copy of this guide and have not installed the software yet or do not know how to activate the license, please see the installation and licensing instructions online or in the *Agilent Genomic Workbench Quick Start Guide*.

Working with Methylation Options

For a detailed description of the Methylation (CH3) analysis tabs and their commands – Preprocessing, Analysis, Reports – see Chapter 4, "Methylation (CH3) Analysis Reference".

Figure 2 CH3 Analysis tabs

Table 1 lists the sections in this chapter along with the tasks you do to prepare for and use the analysis options.

Table 1 Methylation analysis topics

Subject	See these topics
"Changing Preprocessing and Analysis options"	"To combine (fuse) arrays" on page 71 "To set up a moving average (log ratio) calculation" on page 73 "To apply methylation (CH3) event detection" on page 73 "To set up a moving average (ZScore) calculation" on page 74
"Visualizing results and generating reports"	"To view results of analysis" on page 76 "To locate and view data (or results) within the Views" on page 55 "To save a result" on page 78 "To restore a saved result to the display" on page 79 "To generate a report" on page 79

Changing Preprocessing and Analysis options

To combine (fuse) arrays

If you have two arrays that use different design files, you can combine them into one larger virtual array. You do this to increase the coverage of the genome in your design. If for example, you have a catalog array, you can design another array to add probes between the catalog probes to increase the density of coverage. With the Fuse function, you can combine the array data to see all the probe data in the display at once.

The program cannot combine arrays from more than two different design files. See the requirements for fusing arrays after the instructions:

1 Create and activate a new experiment. See "To create a new experiment" on page 26 and "To activate an experiment" on page 48.

2 Add the arrays you want to fuse to the experiment. See "To add arrays to an experiment" on page 27.

3 Assign the same value to the **ArraySet** attribute of every array you want to fuse. Follow these steps for each array:

   a In the **Experiment** pane of the Navigator, right-click the name of the array, then click **Show Properties.**
   The Microarray Properties dialog box appears.

   b Next to the ArraySet attribute, under **Value,** click  .
   A text box appears.

   c Type a value in the text box. Type the same value for the ArraySet attribute of every array.

   d Click **Close.**

4 In the **Preprocessing** tab, under **Combine,** click **Fuse.**

   The Array Set dialog box lists the arrays to be fused. See "Array Set" on page 131.

NOTE Double-check the values in the ArraySet Attribute column of the dialog box. DNA Analytics fuses all of the array pairs that have the same value for this attribute.

5 Set any of these options, as desired:

- **Select Normalization** – Select **None** or **Centralization.** Centralization adds or subtracts a constant value from each log ratio measurement. This re-centers the log ratio values, and ensures that the zero-point reflects the most common ploidy state.
- **Remove arrays from experiment after fuse** – To delete the initial un-fused arrays from the experiment, mark this option. This reduces the duplication of data within the experiment.

6 Click **Continue.**

The program fuses the arrays. The fused array appears in the Experiment pane of the Navigator in a new design folder within the active experiment. The folder name contains the names of both designs.

Requirements for fusing arrays:

- Each array must be associated with a different design file.
- All of the arrays must be of the CH3 application type.
- None of the arrays can be fused arrays.
- The samples you hybridize to the arrays must all be aliquots from the same preparation.
- (Preferred) Hybridization and labeling occur for all samples together under the same conditions.

NOTE

- If the original arrays you fused change, you can manually update the fused array. Fuse the same set of arrays again.
- If you want to fuse many arrays, consider importing an array attributes file. See "To import array attributes" on page 23.
- When you combine (fuse) arrays, the centralization algorithm is applied to the individual arrays before they are fused.

To set up a moving average (log ratio) calculation

With the Moving Average (Log Ratio) group you can have a moving average computed for each point in the data set using a window size centered on that point. See "Visualization Algorithms" on page 209.

1 In the **Analysis** tab, in the **Moving Average (Log Ratio)** group, mark or clear **Show** to show or hide the Moving Average Line Plot.

2 In **Algorithm,** select **Linear** or **Triangular.**

Linear – The linear algorithm calculates a standard, unweighted average using every Log Ratio score within a specified window size (specified by either a number of adjacent measurements or a positional range). This average is applied sequentially to every point by moving the window along the chromosome.

Triangular – The triangular algorithm calculates a weighted average using every Log Ratio within a specified window size. The triangular algorithm is more sensitive to localized variations in the data. See "Triangular Smoothing" on page 209.

3 In **Line width,** select the desired thickness (in pixels) of the Moving Average Line Plot.

The range is 1-5 pixels.

4 In **Window,** select the size of the moving average window.

Moving averages can be computed with windows of sizes based on either a specific length of base-pairs (5 Kb to 50 Kb or .1Mb to 50 Mb) or a fixed number of data points (1pt. to 60 pt.).

To apply methylation (CH3) event detection

The purpose of applying this algorithm is to identify methylation sites located within genomic CpG islands and gene promoter regions. It is used in conjunction with CpG Island tracks that either already exist in the program's database or you must import.

1 Set up CpG Island tracks in Gene View.

See chapters 1 and 2 of this guide for instructions on how to import and visualize these tracks in Gene View. and in the report.

2 Activate the experiment, or wait to activate it after step 3.

See “To activate an experiment” on page 48.

3 Click **Analysis**, and mark **Apply** in the Probe Methylation group.

The Probe Methylation Status Setup dialog box appears.

See “Probe Methylation Status Setup” on page 178.

4 To select a Tm (melting temperature) mapping file, click **Browse** and go to the folder containing the file.

5 Highlight the file, and click **Open**.

6 Click **Continue**.

The algorithm is automatically applied if you’ve activated the experiment, or will be applied when you do. To learn how the algorithm works see “Methylation Status Detection Algorithm” on page 199.

The Z-score results appear in a second channel of Gene View next to the log ratio data. You can now calculate a moving average for the Z-scores. See the next task.

To set up a moving average (ZScore) calculation

The Moving Average (ZScore) group applies a moving average to each point in the combined methylation Z-score data set using a window size centered on that point. See “Methylation Detection and Visualization Algorithms Overview” on page 198.

1 In the **Analysis** tab, in the **Moving Average (ZScore)** group, mark or clear **Show** to show or hide the Moving Average Line Plot.

2 In **Algorithm,** select **Linear** or **Triangular.** See **“Visualization Algorithms” on page 209.**

Linear – The linear algorithm calculates a standard, unweighted average using every combined Z-score within a specified window size (specified by either a number of adjacent measurements or a positional range). This average is applied sequentially to every point by moving the window along the chromosome. See “Visualization Algorithms” on page 209.

Triangular – The triangular algorithm calculates a weighted average using every combined Z-score within a specified window size. The triangular algorithm is more sensitive to localized variations in the data. See “Triangular Smoothing” on page 209.

3 In **Line width,** select the desired thickness (in pixels) of the Moving Average Line Plot.

The range is 1-5 pixels.

4 In **Window,** select the size of the moving average window.

Moving averages can be computed with windows of sizes based on either a specific length of base-pairs (5 Kb to 50 Kb or .1Mb to 50 Mb) or a fixed number of data points (1pt. to 60 pt.).

Visualizing results and generating reports

The section shows you how to make sure the analysis results are displayed after experiment activation. To learn more about what the results mean, see "Visualization Algorithms" on page 209.

To view results of analysis

After you set up an experiment and analyze it, the program displays the results automatically in Genome, Chromosome, and Gene Views. Figure 3 shows an example of a result displayed in the Gene View. To customize the way the results appear, see the *Data Viewing Guide*.

Results for probe methylation and log ratios compared to a CpG Island track

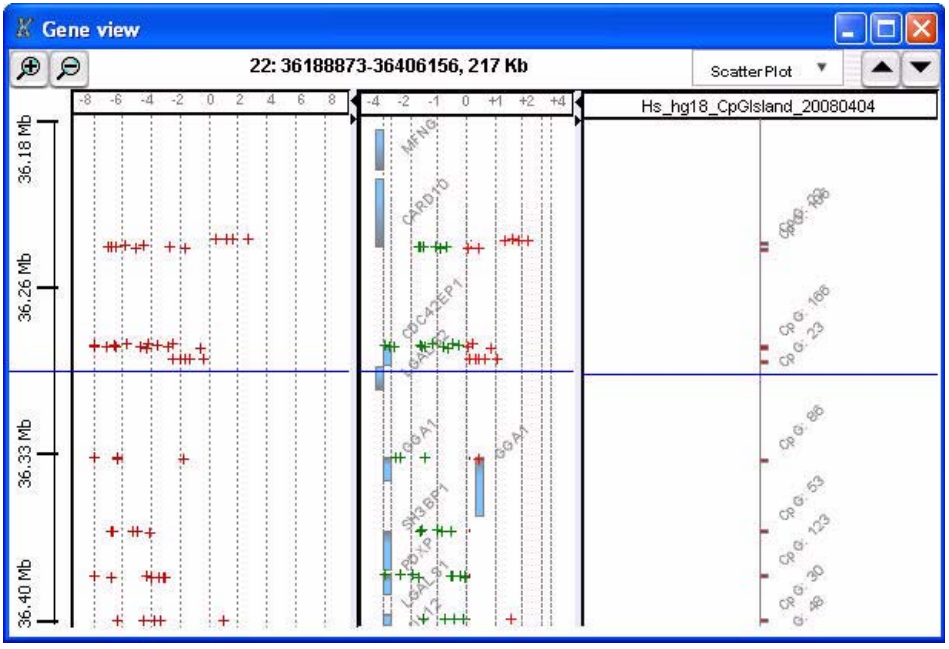


Figure 3 Gene View, showing Z-score data next to log ratios and a CpG Island track

Results for moving averages within genomic boundaries of CpG Island track

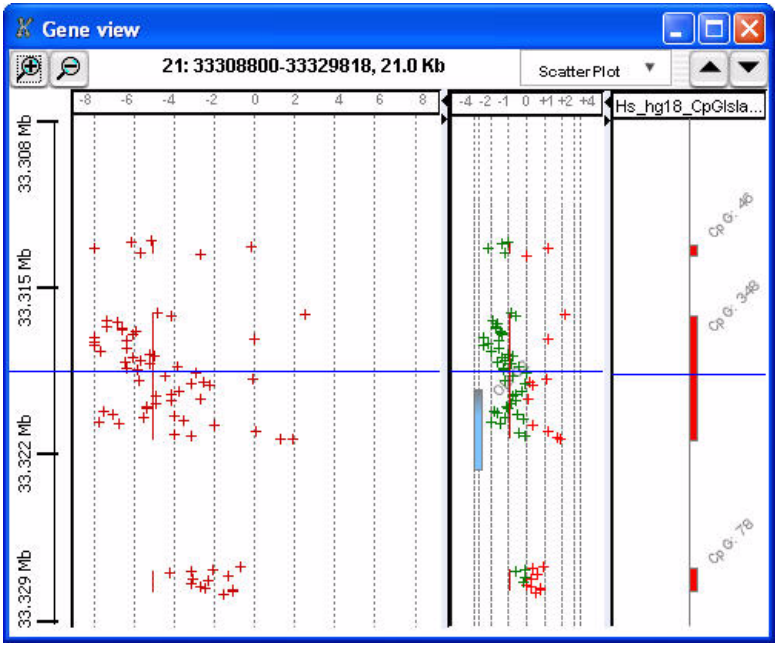


Figure 4 Gene View, showing moving averages within genomic boundaries of tracks

If the analysis results do not appear in any of the Views

Check these four potential problem areas in order:

- You may not have marked the Apply check box. Under Analysis, mark **Apply** for Probe Methylation.
- You may not have activated the experiment or selected the arrays of interest. See "To activate an experiment" on page 48.
- You may not have turned on the display for the moving averages or the Z-score algorithm. Do the following:

  1 Right-click any of the Views, and click **Preferences**.

  2 Click the **View** tab.

  3 Under **Data Visibility,** in **View,** select **All views,** then mark any of the applicable check boxes whose information you want to see.

You can also select just one View.

4 Click **OK.**

- You may not be looking in the right region of Chromosome View.

See "Visualizing CH3 Data and Other Content" on page 47 to find out how to display the Scatter Plot, if you want to see it, and how to navigate the View displays to observe results.

To save a result

The program allows you to save the current result of the active experiment. You can run many different analyses in the same experiment, and save each one. Later, you can restore any of your saved results.

If you are saving a result for the first time for the experiment:

1 In the **Experiment** pane of the Navigator, right-click the experiment, and click **Save Experiment Result**, or

Click **Home > Save Experiment Result.**

A dialog box asks if you want to save the results of the current experiment.

2 Click **Yes**.

The Save experiment result dialog box appears.

3 Type a name for the result, then click **OK.**

If you have already saved at least one result for the experiment:

1 In the **Experiment** pane of the Navigator, expand the folders of the current experiment.

The currently selected result, if any, appears in blue in the Results folder.

2 Click **Home > Save Experiment Result or**

Right-click the experiment, and click **Save Experiment Result**.

A dialog box asks if you want to take one of three actions:

- If you would like to replace the current result with another saved result, choose **Overwrite Current Result.**

- If you would like to add the current results to the list of experimental results, choose **Create New Result.**
- If you would like to change views to another result without changing the current result, choose **Continue Without Saving.**

To restore a saved result to the display

1 If necessary, activate the experiment that contains the result that you want to see. See "To activate an experiment" on page 48.

2 In the **Experiment** pane of the Navigator, expand the folder of the active experiment, then expand its **Results** folder.

3 Right-click the desired result, then click **Restore result.**

The restored result appears in Genome, Chromosome, and Gene Views.

To generate a report

The report includes the following information either for single chromosomes or the complete genome:

Log ratio for each event

Number of probes

Start and stop of each event

Whether the probe is in a CpG island, a promoter, or inside a gene boundary

Likely gene the probe is associated with

1 Click **Reports**, then click **Generate Report**.

The Probe Methylation Report Setup dialog box appears.

2 To specify if you want a report listing the results for the whole genome or for individual chromosomes, click either **Complete Genome** or **Per-Chromosome**.

3 Click **Browse** to select the folder for containing the report(s).

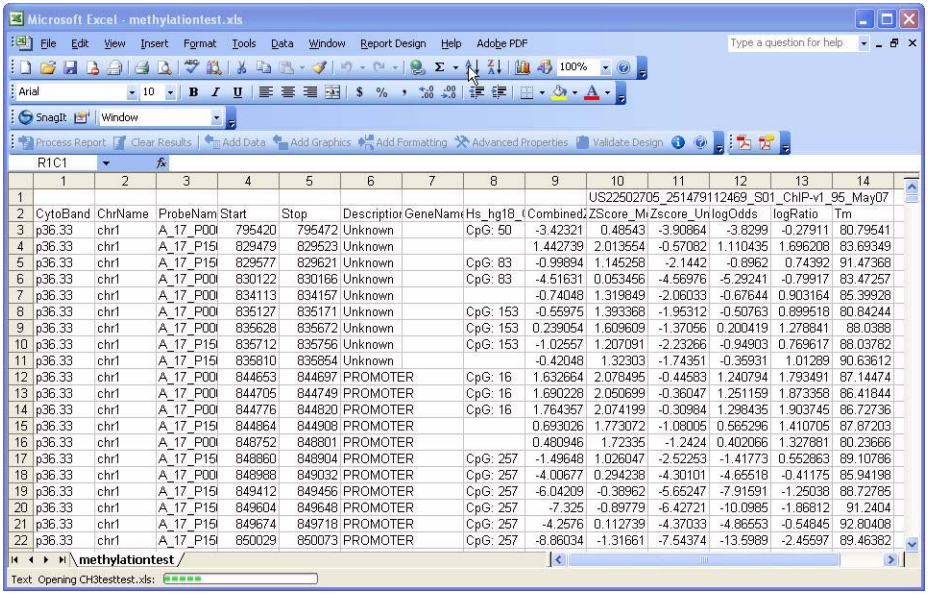
4 Enter the report name, and click **Open**.

The report name contains the suffix, xls.

5 Click **OK**.

6 When a question pops up to ask you if you want to view the report, click **Yes** or **No**.

If you selected Complete Genome, an Excel spreadsheet appears containing all the data:



	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1										US22502705_251479112469_S01_ChIP-v1_95_May07				
2	CytoBand	ChrName	ProbeNam	Start	Stop	Descriptior	GeneNam	Hs_hg18_(	Combined	ZScore_Mi	Zscore_Un	logOdds	logRatio	Tm
3	p36.33	chr1	A_17_P00	795420	795472	Unknown		CpG: 50	-3.42321	0.48543	-3.90864	-3.8299	-0.27911	80.79541
4	p36.33	chr1	A_17_P15	829479	829523	Unknown			1.442739	2.013554	-0.57082	1.110435	1.696208	83.69349
5	p36.33	chr1	A_17_P15	829577	829621	Unknown		CpG: 83	-0.99894	1.145258	-2.1442	-0.8962	0.74392	91.47368
6	p36.33	chr1	A_17_P00	830122	830166	Unknown		CpG: 83	-4.51631	0.053456	-4.56976	-5.29241	-0.79917	83.47257
7	p36.33	chr1	A_17_P00	834113	834157	Unknown			-0.74048	1.319849	-2.06033	-0.67644	0.903164	85.39928
8	p36.33	chr1	A_17_P00	835127	835171	Unknown		CpG: 153	-0.55975	1.393368	-1.95312	-0.50763	0.899518	80.84244
9	p36.33	chr1	A_17_P00	835628	835672	Unknown		CpG: 153	0.239054	1.609609	-1.37056	0.200419	1.278841	88.0388
10	p36.33	chr1	A_17_P15	835712	835756	Unknown		CpG: 153	-1.02557	1.207091	-2.23266	-0.94903	0.769617	88.03782
11	p36.33	chr1	A_17_P15	835810	835854	Unknown			-0.42048	1.32303	-1.74351	-0.35931	1.01289	90.63612
12	p36.33	chr1	A_17_P00	844653	844697	PROMOTER		CpG: 16	1.632664	2.078495	-0.44583	1.240794	1.793491	87.14474
13	p36.33	chr1	A_17_P00	844705	844749	PROMOTER		CpG: 16	1.690228	2.050699	-0.36047	1.251159	1.873358	86.41844
14	p36.33	chr1	A_17_P00	844776	844820	PROMOTER		CpG: 16	1.764357	2.074199	-0.30984	1.298435	1.903745	86.72736
15	p36.33	chr1	A_17_P15	844864	844908	PROMOTER			0.693026	1.773072	-1.08005	0.565296	1.410705	87.87203
16	p36.33	chr1	A_17_P00	848752	848801	PROMOTER			0.480946	1.72335	-1.2424	0.402066	1.327881	80.23666
17	p36.33	chr1	A_17_P15	848860	848904	PROMOTER		CpG: 257	-1.49648	1.026047	-2.52253	-1.41773	0.552863	89.10786
18	p36.33	chr1	A_17_P00	848988	849032	PROMOTER		CpG: 257	-4.00677	0.294238	-4.30101	-4.65518	-0.41175	85.94198
19	p36.33	chr1	A_17_P15	849412	849456	PROMOTER		CpG: 257	-6.04209	-0.38962	-5.65247	-7.91591	-1.25038	88.72785
20	p36.33	chr1	A_17_P15	849604	849648	PROMOTER		CpG: 257	-7.325	-0.89779	-6.42721	-10.0985	-1.86812	91.2404
21	p36.33	chr1	A_17_P15	849674	849718	PROMOTER		CpG: 257	-4.2576	0.112739	-4.37033	-4.86553	-0.54845	92.80408
22	p36.33	chr1	A_17_P15	850029	850073	PROMOTER		CpG: 257	-8.86034	-1.31661	-7.54374	-13.5989	-2.45597	89.46382

Figure 5 Excel methylation report for Complete Genome

If you selected Per-Chromosome, the text files for the individual chromosomes appear in the folder you specified.

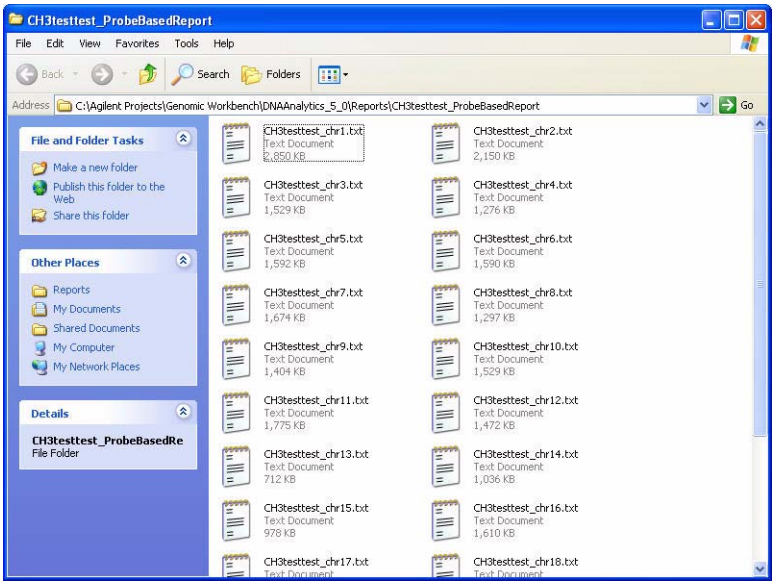


Figure 6 List of text methylation reports for each chromosome

4
Methylation (CH3) Analysis Reference

Genomic Workbench CH3 Application Main Window 84
Switch Application Menu 85
Command Ribbons 86
Navigator 98
Genomic Viewer 113
Status Bar 127
Dialog Boxes 128
Report Format 195

This chapter describes the parts of the Genomic Workbench main window that you use to import, organize, manage, export and display array data and other content. It also details the relevant tab commands, shortcut menus, and dialog boxes that can appear.

For specific instructions on how to use Genomic Workbench to manage data and accomplish related tasks, see Chapter 1, “Importing, Managing, and Exporting CH3 Data and Other Content.” To learn how to visualize data and content, see Chapter 2, “Visualizing CH3 Data and Other Content”.

This chapter also describes the tabs, commands, shortcut menus and dialog boxes specific to DNA Analytics methylation (CH3) data analysis. Another section describes the format of the reports the program generates. For instructions on how to analyze CH3 data interactively, see Chapter 3, “Setting Up Methylation (CH3) Analysis”.

To understand how all the guides work together to help you use Agilent Genomic Workbench, see the *Agilent Genomic Workbench 5.0 Quick Start Guide*.

Genomic Workbench CH3 Application Main Window

The main window of the CH3 Application contains the major components illustrated in Figure 1.

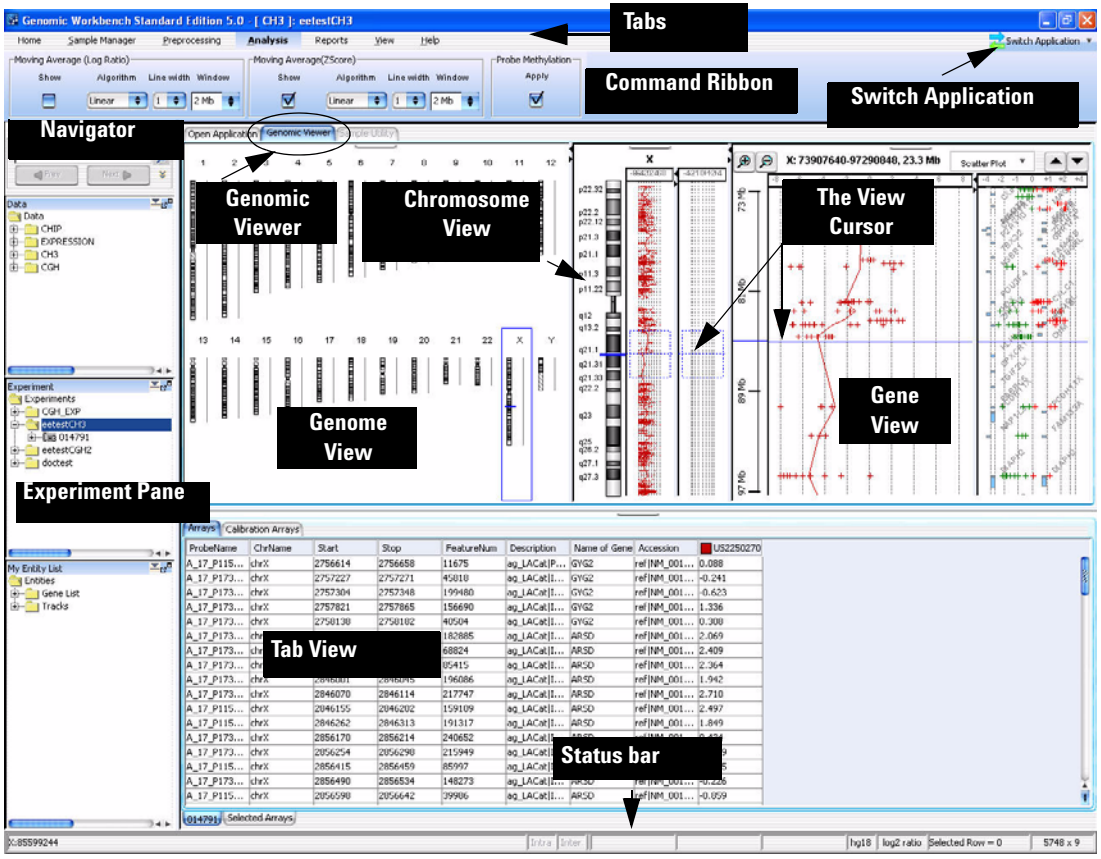


Figure 1 Genomic Workbench Standard Edition – CH3 Application main window

Switch Application Menu

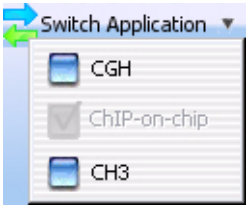


Figure 2 Switch Application menu

The Switch Application menu allows you to switch to the other data display and analysis applications in DNA Analytics. Mark the desired application.

CGH (Separate license required) Imports, displays, and analyzes array-based comparative genomics hybridization (aCGH) data in both an interactive "analyze as you go" mode, and an automated workflow mode.

ChIP (Separate license required) Imports, displays, and analyzes ChIP-on-Chip microarray data in both an interactive "analyze as you go" mode, and an automated workflow mode.

CH3 (Separate license required) Imports and displays data from microarray-based studies of genomic methylation patterns.

Command Ribbons

When you click a tab, groups of commands or single commands appear. The entire string of commands is called a command ribbon. This section describes the tab commands used to import, manage, export and display CH3 data/results, and those specific to methylation (CH3) analysis.

The interactive tabs of the CH3 Application include Home, View, Preprocessing, Analysis and Reports. This section provides descriptions of the commands for each of these tabs, as well as for the Help tab. Descriptions of the remaining tabs are located in the following guides:

Sample Management — *Sample Manager Guide*

Workflow — *Workflow Guide*

Figure 3 CH3 Application tabs

Home command ribbon

The Home command ribbon offers the functions that let you import, manage, export and display CH3 data and content (gene lists, tracks) for further CH3 analysis.

Figure 4 Command ribbon in the Home tab of Genomic Workbench

User Preferences... Opens the User Preferences dialog box with four tabs:

Tab	Description
View	Opens a dialog box that lets you change in what form the data will appear in Genomic Viewer. See "Preferences – View Tab" on page 176.
Tracks	Opens a dialog box that lets you manage which tracks to display in Genomic Viewer and how they appear. See "Preferences – Tracks Tab" on page 174.
Miscellaneous	Opens a dialog box where you can select a new location for your data files and set up access to the eArray web site. See "Preferences – Miscellaneous Tab" on page 173.
License	Opens a dialog box where you can add a DNA Analytics application license, if you choose to purchase one after using the unlicensed version. Preferences – License 171.

Import/Export

Import Opens a menu of file types that you can import:

Option	Description
Array Files	Opens a menu with these options: • **FE File...** – Opens the Import FE Files dialog box, where you can select an Agilent Feature Extraction array data file to import. See "Import" on page 161 and "To import Agilent FE or GenePix/Axon data files" on page 14. • **Axon File...** – Opens the Import Axon Files dialog box, where you can select Axon (*.gpr) files for import. See "Import" on page 161 and "To import Agilent FE or GenePix/Axon data files" on page 14. • **UDF File...** – Opens the UDF Files dialog box, where you can select a Universal Data File (UDF) to import. See "Import" on page 161 and "To import a UDF file" on page 16.

Option	Description
Design Files	Opens a menu with these options: • **GEML File...** – Opens the Import Design Files dialog box, where you can select Agilent GEML-based (*.xml) array design files for import. See "Import" on page 161 and "To import Agilent GEML design files" on page 19. • **Axon Design File** – Opens the Import Axon Design Files dialog box, where you can select Axon (*.gal) array design files for import. See "Import" on page 161 and "To import GenePix/Axon design files" on page 20.
Genome Build...	Opens the Import Genome Build dialog box, where you can import Agilent-provided genome build files. See "Import Genome Build" on page 166 and "To import a genome build" on page 21.
Probe Upload	Allows you to import a file of probe sequences and annotation. For details, see the *eArray XD User Guide.* You see this only if you have installed Genomic Workbench Enterprise Edition.
Track...	Opens the Import Track dialog box, where you can select a BED format track file for import, and create a display name for the track. See "Import Track" on page 167 and "To import tracks" on page 22.
Experiments	Opens the Import Experiments dialog box, where you can select a ZIP format experiment file for import. See "Import" on page 161 and "To import an experiment file" on page 24.

Export Opens a menu that allows you to export several kinds of files.

Option	Description
Experiments...	Opens the Export Experiments dialog box, where you can select one or more experiments for export as a single ZIP file. See "Export Experiments" on page 155 and "To export experiments" on page 43.
Tracks...	Opens the Export Tracks dialog box, where you can select one or more tracks to export as a single BED format file. See "Export Tracks" on page 156 and "To export tracks" on page 44.
Array Attributes	Opens the Export Array Attributes dialog box, where you can select arrays and their attributes for export. See "Export Array Attributes – Array" on page 151.

Create Experiment... Opens the Create Experiment dialog box, where you can create a new, empty experiment and populate it with data. See "Choose Gene List Color" on page 132 and "To create a new experiment" on page 26.

Go to Gene/Genomic Location Moves the cursor to the location in Chromosome and Gene Views that you specify. See "Go To Gene/Genomic Location" on page 160.

Print Opens the Print window to print the display.

Exit Closes the program.

View command ribbon

The View command ribbon lets you change the display of data/results in Genomic Viewer.

Figure 5 View command ribbon

Orientation Select one of these options:

- Horizontal – Stacks Genome, Chromosome, and Gene views horizontally. Chromosomes and chromosomal locations appear in left to right orientation.
- Vertical – Stacks Genome, Chromosome, and Gene views vertically. Chromosomes and chromosomal locations appear in top to bottom orientation.

Copy This command opens a menu with the options listed below. In general, the Copy command copies pane(s) of the main window to the Clipboard as an image. You can then paste the image into a document in another program. See "To copy what you see in the main window" on page 45.

Option	Description
All	Copies all panes of the main window to the Clipboard as an image.
Navigator	Isolates and copies the Navigator to the Clipboard as an image.
Tab View	Isolates and copies Tab View to the Clipboard as an image.
Genome view	Isolates and copies Genome View to the Clipboard as an image.
Chromosome view	Isolates and copies Chromosome View to the Clipboard as an image.
Gene view	Isolates and copies Gene View to the Clipboard as an image.

Show Opens a menu with all available elements of the main window. Mark the check box for the one or ones you want to display in Genomic Viewer.

Scatter Plot See "Scatter Plot" on page 119.

Preprocessing command ribbon

The Preprocessing ribbon contains only one command you use to prepare for CH3 analysis: Fuse.

Figure 6 CH3 Preprocessing command

The Fuse command enables you to combine multiple array designs into a larger virtual combined design. They also allow you to combine both intra- and inter-array replicate probes.

Click to open the Array Set dialog box, where you can view the names of arrays to be fused, set options, and fuse designs together to form a larger virtual design. See "Array Set" on page 131, and "To combine (fuse) arrays" on page 71.

Analysis command ribbon

The Analysis ribbon contains three commands you use for CH3 analysis.

Figure 7 CH3 Analysis ribbon

Moving Average (Log Ratio)

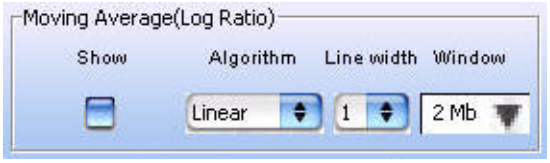


Figure 8 Moving Average (Log Ratio) command

The settings under Moving Average(Log Ratio) control the calculation and display of line plots that represent smoothed log ratio data for each selected array. These plots can appear in Genome, Chromosome, and Gene views.

Show Mark this check box to show moving average line plots for the log ratio data, or clear the check box to hide them. You enable or disable the display of moving average line plots for specific views in the User Preferences — In the **View** tab, under **Data Visibility,** select the desired view, then mark or clear **Moving Average.** See "Preferences – View Tab" on page 176.

Algorithm Select one of these options:

- **Linear** – The linear algorithm calculates a standard, unweighted average using every Log Ratio score within a specified window size (specified by either a number of adjacent measurements or a positional range). This average is applied sequentially to every point by moving the window along the chromosome.

- **Triangular** – The triangular algorithm calculates a weighted average using every Log Ratio within a specified window size. The triangular algorithm is more sensitive to localized variations in the data. See "Methylation Detection and Visualization Algorithms Overview" on page 198.

Line width Select the desired thickness (in pixels) for the moving average line plots. You can select a thickness from 1 to 5 pixels.

Window Select the desired size of the moving average window. You can select either a specific number of base pairs (5 Kb to 50 Mb), or a specific number of data points (1 pt to 60 pt). You can also type a value. The program calculates a moving average for each selected array based on a window of the given size centered on each point in the array.

Probe Methylation

You apply this command to make Z-score calculations to assess whether a Methylation detection array probe is methylated or unmethylated. See "Methylation Detection and Visualization Algorithms Overview" on page 198.

Figure 9 Probe Methylation command

Click **Apply** to bring up the Probe Methylation Status Setup dialog box to assign a melting temperature (Tm) mapping file, if necessary, to the

experiment. See “Probe Methylation Status Setup” **on page 178**.**Moving Average (ZScore)**

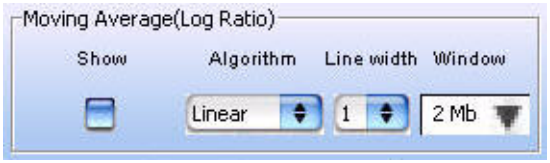


Figure 10 Moving Average (ZScore) command

The settings under Moving Average(ZScore) control the calculation and display of line plots that represent smoothed Z-score data for each selected array. These plots can appear in Genome, Chromosome, and Gene views. You must apply the Probe Methylation algorithm first.

Show Mark this check box to show moving average line plots for the Z-score data, or clear the check box to hide them. You enable or disable the display of moving average line plots for specific views in the User Preferences — In the **View** tab, under **Data Visibility,** select the desired view, then mark or clear **Moving Average.** See “Preferences – View Tab” on page 176.

Algorithm Select one of these options:

- **Linear** – The linear algorithm calculates a standard, unweighted average using every combined Z-score within a specified window size (specified by either a number of adjacent measurements or a positional range). This average is applied sequentially to every point by moving the window along the chromosome.
- **Triangular** – The triangular algorithm calculates a weighted average using every Z-score within a specified window size. The triangular algorithm is more sensitive to localized variations in the data. See “Triangular Smoothing” on page 209.

Line width Select the desired thickness (in pixels) for the moving average line plots. You can select a thickness from 1 to 5 pixels.

Window Select the desired size of the moving average window. You can select either a specific number of base pairs (5 Kb to 50 Mb), or a specific number of data points (1 pt to 60 pt). You can also type a value. The program calculates a moving average for each selected array based on a window of the given size centered on each point in the array.

Reports command ribbon

Figure 11 CH3 Report command

The Generate Report command enables you to save all the important information about the Z-score statistics and CpG Island tracks that can help you assess whether regions of the genome are methylated or not. You use a spreadsheet program or your Internet browser to open the report files the program generates. This allows you to further view, analyze, and organize the result.

Opens the Probe Methylation Report Dialog dialog box, where you can define a location and a name for the probe report file and select whether you want individual reports for each chromosome or one report on the complete genome. See "Probe Methylation Report Dialog" on page 179.

The Methylation Report makes information about the results of the probe methylation algorithm available in .xml format (complete genome) or in .txt format (individual chromosomes). You can use a spreadsheet program to open all these file.s See "Report Format" on page 195 for a description of the columns in the report.

Help command ribbon

The Help command ribbon provides access to the Agilent Genomic Workbench Quick Start Guide and all the other user guides.

Figure 12 Help command ribbon for Genomic Workbench Standard Edition

Table 1 Table of Agilent Genomic Workbench Help

Help Command	Action
Help	Opens the DNA Analytics application user guide for which you have the associated license. Methylation (CH3) Analysis --(CH3 license required) Presents instructions for preprocessing, analysis and reporting options for assessing if segments of a chromosome are methylated or not. Includes details on the algorithms.
Quick Start	Opens the *Genomic Workbench Quick Start Guide* in Adobe Reader. This guide provides brief instructions on how to install and start the program, and how to use the basic features of the program to create custom microarray designs, and to analyze microarray data.
Data View	Shows you how to import, manage, export and display log ratio data from Agilent and other sources
Workflow	Gives instructions on how to set up a workflow for automated feature extraction and/or analysis. Also shows you how to set up the CGH and ChIP analysis methods to be used in a workflow. Note: Workflow is not used with CH3 analysis.
Sample Manager	Shows you how to assign identification and attribute information to image files, imported feature extraction (FE) files or UDF files.

Table 1 Table of Agilent Genomic Workbench Help

Help Command	Action
eArray	(Genomic Workbench Enterprise Edition must be installed) Gives instructions on how to design your own microarrays on your desktop, not the web site.
About	Opens a message with details about the version number and copyright of the program.

Navigator

This section describes the parts of the Navigator, and the shortcut menus and other functionality available within it.

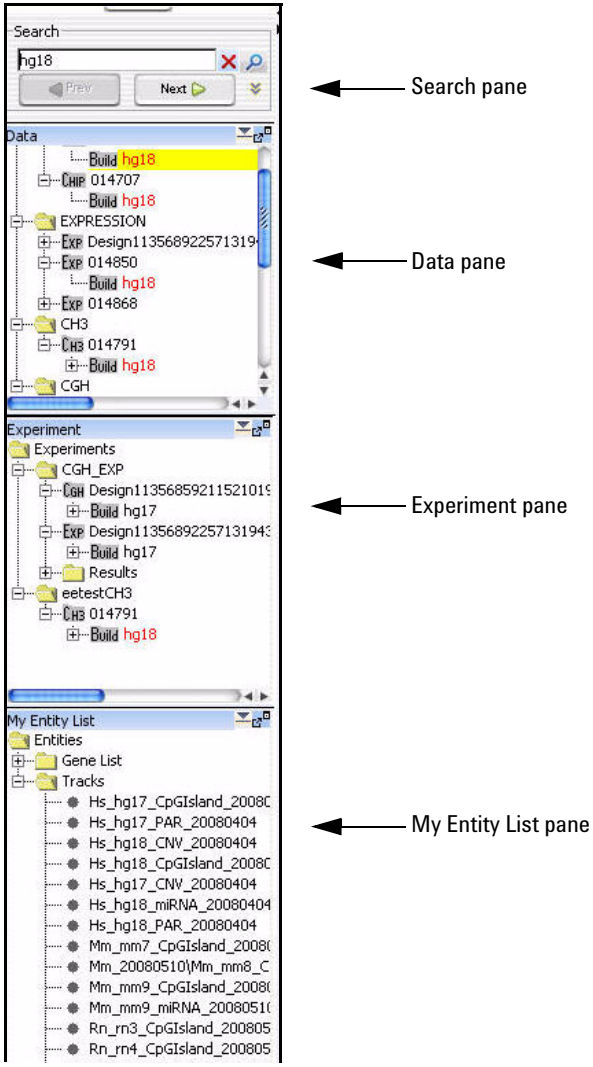


Figure 13 Navigator – note the four panes within the Navigator

The Navigator (Figure 13) catalogs the array data, experiments, and other content stored in Genomic Workbench. It contains four panes:

Pane	Comments
Search	Allows you to find all occurrences of a specific search term in the Data, Experiment, and/or My Entity List panes.
Data	Contains microarray data files, organized by application type and design, and then by genome build.
Experiment	Contains Genomic Workbench experiments. Experiments are organizational units within the program that contain links to microarray data and design files. In data analysis modules, experiments also contain saved results.
My Entity List	Contains gene lists and tracks: • **Gene Lists** are collections of genes of interest. You can create them within the program, import and export them, and apply them to Gene View and Chromosome View. • **Tracks** are collections of annotation or other information that is correlated with specific genomic locations. You can import, export, and combine tracks, and display them in Gene View alongside your array data and analysis results.

Search pane

The Search pane allows you to find all occurrences of a specific search term in the Data, Experiment, and/or My Entity List panes. See "To find specific content items in the Navigator" on page 33. It also contains several buttons that you can use to detach, hide, show or resize the Navigator.

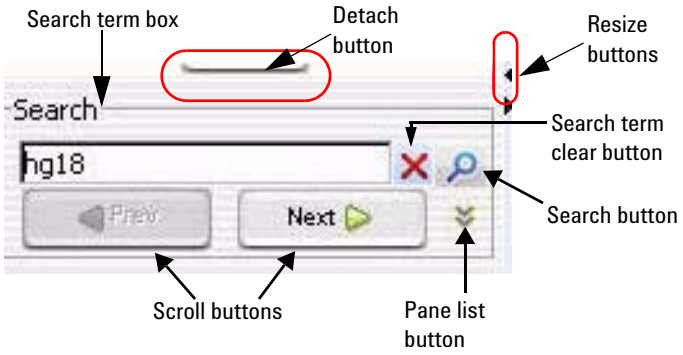


Figure 14 Navigator – Search pane

Detach button Click to detach the Navigator from the main window of the program and open it in a new, separate window.

Resize buttons Click to hide, show, or expand the Navigator.

Search term box Provides a box for you to type your desired search term. Search terms are not case-sensitive, but they must reflect the entire name of an array or other content item that you want to find. You can use asterisks (*) as wildcards to represent groups of unspecified characters.

Pane list Allows you to restrict a search to a specific pane. Select the name of the desired pane from the list. To select all panes, select **All Panels.** By default, the program searches all panes.

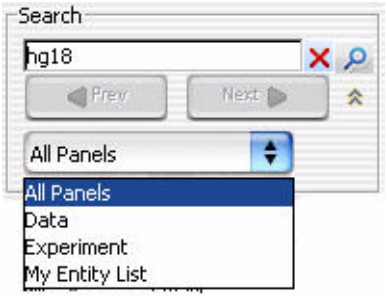


Figure 15 Open Pane List

(Show Pane List button, available only if the Pane list is not visible) Makes the Pane list visible.

(Hide Pane List button, available only if the Pane list is visible) Hides the Pane list.

(Search button) Searches the pane(s) selected in the Pane list for all occurrences of the term you typed in the Search term box. If the program finds a matching item, it expands the folder structure to make the matching item(s) visible, makes the lettering of each item red and highlights the item in yellow. Note: The search term is not case-sensitive, but it must reflect the entire name of the desired items.

Scroll buttons (Available only after a search) Lets you scroll up and down the lists of highlighted search items after a search.

(Clear button, available only after a search) Clears the search term from the Search term box, and resets the color of any matching item to its original color.

Data pane – icons, special text, and buttons

Item	Comments
	Click to expand a folder and display its contents.
	Click to collapse a folder and hide its contents.
	A folder that contains files or other folders.
	A methylation array design. This folder contains array data associated with the design, organized by genome build.
	A CGH array design. This folder contains array data associated with the design, organized by genome build.
	A gene expression array design. This folder contains array data associated with the design, organized by genome build.
	A ChIP array design. This folder contains array data associated with the design, organized by genome build.
	A genome build folder within a specific design folder. This folder contains arrays associated with the specific genome build and design.

Item	Comments
	A single array data file.
	Data created from a multi-pack array.
red text	An item that matches the search term in a search.
	(Dock out button) Detaches the Data pane from the Navigator, and opens it in its own, separate window.
	(Collapse button, available only if the Data pane is not collapsed) Collapses the Data pane, and shows its title bar at the bottom of the Navigator.
	(Expand button, available only if the Data pane is collapsed) Expands the Data pane.

Data pane – actions and shortcut menus

- Double-click any folder to expand or collapse it.

Data Folder

- Right-click the **Data** folder to open a shortcut menu with an Import option. When you select this option, a menu appears with these options for file import:

Option	Description
Design File...	Opens the Import Design Files dialog box, where you can select an Agilent GEML-based (*,xml) file for import. See "Import" on page 161 and "To import Agilent GEML design files" on page 19.
Axon Design File...	Opens the Import Axon Files dialog box, where you can select GenePix/Axon design (*.gal) files for import. See "Import" on page 161 and "To import GenePix/Axon design files" on page 20.
FE File...	Opens the Import FE Files dialog box, where you can select an Agilent Feature Extraction array data file to import. See "Import" on page 161 and "To import Agilent FE or GenePix/Axon data files" on page 14.

Option	Description
Axon File...	Opens the Import Axon Files dialog box, where you can select Axon (*.gpr) files for import. See "Import" on page 161 and "To import Agilent FE or GenePix/Axon data files" on page 14.
UDF Files...	Opens the UDF Files dialog box, where you can select a Universal Data File (UDF) to import. See "Import" on page 161 and "To import a UDF file" on page 16.

Design Folder

- Right-click the name of design folder to open a shortcut menu with these options:

Option	Description
Update from eArray	(Available only for Agilent microarrays) Updates the annotations for your array design from the eArray Web site. Agilent regularly updates annotations in eArray as new ones become available. See "To update probe annotation in design files" on page 34.
Delete	Opens a Confirm dialog box. If you click **Yes,** the program permanently deletes the design and all arrays associated with it.

Genome Build Folder

- Right-click the name of a genome build folder to open a shortcut menu with these options:

Option	Description
Show Properties...	Opens the Design Properties dialog box. See "Design Properties" on page 141.
Delete	Opens a Confirm dialog box. If you click **Yes,** the program permanently deletes all of the arrays in this genome build folder.

Specific Arrays

- Right-click the name of an array to open a shortcut menu with these options:

Option	Description
Show Properties	Opens the Microarray Properties dialog box. See"Microarray Properties - Attribute Tab" on page 168 and "To edit the attribute values of a specific array" on page 31.
Rename...	Opens an Input dialog box, where you can type a new name for the array. Click **OK** to rename the array.
Delete	Opens a Confirm dialog box. If you click **Yes,** the program permanently deletes the array.

- Drag an array from the Data pane to an experiment folder in the Experiment pane to associate it with an experiment. You can drag multiple arrays at once from one genome build in a design. Control-click the additional arrays to select them. You can also select a contiguous block of arrays—click the first array in the block, then shift-click the last one.

Experiment pane – icons, special text, and buttons

Item	Comments
⊞	Click to expand a folder and display its contents.
⊟	Click to collapse a folder and hide its contents.
	A folder that contains files or other folders.
CH3	A methylation array design. This folder contains array data associated with the design, organized by genome build.
CGH	A CGH array design. This folder contains array data associated with the design, organized by genome build.
EXP	A gene expression array design. This folder contains array data associated with the design, organized by genome build.

Item	Comments
	A ChIP array design. This folder contains array data associated with the design, organized by genome build.
	A genome build folder within a specific design folder. This folder contains arrays associated with the specific genome build and design.
	An array that is not selected for view and analysis.
	An array that is selected for view and analysis. The specific color of this icon can vary.
	An array that has been designated as a calibration array.
	An empty folder.
	Data created from a multi-pack array.
blue text	The currently active experiment. All data and results that appear in Chromosome, Gene, and Tab Views are derived from this experiment.
red text	An item that matches the search term in a search.
	(Dock out button) Detaches the Experiment pane from the main window, and opens it in its own, separate window.
	(Collapse button, available only if the Experiment pane is not collapsed) Collapses the Experiment pane, and shows its title bar at the bottom of the Navigator.
	(Expand button, available only if the Experiment pane is collapsed) Expands the Experiment pane.

Experiment pane — actions and shortcut menus

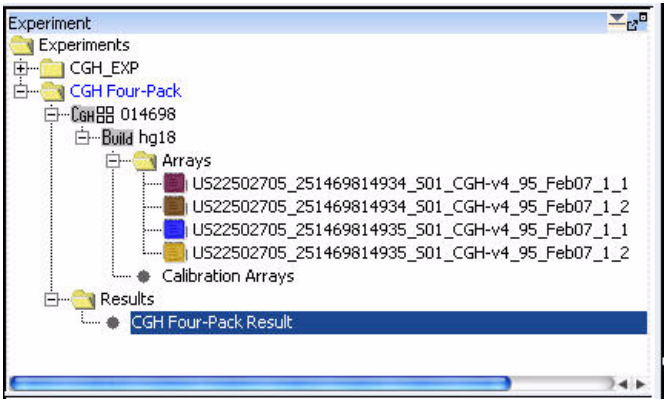


Figure 16 Expanded Experiment Pane

You can use many items in the Experiment Pane of the Navigator to open shortcut menus or take other actions.

- In general, double-click the **Experiments** folder within the Experiment Pane, and the folders within it, to expand and collapse them. Exception: Double-click the name of an unselected experiment to select it for analysis. This action opens the Experiment Selection dialog box. To select the experiment for analysis, click **Yes.**

Experiments Folder

- Right-click the **Experiments** folder to open a shortcut menu with these options:

Option	Description
New Experiment	Opens the Create Experiment dialog box (see "Create Experiment" on page 136), where you can name the new experiment, and open another dialog box that allows you to populate the experiment with microarrays. See "To create a new experiment" on page 26.
Export	Opens the Export Experiments dialog box, where you can export one or more experiments as a single ZIP file. See "Export Experiments" on page 155 and "To export experiments" on page 43.

Specific Experiment Folder

- In the **Experiments** folder, right-click the name of a specific experiment folder to open a shortcut menu with these options:

Option	Description
Select Experiment	(Appears only if the experiment is not selected.) Opens the Experiment Selection dialog box, which asks if you want to select the experiment. Click **Yes** to select the experiment for display and analysis. Or In the Experiments folder, double-click the name of an experiment that is not selected to open the Experiment Selection dialog box. To select the experiment for analysis, click **Yes.** If you switch experiments, a Confirm dialog box asks if you want to save the current result. Select one of these options: • **Overwrite Current Result** – Replaces the currently selected experiment result in the Data Navigator with the result that appears on your screen. • **Create New Result** – Opens the Save experiment result dialog box, where you can save the result that appears in the main window as a new experiment result. See "To save a result" on page 78. • **Continue Without Saving** – The program does not save the result that appears on your screen. In some cases, the Confirm dialog box offers only Yes and No choices. If you click **Yes,** the Save experiment result dialog box appears, where you can save the result that appears on your screen with the name of your choice.
Deselect Experiment	(Appears only if the experiment is selected.) If there are unsaved results, a Confirm dialog box opens with these options: • **Overwrite Current Result** – Replaces the currently selected experiment result in the Data Navigator with the result that appears on your screen. • **Create New Result** – Opens the Save Experiment result dialog box, where you can save the result that appears on your screen as a new experiment result. • **Continue Without Saving** – The program does not save the result that appears on your screen. In all three cases, the program then removes the experiment data and results from all views.

Option	Description
Show Properties	Opens the Experiment Properties dialog box. Use this dialog box to see the names of the arrays in the experiment, and also to add or remove arrays from the experiment. See "Experiment Properties" on page 147.
Edit Attributes	Opens the Experiment Attributes dialog box, where you can change the values for the attributes assigned to the arrays in the experiment. See "Experiment Attributes" on page 146. To add new attributes you must use the Sample Manager. See *Sample Manager User Guide*.
Export	Opens the Export Experiments dialog box, where you can export this and other experiments as a single ZIP file. See "Export Experiments" on page 155, and "To export experiments" on page 43.
Export Attributes	Opens the Export Array Attributes dialog box, where you can save a file that contains selected attributes of the arrays in your experiment. See "Export Array Attributes – Array" on page 151.
Edit Array Color	Opens the Edit Array Color dialog box, where you can select a display color for each of the arrays in the experiment. See "Edit Array Color" on page 144.
Edit Array Order	Opens the Edit Array Order dialog box, where you can change the order of the arrays in the experiment pane of the Navigator, and in Chromosome, Gene, and Tab Views. See "Edit Array Order" on page 145.
Rename	Opens an Input dialog box, where you can type a new name for the experiment. Click **OK** to rename the experiment.
Delete	Opens a Confirm dialog box that asks if you want to delete the Experiment. Click **Yes** to delete it. Note: You can delete any experiment except the selected one.

Design Folder

- In the folder of a specific experiment, right-click the name of a design to open a shortcut menu with a Delete command. If you click **Delete,** a Confirm dialog box opens. Click **Yes** to disassociate all of the arrays under the design from the experiment.

Genome Build Folder

- In the folder of a specific experiment, in a specific design folder, right-click the name of a genome build to open a shortcut menu with these options:

Option	Description
Set for Calibration	Designates all arrays associated with the specific genome build and design as calibration arrays. The program transfers their names to the Calibration Arrays folder of the experiment, and moves the array data in Tab View for those arrays to the Calibration Arrays tab.
Save As Text File	Opens the Save Design dialog box, where you can save all of the data associated with the genome build and design within the experiment as a tab-delimited text file.
Delete	Opens a Confirm dialog box that asks if you want to disassociate all arrays under the design from the experiment. Click **Yes** to remove the links between the arrays and the experiment. Note: • If you delete a design from an experiment, the program removes the links between the experiment and the design and its arrays. The actual design and array data remain in the Data folder. • Saved results become unavailable if they involve arrays you delete with this command.

Individual Arrays

- In a specific experiment, right-click the name of an individual array to open a shortcut menu with these options:

Option	Description
Select	(Available only if the array is not already selected) Selects the array for display and analysis.
Deselect	(Available only if the array is selected) Removes the array data from Genome, Chromosome, and Gene views, and excludes it from the analysis. Also removes the array data from the Selected Arrays tab in Tab View.

Option	Description
Select for Calibration	(Available only in the **Arrays** folder.) Designates the array as a calibration array. Within the specific experiment, the program moves the name of the array to the Calibration Arrays folder. Within the specific design tab in Tab View, the program also moves the data for the array from the Arrays tab to the Calibration Arrays tab.
Deselect from Calibration	(Available only within the **Calibration Arrays** folder.) Designates the array as a "regular" non-calibration array. Within the specific experiment, the program moves the name of the array to the Arrays folder. Within the specific design tab in Tab View, the program also moves the data for the array from the Calibration Arrays tab to the Arrays tab.
Rename	Opens an Input dialog box, where you can type a new name for the array. Click **OK** to accept the new name for the array.
Delete	Opens a Confirm dialog box that asks if you want to disassociate the array from the experiment. Click **Yes** to remove the link between the array and the experiment. **Note:** • If you delete an array from an experiment, the program removes the link between the experiment and the array. The actual array data remains in the Data folder. • You cannot restore an experiment result that includes a deleted array.
Show Properties	Opens the Microarray Properties dialog box, where you can view and edit microarray attributes. See "Microarray Properties - Attribute Tab" on page 168. For array files from the Agilent Feature Extraction program, you can also view the headers and feature data from the file. See "To edit the attribute values of a specific array" on page 31.
Edit Array Color	Opens the Select Color dialog box, where you can select a display color for the array. See "Search probes in eArray" on page 181.
Edit Array Order	Opens the Array Order dialog box, where you can change the order of the arrays in the Experiment pane of the Navigator, and in Chromosome, Gene, and Tab Views. See "Edit Array Order" on page 145.

My Entity List pane – Icons, buttons, and special text

Item	Comments
	Click to expand a folder and display its contents.
	Click to collapse a folder and hide its contents.
	A folder that contains files or other folders.
	An individual gene list or track.
red regular text	An item that is an exact match with the search term in a search, or an unapplied gene list that has red chosen as its custom color.
colored italics	A gene list that has been applied.
red bold italics	A track that is selected for display in Gene View.
black bold italics	A "combined" track that is selected for display in Gene View. A combined track contains information from two or more individual tracks associated by logical criteria.
	(Dock out button) Detaches the My Entity List pane from the main window, and opens it in its own, separate window.
	(Collapse button, available only if the My Entity List pane is not collapsed) Collapses the My Entity List pane, and shows its title bar at the bottom of the Navigator.
	(Expand button, available only if the My Entity List pane is collapsed) Expands the My Entity List pane.

My Entity List pane – Actions and shortcut menus

- Right-click the **Gene List** folder to open a shortcut menu with an **Import Gene List** option. This command opens an Import dialog box that you can use to import a gene list into the program. See "To import a gene list" on page 24 and "Import" on page 161.
- Double-click the **Gene List** folder to show or hide its gene lists.
- In the **Gene List** folder, right-click the name of a gene list to open a shortcut menu with these options:

Option	Description
View in Table	Opens the Gene List dialog box, where you can view the list of genes. You can also edit the description of the gene list, and change the display color of the genes. See "Gene List" on page 159 and "To view the genes in a gene list" on page 38.
Rename...	Opens an Input dialog box, where you can type a new name for the gene list. Click **OK** to accept the new name.
Delete	Opens a confirm dialog box that asks if you are sure you want to delete the gene list. Click **Yes** to confirm.
Save As	Opens a Save As dialog box, where you can save the gene list as a text (*.txt) file. See "To export a gene list" on page 44.
Add to gene list	Opens the Add gene list dialog box, where you can add the gene list to any other one in the Gene List folder. See "Add Gene List <name> to" on page 128 and "To add one gene list to another" on page 38.
Highlight	(Available if the gene list is not selected.) Displays all genes in Gene View, and highlights the genes from the gene list in their display color. See "To show gene lists in Gene View" on page 58.
Show only	(Available only if all genes appear in Gene View, or if the gene list is not selected) Restricts the genes in Gene View to those on the gene list. No other genes appear. The program displays the genes in their display color. See "To show gene lists in Gene View" on page 58.
Show All	(Available only for the selected gene list.) In Gene View, displays all genes, without highlighting. See "To show gene lists in Gene View" on page 58.

- Right-click the Tracks folder to open a shortcut menu with these options:

Option	Comments
Import Tracks	Opens the Import Track dialog box, where you can import a BED format track file into the program. See "Import Track" on page 167 and "To import tracks" on page 22.

Option	Comments
Export Tracks	Opens the Export Tracks dialog box, where you can select tracks for export as a single BED format track file. See "Export Tracks" on page 156 and "To export tracks" on page 44.
Combine Tracks	Opens the Combine Tracks dialog box, where you can associate two or more individual tracks by logical criteria to create a new combined track. See "Combine Tracks" on page 133 and "To combine tracks" on page 39.

- Right-click the name of a track to open a shortcut menu with these options:

Option	Comments
Show in UI	Mark this option to display the track in Gene View alongside the data and results of the selected experiment. See "To show tracks in Gene View" on page 70.
Show in Report	Mark this option to display the track in the reports.
Genomic Boundaries	Click to analyze data invoking the genomic boundaries for only that track. You can choose to do this for only one track.
Show in UCSC	Opens the UCSC Genome Browser in your Web browser and uploads the track. You can then view the track.
View Details	Opens a dialog box that displays information about the track. See "Track" on page 189.
Rename	Opens an Input dialog box, where you can type a new name for the track. Click **OK** to rename the track.
Delete	Opens a Delete Track dialog box that asks if you are sure you want to delete the track. Click **Yes** to delete the track.

Genomic Viewer

Genomic Viewer is the display for the DNA Analytics applications. It includes the three Views – Genome, Chromosome and Gene Views – the Tab View and the Cursor.

Genome View

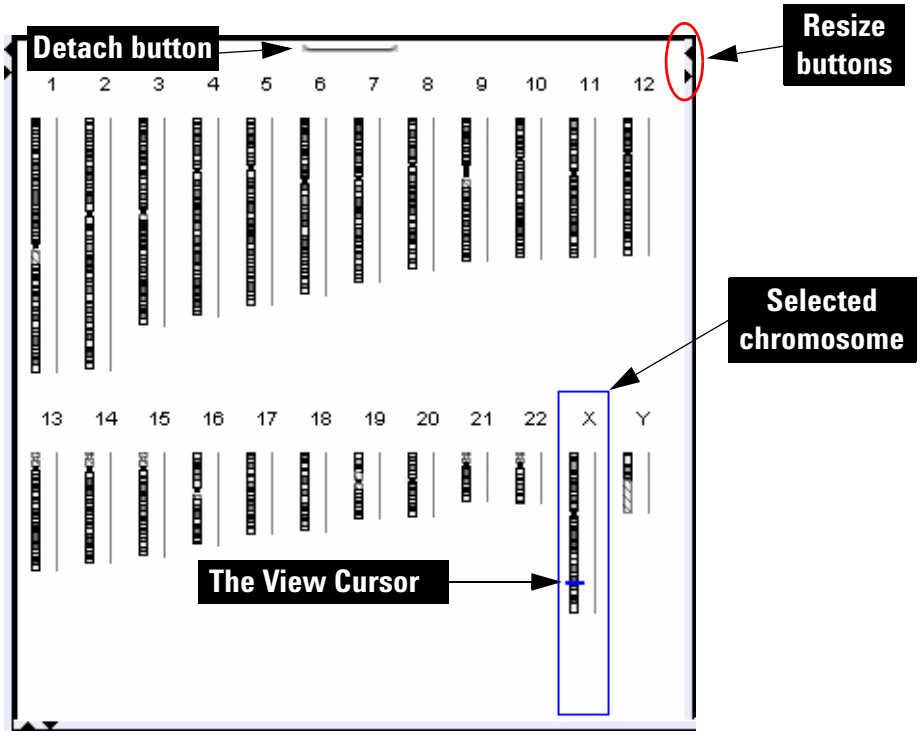


Figure 17 Genome View, with human chromosomes. The X chromosome is selected.

Genome View shows pictures of each of the distinct types of chromosomes in the selected genome. A blue box encloses the currently selected chromosome, and the View cursor appears as a blue line across the chromosome.

Genome View actions and shortcut menus

- Click a chromosome to select it. When you select a chromosome, Chromosome, Gene, and Tab Views show only genomic regions, genes, and data associated with it. The specific location in which you click the chromosome sets the position of the cursor. See "The View Cursor" on page 122.

- On the selected chromosome, click anywhere to reposition the cursor. See "The View Cursor" on page 122. This also repositions the cursor in Chromosome, Gene, and Tab Views.
- Right-click anywhere within Genome View to open a shortcut menu with a Preferences option. If you click **Preferences,** the Preferences dialog box opens, where you can set user preferences on four separate tabs. See "Preferences – License" on page 171 and the topics that follow.
- Click the **Detach** button (located at the top center of the pane) to remove Genome View from the main window and open in its own separate window. To reattach the view, click its **Close** button  . See "Resizing and detaching panes from the Genomic Workbench main window" on page 19 and "Maximizing and reattaching panes to the Genomic Workbench main window" on page 20 of the *Quick Start Guide*.
- Drag the side or bottom borders of the pane to resize it.
- On a border of the pane, click a resize button (for example,  or ) that points away from the pane to move that border all the way to the edge of the main window. To move the border back to its previous location, click the other resize button. See "Resizing and detaching panes from the Genomic Workbench main window" on page 19 of the *Quick Start Guide*.

Chromosome View

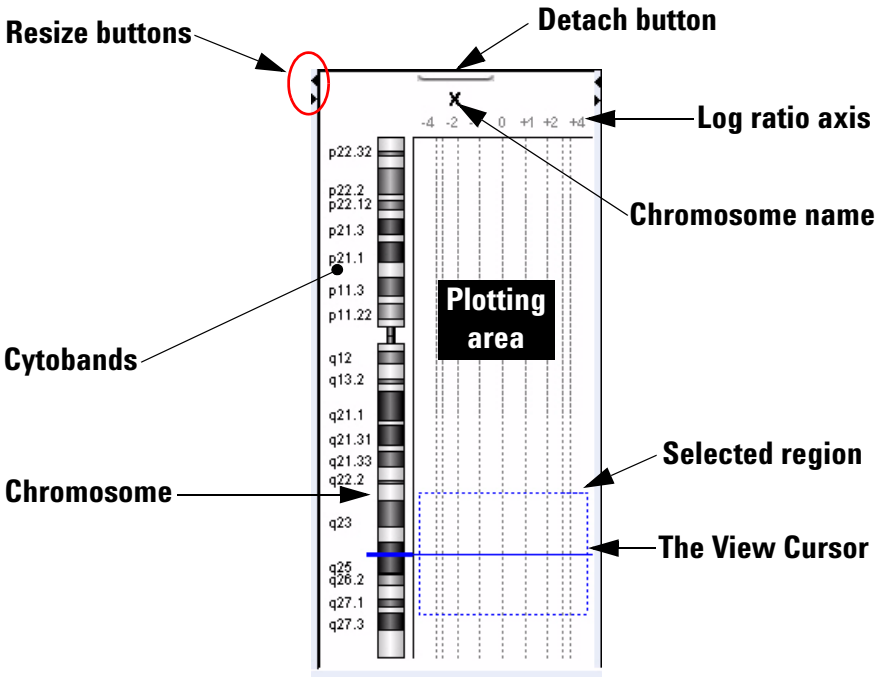


Figure 18 Chromosome View, human X chromosome shown

Chromosome View shows a more detailed diagram of the chromosome you select in Genome View.

- Cytobands and a plotting area appear alongside the chromosome.
- When you select arrays for display, their data appear in the plotting area.
- The cursor appears as a solid blue line across the chromosome and the plotting area.
- The selected region of the chromosome (if any) appears as a dotted blue box in the plotting area.

Chromosome View actions and shortcut menus

- Click a cytoband, any part of the chromosome, or anywhere in the plotting area to reposition the View cursor at that location. See "The View Cursor" on page 122.
- Drag the pointer over any part of the plotting area to select a chromosomal region for display in Gene View. Drag parallel to the chromosome. This also repositions the cursor to the center of the selected region. See "The View Cursor" on page 122.
- Right-click anywhere in the plotting area of Chromosome View to open a shortcut menu with a Preferences option. If you click **Preferences,** the Preferences dialog box opens, where you can set user preferences on four separate tabs. See "Preferences – License" on page 171 and the topics that follow.
- Click the **Detach** button (located at the top center of the pane) to remove Chromosome View from the main window and open in its own separate window. To reattach the view, click its **Close** button  . See "Resizing and detaching panes from the Genomic Workbench main window" on page 19 and "Maximizing and reattaching panes to the Genomic Workbench main window" on page 20 of the *Quick Start Guide*.
- Drag the side or bottom borders of the pane to resize it.
- On a border of the pane, click a resize button (for example,  or ) that points away from the pane to move that border all the way to the edge of the main window. To move the border back to its previous location, click the other resize button. See "Resizing and detaching panes from the Genomic Workbench main window" on page 19 of the *Quick Start Guide*.

Gene View

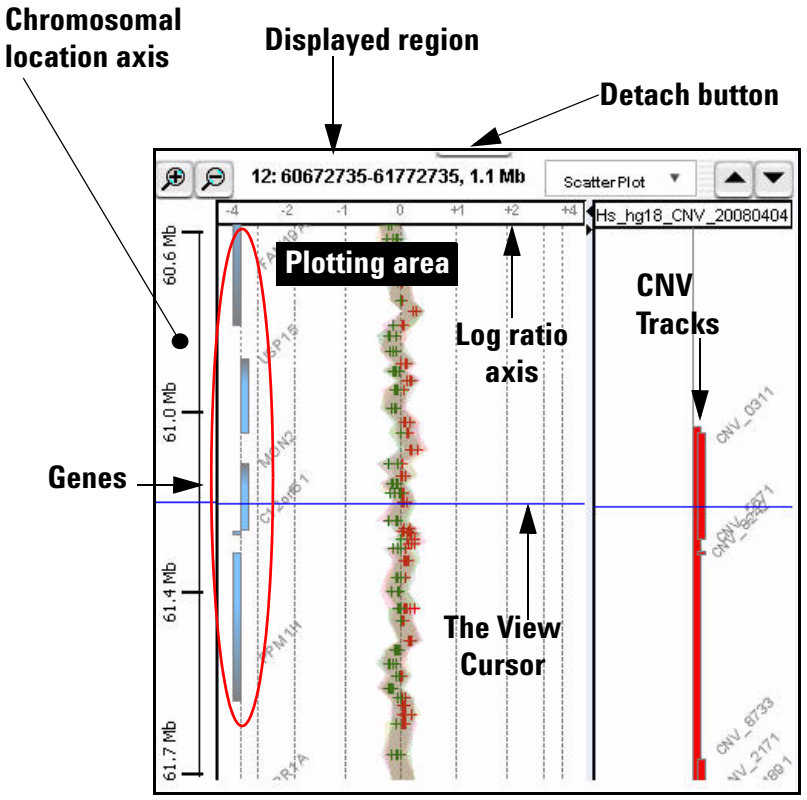


Figure 19 Gene View, with log ratio data from an experiment and CNV tracks.

Gene View shows a more detailed view of the chromosomal region you select in Chromosome View. See "Chromosome View" on page 116.

- Regions occupied by genes appear as small blue boxes. Gene names appear nearby. You can customize the appearance of gene names. Also, you can use a gene list to highlight genes of interest, or to restrict the genes that appear to those in the list. See "To change the appearance of genes in Gene View" on page 60, and "To show gene lists in Gene View" on page 58.

- Log ratio data from selected arrays in the active experiment appear as a scatter plot. Points appear in up to three different colors. See "To show significant data points in a scatter plot" on page 53. You can also customize the scatter plot. See "To change scatter plot appearance" on page 54.
- The location of the cursor matches the location of the cursors in other views. See "The View Cursor" on page 122.
- The name of the chromosome, and the coordinates and size of the displayed chromosomal region appear at the top of the pane.
- Imported tracks can also appear in Gene View. See "To show tracks in Gene View" on page 60.

Scatter Plot

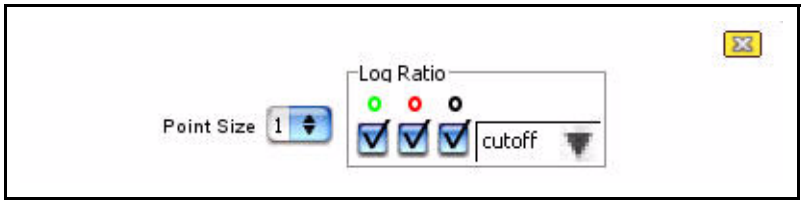


Figure 20 Scatter Plot command group in CH3 Gene View

You access the scatter plot command group through Gene View or through the View tab. The commands differ depending on the DNA Analytics application you are working with. All the scatter plot command groups contain the commands for log ratio data. Scatter plots appear in the Chromosome and Gene Views but only if they have been activated in the Preferences dialog box.

Cutoff Select the threshold of log ratios to give a visual clue about significance. In the above example, all data points whose log ratios are below the cutoff of +0.125 are significant. If you select **none,** the program classifies all points as significant.

Mark this red check box to display, or clear it to hide, points with log ratios above the Cutoff selected. You select the cutoff for significance in **Cutoff.** These points appear in red.

Mark this green check box to display, or clear it to hide, points that reflect significant decreases. You select the cutoff for significance in **Cutoff.** These points appear in green.

Mark this black check box to display, or clear it to hide, points that reflect insignificant changes. You select the cutoff for significance in **Cutoff.** These point appear in black.

Point size Sets the size (in pixels) of the points in the scatter plots.

In addition to options for log ratio plotting, the CGH command group includes options for raw intensity plotting.

Gene View buttons

Zooms in to see a smaller region in more detail.

Zooms out to see a larger region in less detail.

(Available when Gene View is in vertical orientation.) Scrolls up through the genes and data to lower-numbered chromosomal coordinates.

(Available when Gene View is in vertical orientation.) Scrolls down through the genes and data to higher-numbered chromosomal coordinates.

(Available when Gene View is in horizontal orientation.) Scrolls left through the genes and data to lower-numbered chromosomal coordinates.

(Available when Gene View is in horizontal orientation.) Scrolls right through the genes and data to higher-numbered chromosomal coordinates.

(**Resize** buttons) The button that points away from Gene View expands the view. The other button restores the view to its original size. See "Resizing and detaching panes from the Genomic Workbench main window" on page 19 of the *Quick Start Guide*.

(**Detach** button) Removes Gene View from the main window, and opens it in its own separate window. "Resizing and detaching panes from the Genomic Workbench main window" on page 19 of the *Quick Start Guide*.

Gene View shortcut menu and other actions

- Click anywhere in the plotting area of Gene View to move the cursor to that location. See "The View Cursor" on page 122.
- Drag an inside border of Gene View to resize the view.

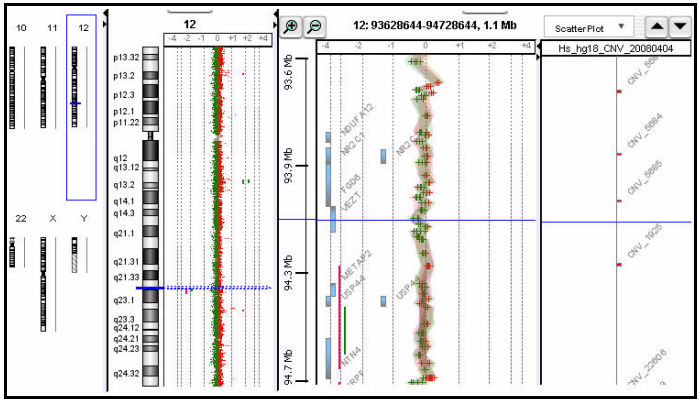
- Right-click anywhere in the plotting area of Gene View to open a shortcut menu with these options:

Option	Description
Create Gene List...	Opens the Create Gene List dialog box, where you can create a new gene list based on the currently selected (or another) chromosomal region. See "Create Gene List" on page 138 and "To create a gene list" on page 37.
Show in UCSC	Opens the UCSC (University of California at Santa Cruz) Genome Browser in your Web browser and uploads the track. You can then view the track.
Search probes in eArray...	Opens the Search probes in eArray dialog box, where you can start a search of the Agilent eArray web site for probes in the selected (or another) chromosomal region. See "Search probes in eArray" on page 181 and "To search Agilent eArray for probe information" on page 66.
Preferences	Opens the Preferences dialog box, where you can set user preferences on four separate tabs. See "Preferences – License" on page 171 and the related pages that follow.

The View Cursor

The View cursor reflects the center of the current chromosomal location of interest. It appears in several Views:

- In Genome View, it appears as a blue bar across the selected chromosome.
- In Chromosome View, it is a blue bar that appears across the chromosome and across the plotting area of the view.
- In Gene View, it is a blue bar that appears across the plotting area and tracks of the view.



The position of the cursor in one view is also the position of the cursor in all views. The exact chromosomal location of the cursor appears in the first cell of the Status bar. Several actions affect the position of the cursor:

- In Genome View, click anywhere on a chromosome to move the cursor to that location.
- In Chromosome View, click a cytoband name, part of the chromosome, or anywhere in the plotting area to move the cursor to that location.
- In Gene View, click anywhere in the plotting area to move the cursor to that location.

  The cursor used in Gene View is the same cursor used for the tracks.
- In Tab View, click a row of a data table to move the cursor to the chromosomal location associated with that row.

Tab View

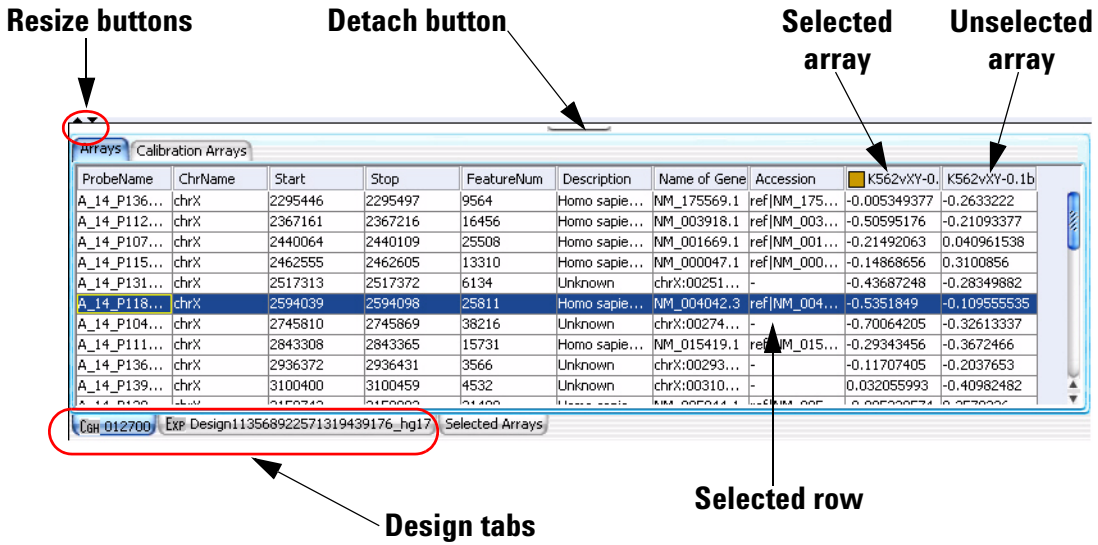


Figure 21 Tab View

Tab View displays design annotation and log ratio data related to the chromosome you select in Chromosome View.

- The exact column content of the tables depends on the specific tab and design, but it always includes chromosomal locations of probes
- The selected row of data appears highlighted in blue. This row represents data that corresponds approximately with the location of the cursor.
- Columns of log ratio data appear below the names of the specific arrays to which they correspond. If an array is selected for display in Chromosome and Gene views, a colored square appears next to its name.

Tab View tabs and buttons

You can see the following tabs and buttons in Tab View. See Figure 21 for a diagram that identifies some of these elements.

Design tabs A separate tab appears for each microarray design included in the active experiment. The name of the design appears on each tab, along with an icon:

- CH3 – A methylation array design
- CGH – An aCGH array design.
- EXP – A gene expression array design.
- CHIP – A ChIP-on-Chip array design.

When you click a design tab, the data and annotation for the arrays in the design appear in Tab View. The program separates the arrays of the design into the Arrays tab and the Calibration Arrays tab (see below).

Arrays tab (Available when you click a specific design tab.) Contains a table of data and annotation for all arrays in a design that contain biological data.

Selected Arrays tab Contains a table of data and annotation for the selected arrays from all designs in the active experiment.

(**Resize** buttons) The button that points away from Tab View expands the view. The other button restores the view to its original size. See "Resizing and detaching panes from the Genomic Workbench main window" on page 19 of the *Quick Start Guide*.

(**Detach** button) Removes Tab View from the main window, and opens it in its own separate window. See "Resizing and detaching panes from the Genomic Workbench main window" on page 19 of the *Quick Start Guide*.

Tab View actions and shortcut menus

- Click the name of an *array in a column heading* to select the array for display.
- Right-click the name of an *array in a column heading* to open a shortcut menu with these options:

Option	Description
Rename Array...	Opens an Input dialog box, where you can type a new name for the array. This only changes the name of the array within the active experiment.
Remove Array From Experiment	Opens a confirmation dialog box. Click **Yes** to remove the link between the array and the active experiment. This command does not delete the data file from the program. To do this, see "To remove data or design files from the program" on page 36.
Select Array	(Available if the array is not selected.) Selects the array for display. A colored square appears next to the name of the array.
Deselect Array	(Available if the array is selected.) Removes the array data from scatter plots, and removes the column of the array from the Selected Arrays tab.
Edit Array Color	Opens the Select Color dialog box, where you can change the display color of the array. See "Edit Array Color" on page 144 and "To change the display color of an array" on page 51.
Edit Array Order	Opens the Edit Array Order dialog box, where you can change the order in which the names of the arrays in a given design of the active experiment appear in Tab View and in the Data Navigator. In Gene View, when you view separate scatter plots for each array, the plots also appear in this order. See "Edit Array Order" on page 145 and "To change the order of arrays in an experiment" on page 28.
Select All Arrays	Selects all arrays in all designs in the active experiment for display. All arrays appear in the Selected Arrays tab.
Deselect All Arrays	Removes all arrays from display, and from the Selected Arrays tab.
Scroll to Column...	Opens the Scroll to Column dialog box, where you can select a column in the current tab. The program then scrolls the data table in the tab so you can see the selected column. See "Scroll to Column" on page 180.

- Right-click a *heading of a column other than an array data column* to open a shortcut menu with a Scroll To Column option. If you click this option, the Scroll to Column dialog box appears, where you can select a column in the current tab. The program then scrolls the data table in the tab so you can see the column.
- Click *an entry in a data table* to select the row in which it appears. This also moves the cursor to the location of the data point corresponding to the selected row.
- Right-click a *data table entry* to open a shortcut menu with these options:

Option	Description
Find in Column	Opens the Find in column dialog box, where you can search for a specific text string within the column you clicked. See "Find in column" on page 157.
Google... LocusLink... PubMed... UCSC HG15(April '03) UCSC HG16(July'03) UCSC HG17(May'04) UCSC HG18(March'06) UCSC mm8(Feb'06) UCSC mm9(July'07)	Opens your Web browser, and passes the column entry you clicked as a search string to the selected site. The UCSC links search the indicated University of California, Santa Cruz database related to the indicated genome build. See "To search the Web for information on probes in Tab View" on page 67.
Customize Link...	Opens the Customize Search link dialog box, where you can create or edit a custom Web link that appears in this shortcut menu. When you click a custom link, the program opens your Web browser, and passes the column entry you clicked as a search string to the site. See "Customize Search Link" on page 140 and "To create a custom Web search link" on page 67.
(other options)	If other options appear in this shortcut menu, they are custom Web search links. Click them to open your Web browser, and pass the column entry you clicked as a search string to the site.

Status Bar

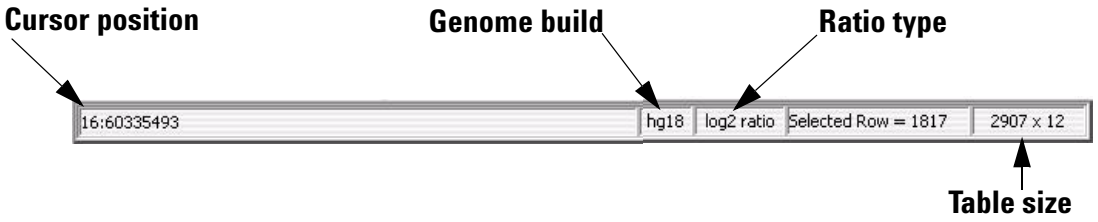


Figure 22 Status bar

The Status Bar displays information related to the currently displayed data.

Cursor position The chromosomal location of the cursor. See "The View Cursor" on page 122.

Genome build The genome build associated with the currently displayed data.

Ratio type The mathematical type of the array data. The possible types are:

- **ratio**
- **$\log_2$ ratio**
- **$\log_{10}$ ratio**
- **ln (natural log) ratio**

Selected Row The row in the currently displayed data table that is selected. The location of the cursor is approximately the chromosomal location associated with this row.

Table size The number of columns in the currently displayed tab. The size appears as <# of rows> x <# of columns>.

Dialog Boxes

This section describes the dialog boxes specific to the interactive analysis of the CH3 Application. They are listed in alphabetical order by title.

Add Gene List <name> to

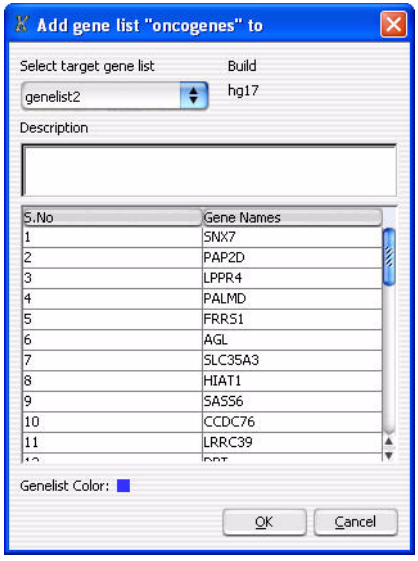


Figure 23 Add Gene List <name> to

Purpose: Adds genes from one gene list (the source gene list) to another (the target gene list).

To open: In the Data pane, right-click the name of a gene list, then click **Add to Gene List.**

Select target gene list The gene list to which genes will be added. Select one from the list.

Build (Read only) The genome build associated with the genes in the list. The builds of the two gene lists must match.

Description (Optional) Description of the combined gene list.

List of genes A list of the genes in the target gene list.

Gene List Color (Read only) The current display color of the target gene list.

OK Adds the genes from the source gene list to the target gene list.

Cancel Closes the dialog box without adding any genes to the target gene list.

Agilent Feature Extraction Importer

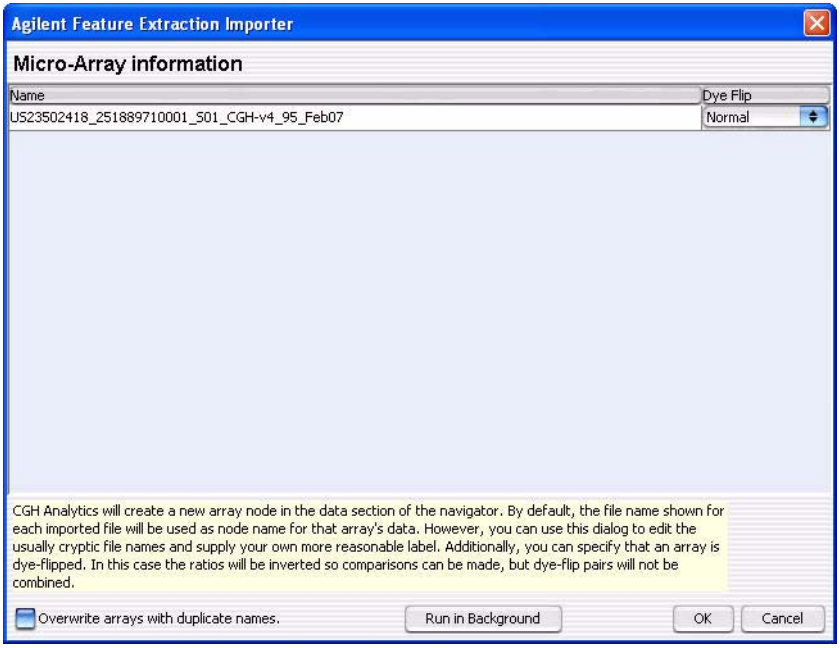


Figure 24 Agilent Feature Extraction Importer

Purpose: Allows you to edit the name of the FE data file you intend to import and to specify if you want to flip the red/green ratio for the data.

To open: In the Home tab, click **Import > Array Files > FE File...,** select the desired FE data file(s), then click **Open.**

Name Allows you to edit the names of the FE files. You can change the names of the files to names that you are more likely to recognize or remember.

Dye Flip For each array:

Select **Normal** if:

- The test samples were labeled with cyanine-5 (red).
- The control samples were labeled with cyanine-3 (green).
- The imported ration (test/control) should be reported directly.

Select **Flipped** if:

- The test samples were labeled with cyanine-3 (green).
- The control samples were labeled with cyanine-5 (red).
- The imported ratio (control/test) should be reported with the ratio inverted (test/control).

The program does not combine dye-flip pairs.

Overwrite arrays with duplicate names Mark this option to replace existing file(s) in the program with the imported one(s), if they have the same name(s).

Run in Background Imports the files, and allows you to use your computer for other purposes while the import occurs. This is especially useful if you have many files to import.

OK Imports the files in the foreground. You cannot use your computer for other purposes while the import occurs.

Cancel Cancels the entire import process without importing anything.

Array Set

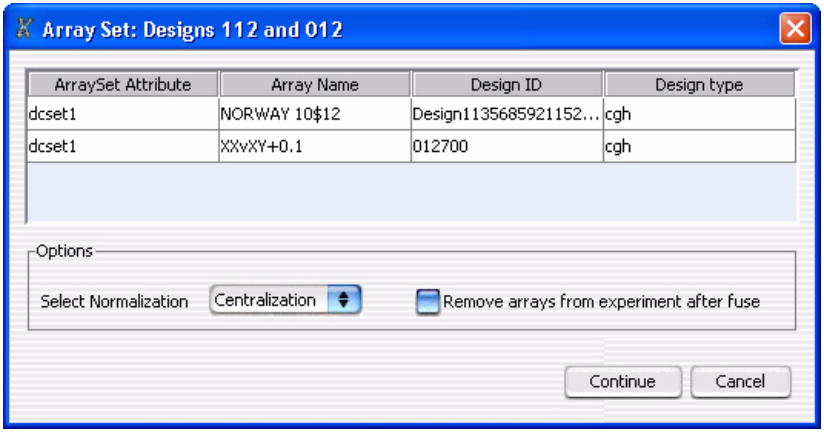


Figure 25 Array Set dialog box.

Purpose: To specify replicate arrays to be combined in the analysis.

To open: Click the Fuse button in the Preprocessing ribbon.

The array set dialog box opens when you fuse designs. See "To combine (fuse) arrays" on page 71.

Table Lists the arrays to be fused, arranged by their values for the ArraySet attribute. The program creates a separate fused array for each group of arrays with a given value for ArraySet. The table also lists the design ID associated with each array, and the design type.

Select Normalization Currently, the Methylation (CH3) Analysis always applies the Centralization normalization algorithm to the arrays in each fused design. See "Methylation Status Detection Algorithm" on page 199.

Remove arrays from experiment after fuse Deletes the original un-fused arrays after creating fused arrays. This can help reduce the duplication of data within the experiment.

Continue Click to create fused designs using the selected options.

Cancel Cancels any selections, and closes the dialog box.

Choose Gene List Color

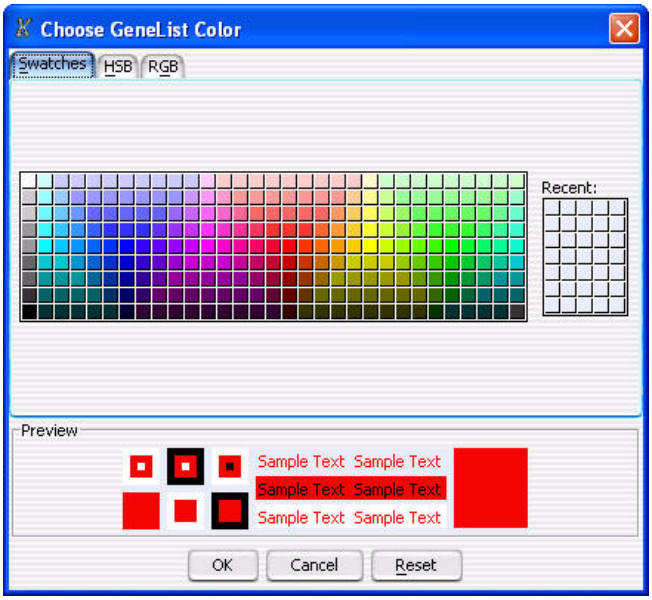


Figure 26 Choose GeneList Color dialog box

Purpose: To distinguish multiple gene list names by color

To open: Right-click Gene View, click **Create Gene List > Change...**, or in Data Navigator, right-click a specific gene list, click **View in Table > Color**.

Swatches Tab Choose color based on color samples (Swatches)

HSB Tab Choose colors based on an HSB schema (Hue, Saturation, and Brightness or Value). See "Select Color (Edit Array Color) – HSB Tab" on page 183.

RGB Tab Choose colors based on an RGB schema (Red-Green-Blue). See "Select Color (Edit Array Color) – RGB Tab" on page 184.

Recent: Choose a recent color selection.

Reset Click to return HSB or RGB values back to default values.

Combine Tracks

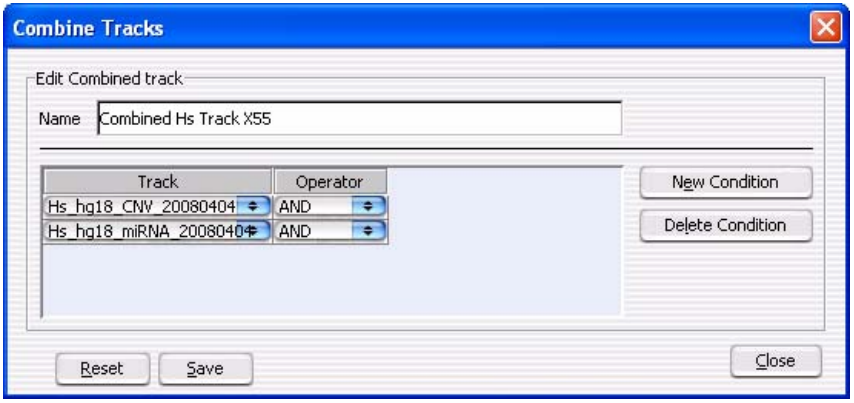


Figure 27 Combine Tracks dialog box

Purpose: Allows you to create a combined track that contains elements of two or more source tracks, associated by logical criteria. See "To combine tracks" on page 39.

To open: In the **My Entity List** pane, right-click the **Tracks** folder, then click **Combine Tracks.**

Name The name of the combined track

New Condition Adds a new, empty row to the Track/Operator table in the dialog box.

Delete Condition Removes the bottom row from the Track/Operator table in the dialog box.

Track In each row, select a track to include in the combined track.

Operator In each row, select the desired logical operator. This operator controls the manner in which the program combines the track in this row with the others. Select one of these options:

Operator	Comments
AND	Places an element in the combined track if it appears in both this track and any of the others.
OR	Places an element in the combined track if it appears in either this track or any of the others. If you set this operator for all tracks in the list, the result is a non-redundant set of elements from all tracks.
MINUS	Removes the elements that appear in this track from the combined track, if they otherwise appear there.

Reset Removes all Track/Operator pairs from the table in the dialog box, and clears the Name of the combined track.

Save Creates the combined track, but does not close the dialog box.

Close Closes the dialog box. Opens the Confirm track save dialog box if you created a combined track, but did not save it.

Confirm overwrite

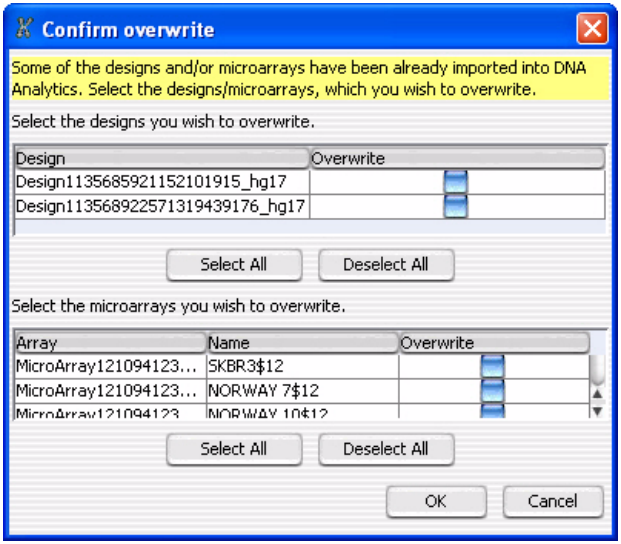


Figure 28 Confirm overwrite dialog box

Purpose: When you import an experiment, it can contain designs and/or arrays that have the same names as those already available in Genomic Workbench. This dialog box allows you to select which designs and/or arrays to overwrite.

To open: This dialog box appears when you import a ZIP format experiment file, and it contains designs and/or arrays that are already available in Genomic Workbench. See "To import an experiment file" on page 24.

Select the designs you wish to overwrite

Design The names of the designs in the imported file that have the same names as designs that are already available in Genomic Workbench.

Overwrite Mark the check box next to each existing design that you want to overwrite.

Select All Marks all of the check boxes under Overwrite.

Deselect All Clears all of the check boxes under Overwrite.

Select the microarrays you wish to overwrite

Array The microarray identification, usually a barcode.

Name The names of the arrays in the imported file that have the same names as arrays that are already available in Genomic Workbench.

Overwrite Mark the check box next to each existing array that you want to overwrite.

Select All Marks all of the check boxes under Overwrite.

Deselect All Clears all of the check boxes under Overwrite.

OK Overwrites the selected files (both designs and arrays) and closes the dialog box.

Cancel Closes the dialog box, and returns you to the Import (experiments) dialog box. See "Import (experiments)" on page 163.

Create Experiment

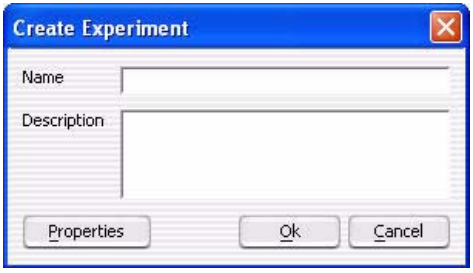


Figure 29 Create Experiment dialog box

Purpose: To create an organizational unit in the Experiment pane of the Navigator to link to array data for viewing and analysis and to create the links to the data (populate) the experiment (in Experiment Properties dialog box)

To open: In the Experiment pane of the Navigator, right-click the Experiments folder, and click New Experiment, or click File > New Experiment.

Name Type a name for your new experiment.

Description Briefly, describe your experiment with information that will help you identify it.

Properties Click to access the Experiment Properties dialog box where you can select microarrays to populate your new experiment. See "Experiment Properties" on page 147.

NOTE Do not click **OK** until you have populated your experiment in the Experiment Properties dialog box or you will have an empty experiment.

Create Gene List

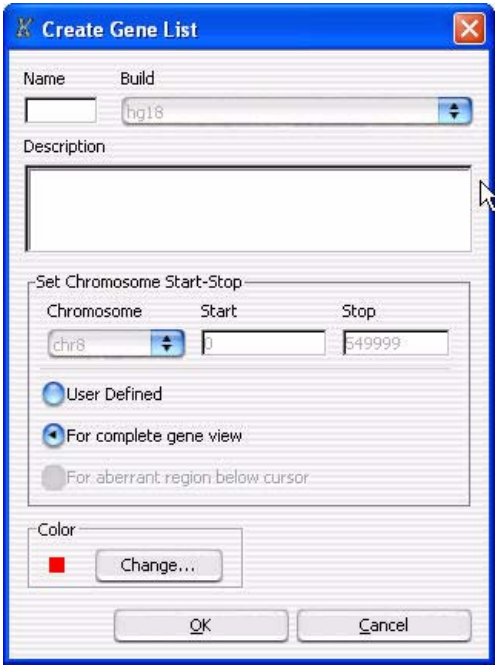


Figure 30 Create Gene List

Purpose: To limit the genes presented in Gene View to a preselected number valuable for interpreting data

To open: Right-click Gene View, and click **Create Gene List**.

Name Type in name of gene list.

Build Select the genome build for the genes to be selected for list.

Description Describe the type or nature of the genes in the list.

Set Chromosome Start-Stop

Select a chromosome and a region in Chromosome View for selecting the genes in the list before you open the Create Gene List dialog box.

User Defined Select to choose region from which the genes in Gene View will be selected. The chromosome selection list and the Start and Stop positions on the Y axis are activated when this option is selected. With this option you can override the selections you made before opening Create Gene List.

For complete gene view Select all the genes in Gene View.

For aberrant region below cursor Select those genes that appear in the aberrant region just below where the cursor sits in Gene View. Not operational in Genomic Viewer; depends on analysis.

Chromosome If you select User Defined, you can select a different chromosome than had been selected before opening the Create Gene List dialog box.

Start If you select User Defined, you can type in a Start position for defining the region contained the genes to be in the list.

Stop If you select User Defined, you can type in a Stop position for defining the region contained the genes to be in the list.

Color

Change... Click to change the color of the gene list name in Data Navigator.

Customize Search Link

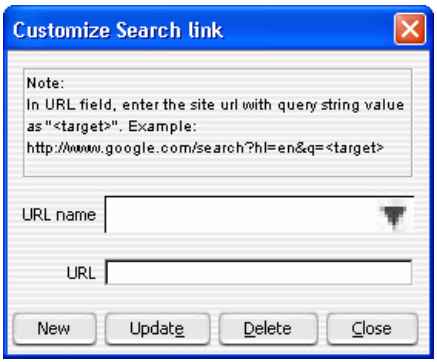


Figure 31 Customize Search Link dialog box

Purpose: This dialog box allows you to create a custom Web search link in the shortcut menu that appears when you right-click a tab entry. The link opens the URL of your choice, and passes the tab entry to it as a search string. See "To create a custom Web search link" on page 67.

To open: Right-click any entry in a tab in Tab View, other than a column heading, then click **Customize Link...**

URL Name The name of the custom Web search link that appears in the shortcut menu (see above). To edit an existing custom Web search link, select it from the list.

URL The full uniform resource locator (URL) of the desired search page. For the query string value, type `<target>`

For example, this URL passes the selected tab view entry to google.com:

```
http://www.google.com/search?hl=eng&q=<target>
```

New Opens an Input dialog box, where you can type a name for a new custom Web search link. Click **OK** to add the name to the URL name list.

Update Saves the settings in the dialog box.

Delete Deletes the currently selected custom Web search link.

Close Closes the dialog box.

Design Properties

Purpose: Provides general and detailed information about a given microarray design. See "To view the properties of a specific design" on page 34.

To open: In the **Data** pane of the Navigator, right-click the name of a genome build within a design folder, then click **Show Properties.** Several tabs are available.

Attribute tab

Lists general identifying attributes of the array design, and statistics such as the total number of features in the design.

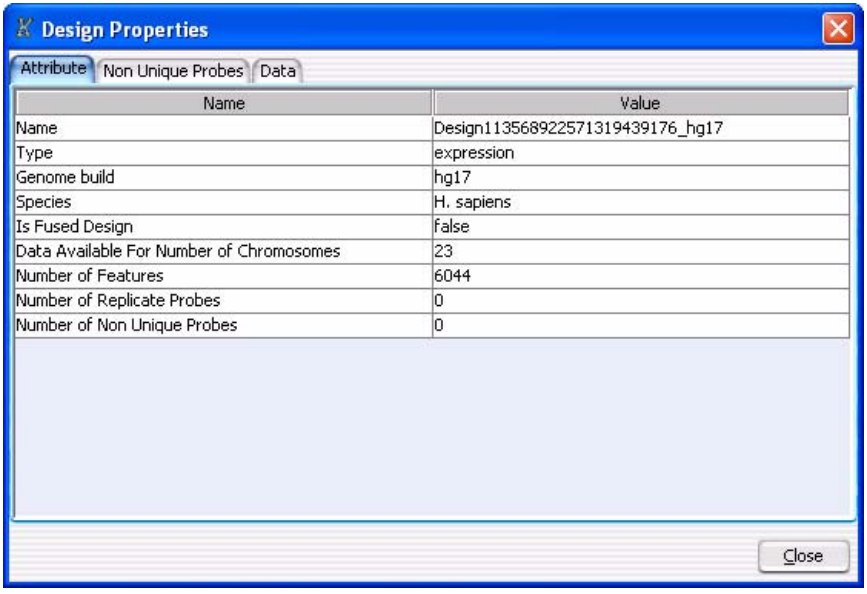


Figure 32 Design Properties dialog box – Attribute tab

Non-Unique Probes tab

Lists the non-unique probes in the design. Non-unique probes bind to more than one location in a target genome.

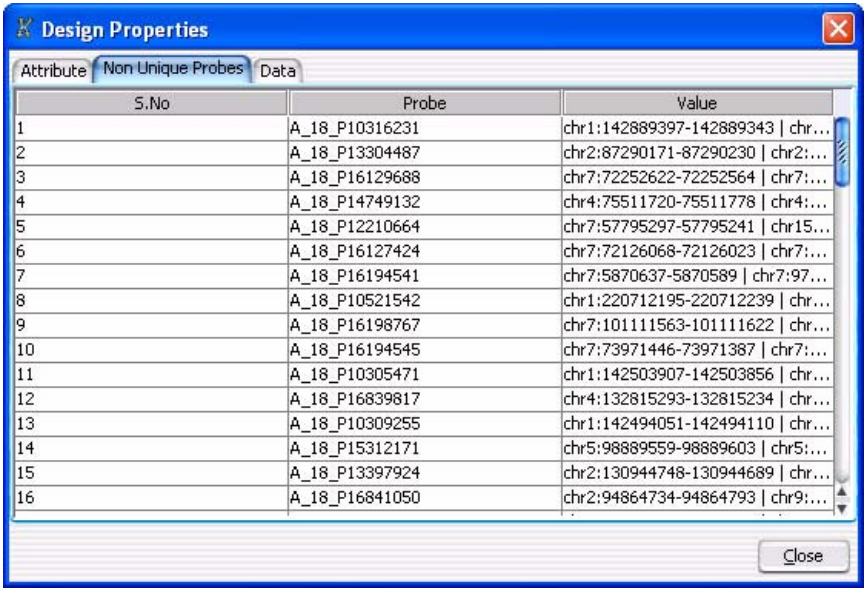


Figure 33 Design Properties dialog box – Non-Unique Probes tab

S. No The sequence order of the probes within the tab.

Probe The name of the each non-unique probe.

Value The chromosomal locations to which each of the probes binds. Because these are non-unique probes, two locations appear for each probe.

Data tab

Lists the names of the probes in the design and the genomic locations to which they are designed. The tab displays the probes for one chromosome at a time.

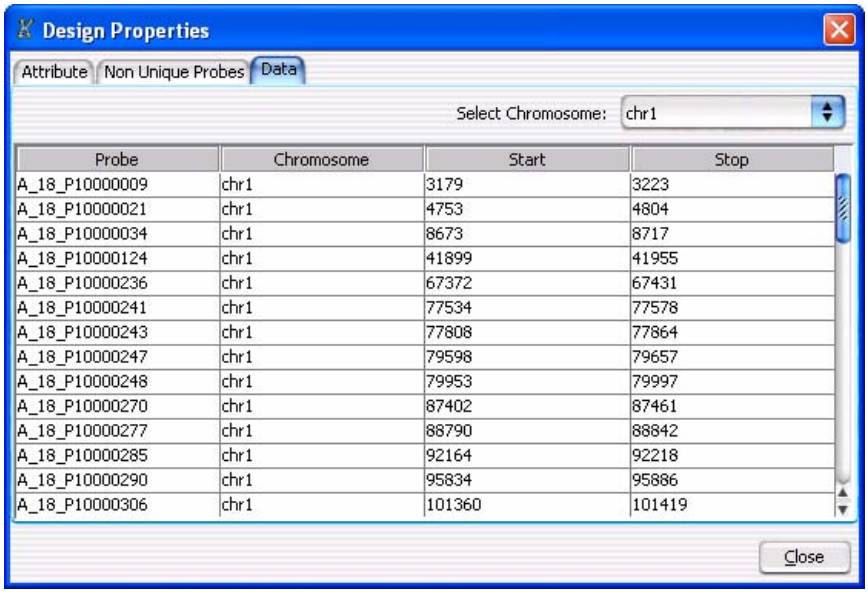


Figure 34 Design Properties dialog box – Data tab

Select Chromosome The chromosome whose probes appear in the list. To view the probes designed to a different chromosome, select one from this list.

Probe The name (Probe ID) of each probe.

Chromosome The name of the chromosome to which the probe is designed.

Start The location on the selected chromosome of the first base pair to which each probe is designed.

Stop The location on the selected chromosome of the last base pair to which each probe is designed.

Edit Array Color

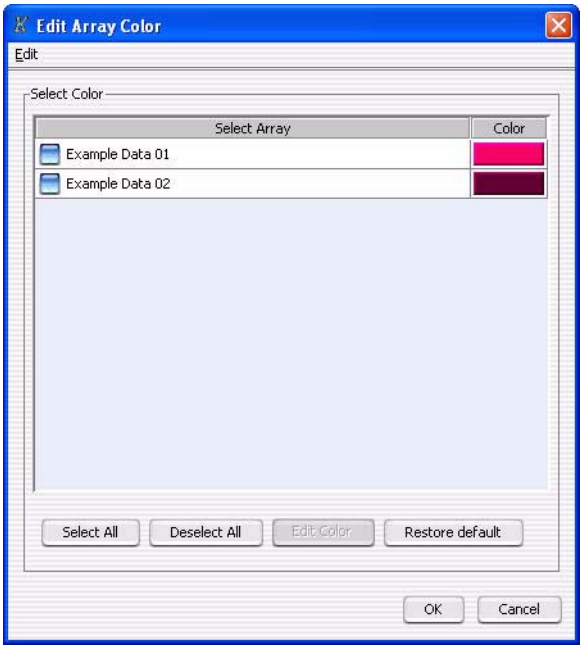


Figure 35 Edit Array Order dialog box

Purpose: To change the color of the arrays in an experiment

To open: Right-click the experiment name, and click Edit Array Color.

Select Array Mark the check box for the array(s) whose color you want to change.

Color Click to change the color for the selected array(s). If you selected more than one, all the selected arrays will change to the same color.

Select All Click to mark all the check boxes.

Deselect All Click to clear all the check boxes.

Edit Color Click to change the color for the selected array(s). Same as Color button.

Restore default Click to restore the original color(s) to the selected array(s).

Edit Array Order

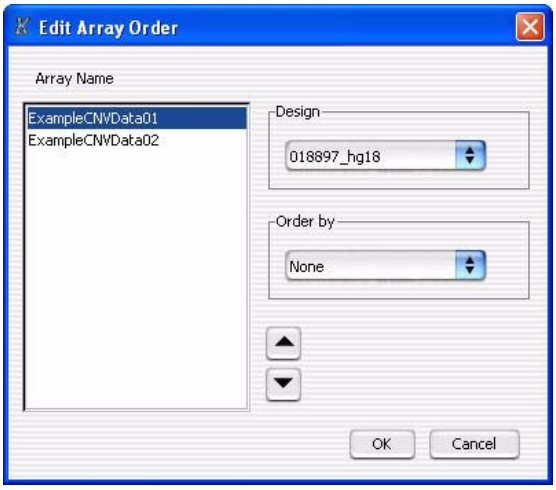


Figure 36 Edit Array Order dialog box

Purpose: To change the order of the arrays in an experiment

To open: Right-click the experiment name, and click Edit Array Order.

Array Name The arrays in the selected design listed in the order used in the Navigator.

Design Select a design from the scroll-down list.

Order by A list box showing the attributes related to the arrays. Select an attribute to use for ordering the list. The arrays are re-ordered based on their respective values for that attribute. For example, in the arrays displayed, the list box showed:

- None
- Array type
- Chip Barcode
- Polarity
- QC Metric Status
- Sample
- Wash Conditions

Experiment Attributes

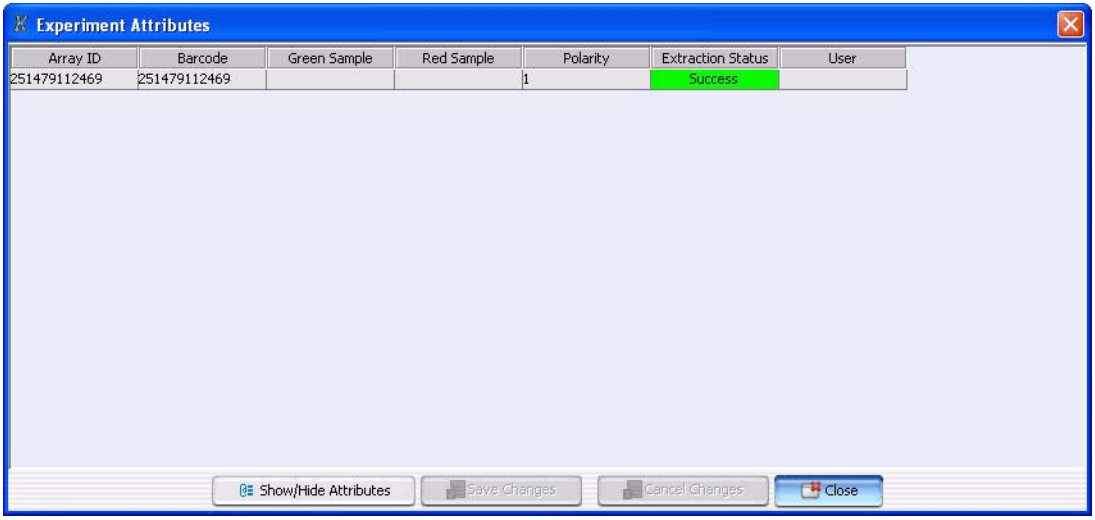


Figure 37 Experiment Attributes dialog box

Purpose: To show, hide, or edit array attributes in the experiment

To open: In the Experiment Pane of the Navigator, right-click the experiment, and click Edit Attributes.

This dialog box lets you enter or change the existing values for the attributes listed for the arrays in the experiment. You can also show or hide attribute columns. The columns that show up initially are the default columns (Array ID, Green Sample, Red Sample, Polarity and Extraction Status) plus any that have been unhidden in the Sample Manager table. See the *Sample Manager User Guide*.

Experiment Properties

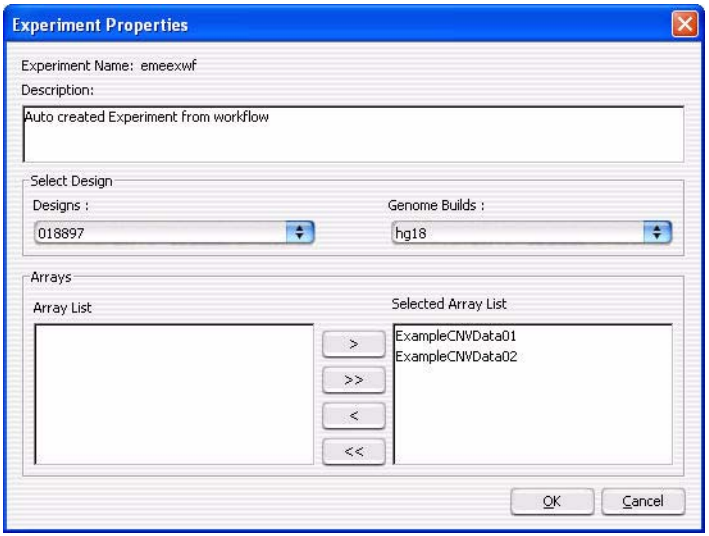


Figure 38 Experiment Properties dialog box

Purpose: To select the arrays to be linked to the experiment

To open: In the Create Experiment dialog box, click Properties, or in the Experiment pane of the Navigator, right-click the experiment name, and click Show Properties.

Experiment Name: The name of the selected experiment is displayed automatically.

Description Text Box: Displays the description of the experiment that was entered when the experiment was created.

Select Design

Designs From the list select the design whose arrays you want to add to the experiment.

Genome Builds From the list select the genome build for the design you selected, if the design has more than one genome build.

Arrays

Array List Lists the arrays in the selected design that are available for the experiment.

- To select an array for subsequent transfer to the Selected Array List, click its name.
- To select additional arrays, control-click their names.
- To select a contiguous block of arrays, click the name of the first array, then shift-click the name of the last one.

Selected Array List Lists the arrays that you have selected for the experiment.

> Moves the selected arrays in Array List to the Selected Arrays List. You can move arrays from as many designs as you like, as long as they are all associated with the same genome build.

>> Moves all of the arrays in Array List to the Selected Arrays List.

< Removes an array from the Selected Array List. To select an array for removal, click its name. If desired, you can re-add an array.

<< Clears the Selected Array List.

OK Accepts the current arrays for the experiment and closes the dialog box.

Cancel Closes the dialog box without making any changes.

Export

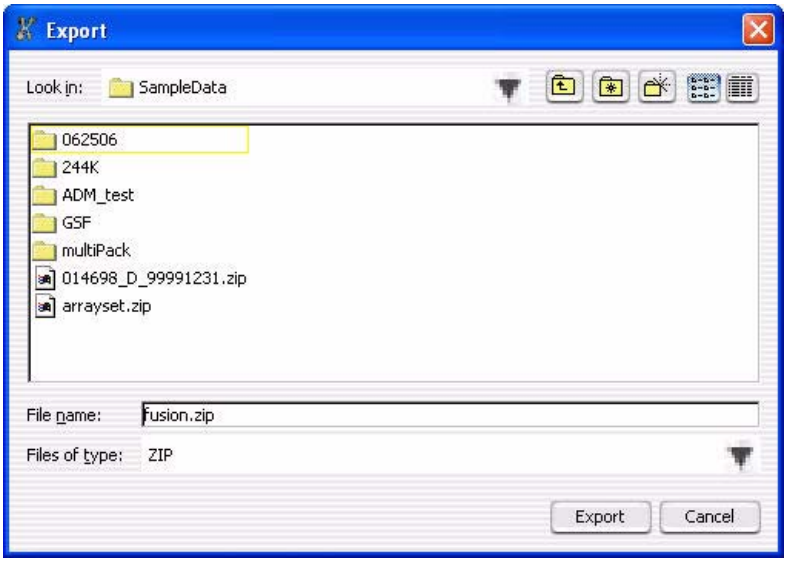


Figure 39 Export dialog box – Several types of file exports use this dialog box. This specific example exports selected experiment(s) as a ZIP format file.

Purpose: Allows you to select a location for an exported file.

To open: This dialog box appears after you select specific experiment(s), track(s), or array attribute(s) to export. See "To export experiments" on page 43, "To export tracks" on page 44, or "To export array attributes" on page 42.

Look in Displays the folder or other location whose contents appear in the main pane of the dialog box. To select another folder or other location, click .

Navigates up one level.

Navigates to the Desktop.

Creates a new folder in the selected location in *Look in*.

Displays the names, only, of folders, files, and other locations in the main pane of the dialog box.

Displays both the names and details of folders, files, and other locations in the main pain of the dialog box.

Main pane Displays the folders, files, and other locations in the selected location in *Look in*. The program restricts listed files to the type selected in *Files of type*. To select file, click its name. To open a folder or other location, double-click its name.

File name Displays the name of the file to which the exported content will be saved. To change the name, you can either select a file in the main pane of the dialog box, or type a new name.

Files of type Restricts the files listed in the main pane to those of the appropriate type for your specific kind of export. To show all files, click ▼, then select **All Files.**

Export Saves the selected content to the location specified in the dialog box.

Cancel Cancels your selections and closes the dialog box.

Export Array Attributes – Array

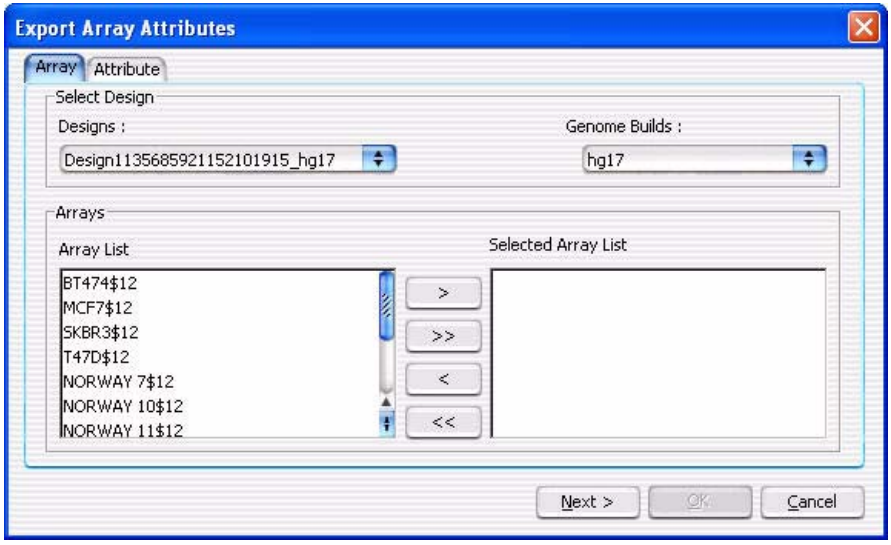


Figure 40 Export Array Attributes – Array tab dialog box

Purpose: Allows you to select array designs whose selected attributes you will then export. See "To export array attributes" on page 42.

To open: In the Home command ribbon, click **Export > Array Attributes...**, or in the **Experiment** pane of the Navigator, right-click the name of an experiment, then click **Export Attributes.**

Select Design

Designs Lists all of the designs available in the program. Select the design associated with arrays that you want to export.

Genome Builds Lists the genome build(s) associated with the design. Select the desired genome build to display the arrays associated with a single genome build.

Arrays

Array List Lists the arrays in the selected design that are available for export.

- To select an array for subsequent transfer to the Selected Array List, click its name.
- To select additional arrays, control-click their names.
- To select a contiguous block of arrays, click the name of the first array, then shift-click the name of the last one.

Selected Array List Lists the arrays that you have selected for export.

> Moves the selected arrays in Array List to the Selected Arrays List. You can move arrays from as many designs as you like, as long as they are all associated with the same genome build.

>> Moves all of the arrays in Array List to the Selected Arrays List.

< Removes an array from the Selected Array List. To select an array for removal, click its name. If desired, you can re-add an array.

<< Clears the Selected Array List.

Next Moves to Attribute tab to select attributes for export. See "Export Array Attributes – Attribute" on page 153.

Cancel Closes the dialog box without selecting any array attributes to be exported.

Export Array Attributes – Attribute

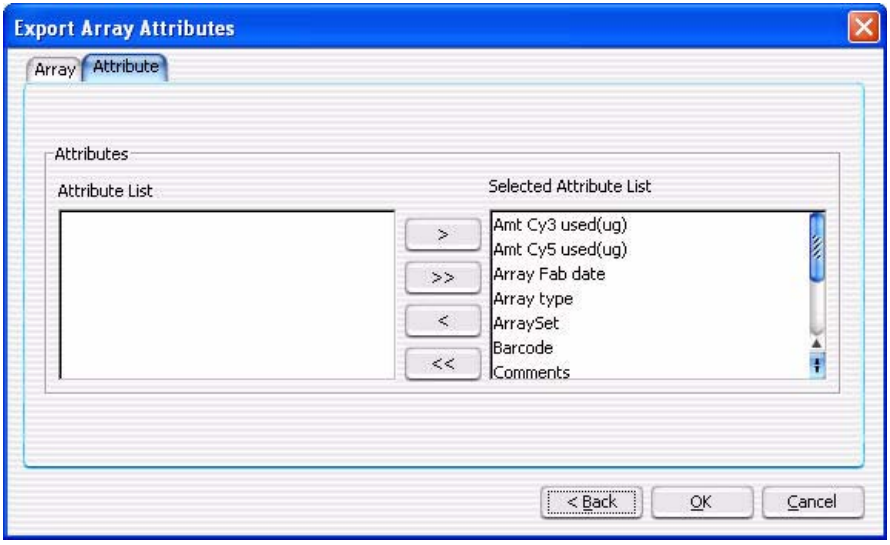


Figure 41 Export Array Attributes – Attribute tab dialog box

Purpose: Allows you to remove selected array attributes from the list you will then export. See "To export array attributes" on page 42.

To open: In the Home command ribbon, click **Export > Array Attributes...**, or in the **Experiment** pane of the Navigator, right-click the name of an experiment, then click **Export Attributes.** After selecting the arrays whose attributes you intend to export, click **Next.**

Attributes

Selected Attribute List

Lists the attributes for the selected arrays that will be exported if you click OK.

- To select an attribute for subsequent removal to the Attribute List, click its name.
- To select additional attributes, control-click their names.
- To select a contiguous block of attributes, click the name of the first attribute, then shift-click the name of the last one.

Attribute List Lists the attributes that will not be exported for the selected arrays.

Removes an attribute from the Selected Attribute List. To select an attribute for removal, click its name. If desired, you can re-add an attribute.

Clears the Selected Attribute List.

Moves the selected attributes in the Attribute List to the Selected Attribute List.

Moves all of the attributes in the Attribute List to the Selected Attribute List.

Back Moves back to the Array tab for array selection or removal.

OK Opens the Export dialog box. See “Export” on page 149.

Cancel Closes the dialog box without adding the attributes in the Selected Attribute List to be exported.

Export Experiments

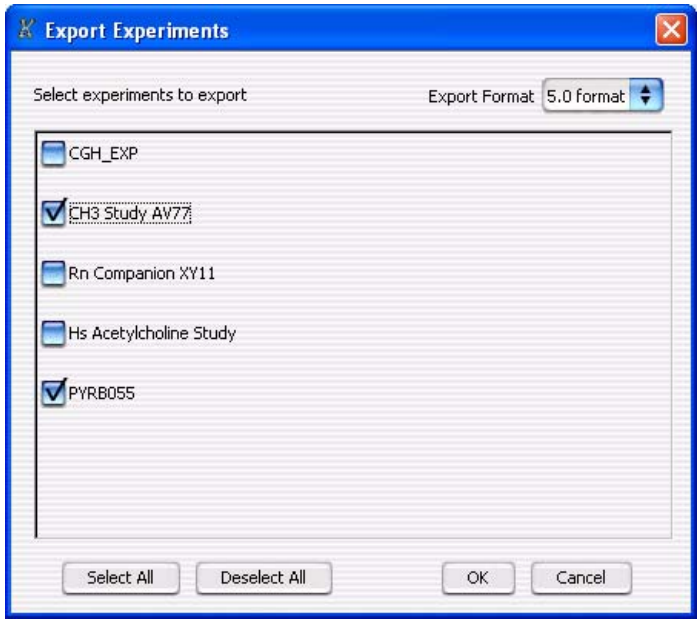


Figure 42 Export Experiments dialog box

Purpose: Allows you to select experiments for export. The program exports all array designs and data associated with the experiments as a single ZIP file. This file does not include any parameter settings, array selections, or results. See "To export experiments" on page 43.

To open: In the Home tab, click **Export > Experiments...**

Export Format The file format for the exported experiment file. Select one of these options:

- **5.0 Format** – Exports the experiment(s) in a format that you can import into Genomic Workbench. This is the most current experiment format, but it is not compatible with previous versions of the program.
- **3.0 Format** – Exports the experiment(s) in a format that you can import into Agilent CGH Analytics 3.0 or later. This is a "legacy" format that you can use to maintain compatibility with earlier versions of the program.

Select experiments to export Lists all experiments available for export. Mark each experiment you want to export.

Select All Selects all experiments for export.

Deselect All Clears all check boxes under Select experiments to export.

OK Opens an Export dialog box. See "Export" on page 149.

Cancel Cancels the export and closes the dialog box.

Export Tracks

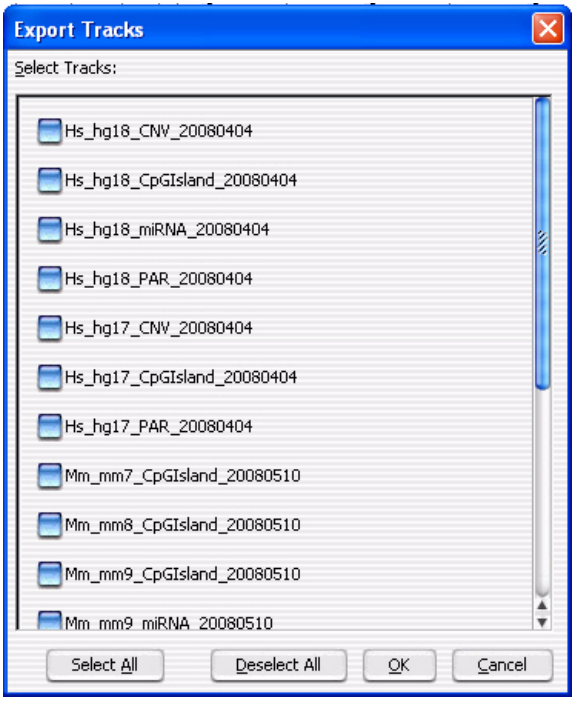


Figure 43 Export Tracks dialog box

Purpose: Allows you to select tracks to export as a single BED format file. See "To export tracks" on page 44.

To open: In the **Home** tab, click **Export > Tracks...**

Select tracks Lists all of the filters available in the program. Mark the check box next to each track that you want to export.

Select All Selects all available tracks for export.

Deselect All Clears all of the check boxes under Select Tracks.

OK Opens the Export dialog box, where you can select a location for the exported BED format file. See "Export" on page 149.

Cancel Cancels the export and closes the dialog box.

Find in column

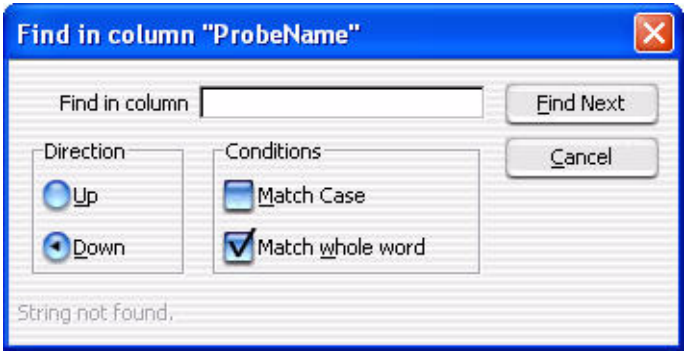


Figure 44 Find in column dialog box

Purpose: This dialog box allows you to set search parameters for a specific column entry. Based on these parameters, the program can highlight the row of the first entry that matches. The cursor then moves to the location defined in the row.

To open: Right click any entry in a tab in Tab View other than a column heading, then click **Find in column** in the shortcut menu.

Find in column Type all or part of the entry you want to find.

Direction Select a search direction:

- **Up** – Sets the search to scan the column you clicked in an upward direction from the currently highlighted row.
- **Down** – Sets the search to scan the column you clicked in an downward direction from the currently highlighted row.

Conditions Mark any of these search options:

- **Match Case** – Mark this option to take case into account. For example, if you mark Match Case, and you type `aa351` in Find in column, the search finds the next entry in the column that contains **aa351.** It does *not* find entries that contain **AA351** or **Aa351.**
- **Match whole word** – Mark this option to only find entries in which the complete entry matches what you type in Find in column. For example, if you type `AA351` in Find in column, and mark **Match whole word,** the program finds the next **AA351** entry. It does not find entries such as **AA3512** or **AA351992.**

Find Next Finds the next matching entry in the selected column, and moves the cursor to the location defined in the row that contains the entry.

Cancel Closes the dialog box.

Gene List

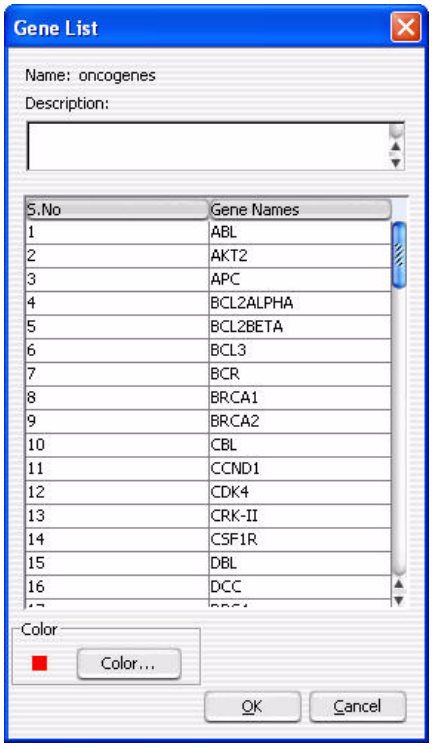


Figure 45 GeneList dialog box

Purpose: Allows you to view the names of the genes in a specific gene list and to change the display color of the gene list. See "To view the genes in a gene list" on page 38.

To open: In the **My Entity List** pane of the Navigator, right-click the name of a gene list, then click **View in Table.**

Name (Read only) The name of the gene list.

Description (Optional) Brief descriptive comments about the gene list, such as how it was created or the nature of the genes in the list. You can edit the description.

S. No The sort order number. This is the index number of each gene within the gene list.

Gene Name The names of the genes in the gene list.

Color... Opens the Choose Gene List Color dialog box, where you can change the display color for the gene list. See "Choose Gene List Color" on page 132.

OK Saves the gene list with any new description or display color, and closes the dialog box.

Cancel Closes the dialog box without making any changes to the gene list.

Go To Gene/Genomic Location

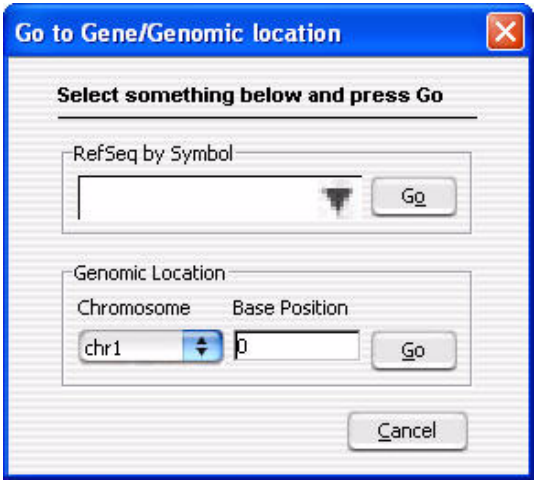


Figure 46 Go To Gene/Genomic location dialog box

Purpose: To find a specific gene location in Gene View by either selecting the RefSeq by Symbol or by selecting the Genomic Location.

To open: Click **Home > Go to Gene/Genomic location**.

RefSeq by Symbol Select the Reference Sequence accession symbol from NCBI, and click **Go.**

Genomic Location

- Chromosome – The chromosome number.

- Base Position – The position on the chromosome.

Click **Go** after selecting the chromosome number and the position of the gene on the chromosome.

Cancel Closes the dialog box.

Import

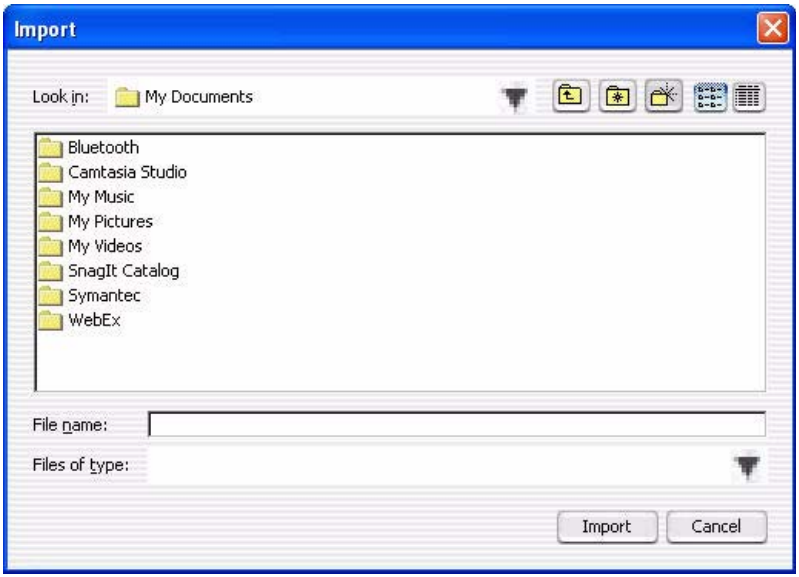


Figure 47 Import dialog box

Purpose: Allows you to select files for import into Genomic Workbench.

To open: In the **Home** tab, click **Import,** then select any kind of import except Genome Build or Track. The type of file to be imported appears in the title of the dialog box. To import a gene list, right-click the **Gene List** folder in the **My Entities List** pane of the Navigator, then click **Import Gene List.**

Use the standard Windows Explorer commands in the dialog box to select a file for import.

For some imports, you can select multiple files. Click the name of the first file, then control-click the names of additional files. To select a contiguous block of files, click the name of the first file in the block, then shift-click the name of the last one.

File name Displays the name of a file you select for import.

Files of type The program restricts the list of files to the specific types expected for the import. To display all files, click ▼,then select **All Files.**

Type of import	File type
FE array File	*.txt
Axon array file	*.gpr
UDF file	*.txt
Design file (GEML)	*.xml
Axon design file	*.gal
Array attributes	*.txt
Experiments	*.zip
Filters	*.xml
Gene list	*.txt

Import Imports the file into the program. In some cases, the name of this button is *Open,* rather than *Import.* Also, when you click **Import,** in many cases one or a series of additional dialog box(es) allows you to further define the content for import. See the instructions for each specific type of import in Chapter 1.

Cancel Cancels the import and closes the dialog box.

Import (experiments)

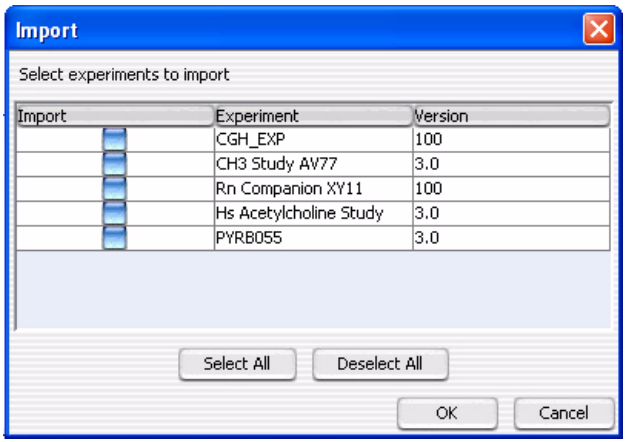


Figure 48 Import dialog box (for experiments)

Purpose: Allows you to select the specific experiments within a ZIP format experiment file to load into the program. See "To import an experiment file" on page 24.

To open: In the **Home** tab, click **Import > Experiments.** In the dialog box that appears, select the desired ZIP format experiment file, then click **Import.**

Select experiments to import

These columns appear:

- **Import** – Mark the check box next to the experiment(s) that you want to import.
- **Experiment** – The names of the experiments available for import in the ZIP format experiment file.
- **Version** – The experiment version.

Select All

Selects all of the experiments in the ZIP file for import.

Deselect All

Clears all of the check boxes under Import.

OK Imports the selected experiments into the program. If the name of an imported array design or data file matches one that is already available in the program, the Confirm overwrite dialog box appears, where you can select the data and/or design files that you want to overwrite. See "Confirm overwrite" on page 135.

Cancel Cancels the upload and closes the dialog box.

Select All Selects all of the filters in the list for import.

Deselect All Clears all of the check boxes under Import.

OK Imports the selected filters and closes the dialog box.

Cancel Cancels the import and closes the dialog box.

Import GEML design files

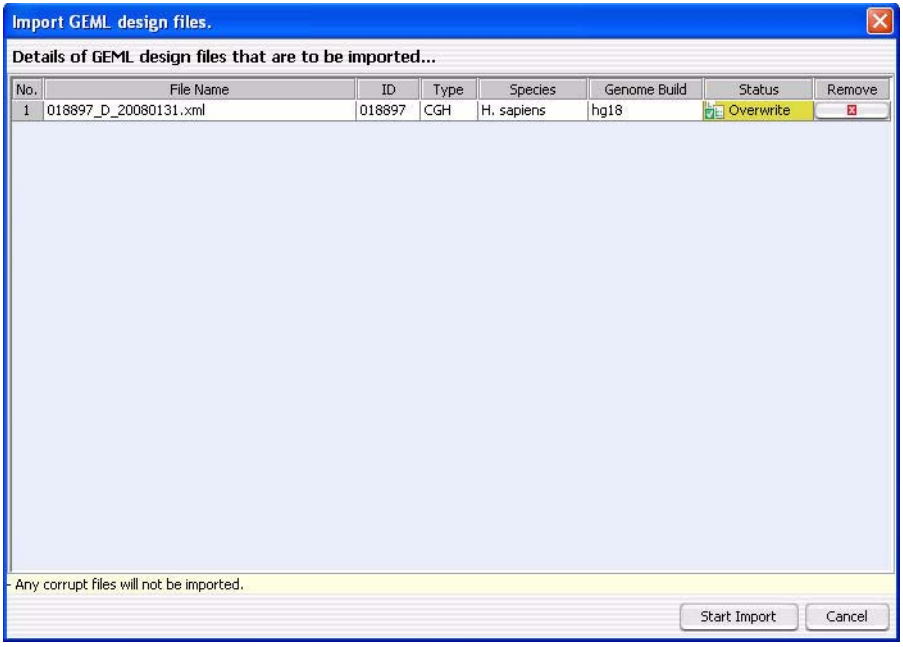


Figure 49 Import GEML design files dialog box

Purpose: To view information in the design file and to remove any files that you don't want to import.

To open: In the Home tab, click **Import > Design Files > GEML File...** Select the desired *.xml design files, then click **Open.**

File Name The name(s) of the design file(s) to be imported.

ID The Agilent ID number for the design file

Type The application type, which can be CGH, ChIP, miRNA, or gene expression.

Species At present, Genomic Workbench supports these species:

- *H. sapiens*
- *M. musculus*
- *R. norvegicus*

Genome Build The genome build with which this design is associated.

Status

- **Overwrite** – Appears when the design file has been updated and will overwrite any existing one of the same name.
- **Valid** – Appears when the file is new.
- **Corrupt** – Appears when the file is corrupt.

Remove Click  to remove a specific design file from the list.

Start Import Starts the import of the design files in the list.

Cancel Cancels the upload and closes the dialog box.

Import Genome Build

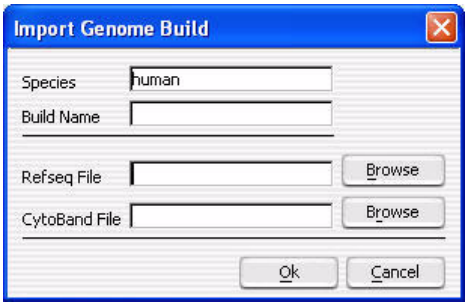


Figure 50 Import Genome Build dialog box

Purpose: To import a new set of genome build files into Genomic Workbench. See "To import a genome build" on page 21.

To open: In the Home tab, click **Import > Genome Build.**

Species The genome's species of origin. The program supports these species:

- *H. sapiens*
- *M. musculus*
- *R. norvegicus*

Build Name The name of the build to be imported.

Refseq File The location of the RefSeq database file. This file contains chromosomal locations of genes. To select a Refseq file, click **Browse...**

CytoBand File The location of the applicable cytoband file. This file contains graphical cytoband information for Gene View and Chromosome View. To select a cytoband file, click **Browse...**

OK Imports the genome build and closes the dialog box.

Cancel Cancels the import and closes the dialog box.

CAUTION Import only Agilent-provided genome build files.

Import Track

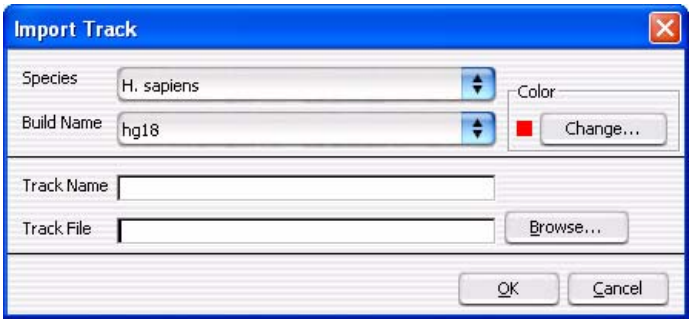


Figure 51 Import Track dialog box

Purpose: Allows you to import a BED format track file. See "To import tracks" on page 22. Track information can appear in Gene view. See "Gene View" on page 118.

To open: In the **Home** tab, click **Import > Track...**

Species Select the species to which the track relates. The program supports these:

- *H. sapiens*
- *M. musculus*
- *R. norvegicus*

Build Name This list contains the available genome builds for the selected species. Select the desired genome build.

Color Shows the currently assigned display color for the track. To change this color, click **Change...**

Track Name Type a name to identify the imported track.

Track File Type the location of the BED track file that you want to import, or click **Browse...** to select a file.

Browse... Opens an Open dialog box, where you can select the BED track file that you want to import.

OK Imports the track into the program.

Cancel Cancels the import and closes the dialog box.

Microarray Properties - Attribute Tab

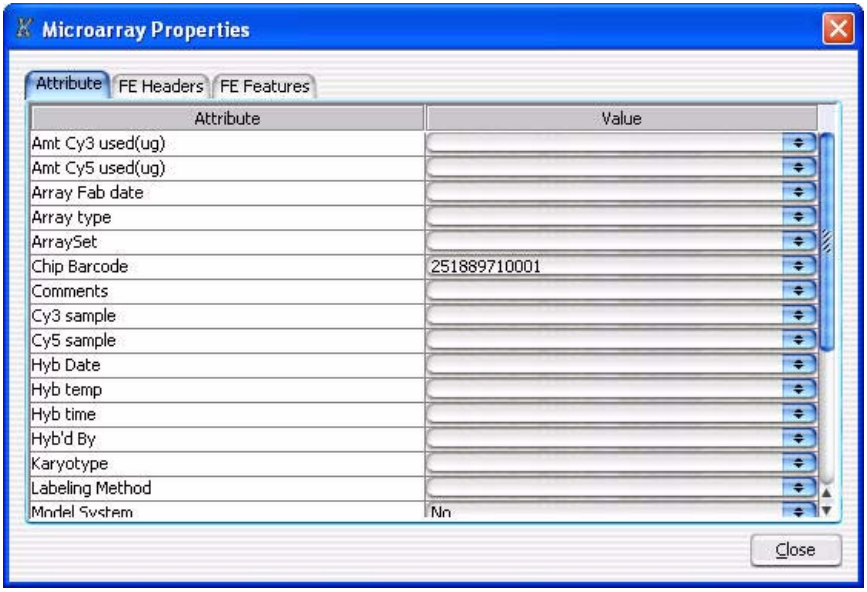


Figure 52 Microarray Properties dialog box listing Attributes and their values

Purpose: To view the list of attributes and their values associated with an array.

To open: For any array in the Data folder or Experiments folder, right-click the array name and click Show Properties. For third-party or non-Agilent arrays, only the Attribute Tab appears.

Attribute Tab

- **Attribute** — Lists the attributes in an array by name. You can load these from an Excel spreadsheet.
- **Value** – Indicates the values, if any, for each array.

Close Closes the dialog box.

For the FE Headers tab options, see "Microarray Properties - FE Headers" on page 170.

For the FE Features tab options, see "Microarray Properties - FE Features Tab" on page 169.

Microarray Properties - FE Features Tab

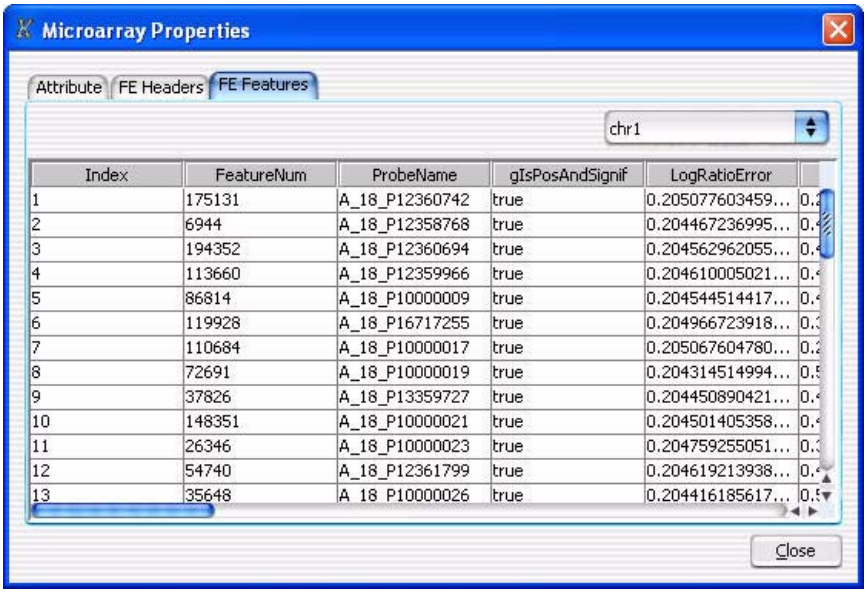


Figure 53 Microarray Properties dialog box listing FE Features and associated data

Purpose: To view feature information for arrays extracted with Agilent Feature Extraction software.

To open: Right-click the array name and click Show Properties.

Selection List Select the chromosome whose feature information you want to display.

List Box Displays FE features and the associated data. The fields are:

Index	FeatureNum	ProbeName
GIsPosAndSignif	LogRatioError	PValueLogRatio
gProcessedSignal	rProcessedSignal	gMedianSignal
rMedianSignal	gBGSubSignal	rBGSubSignal
gIsSaturated	rIsSaturated	gIsFeatNonUnifOL
rIsFeatNonUnifOL	gIsBGNonUnifOL	rIsBGNonUnifOL
rIsPosAndSignif	gIsWellAboveBG	rIsWellAboveBG

Microarray Properties - FE Headers

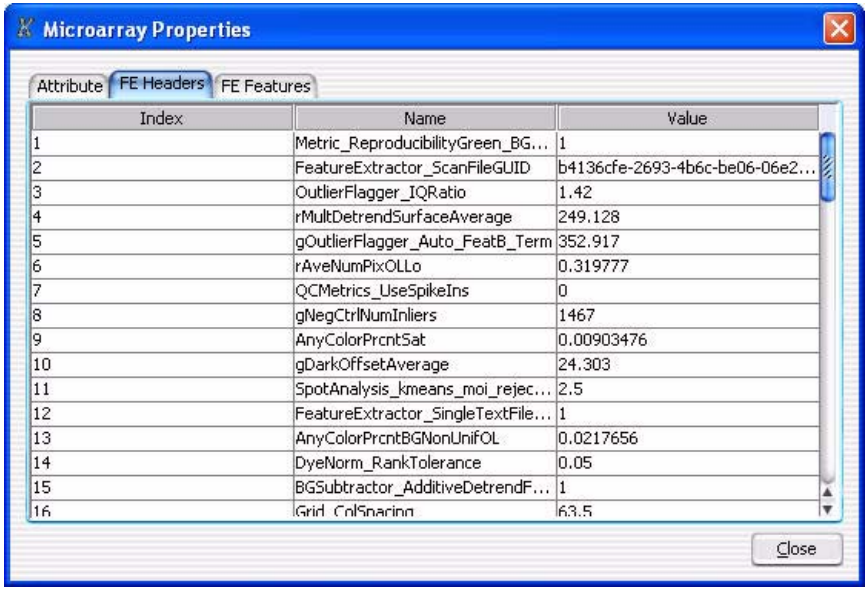


Figure 54 Microarray Properties dialog box listing FE Headers their values

Purpose: To view feature parameters, statistics and constants for the whole array extracted with Agilent Feature Extraction software.

To open: For Agilent arrays in the Data folder or Experiments folder, right-click the array name and click Show Properties.

Index Displays a sequential index to help identify FE properties.

Name Displays feature parameters, statistics and constants for the whole array.

Value Displays the value for each parameter, statistic and constant.

Close Closes the dialog box.

Preferences – License

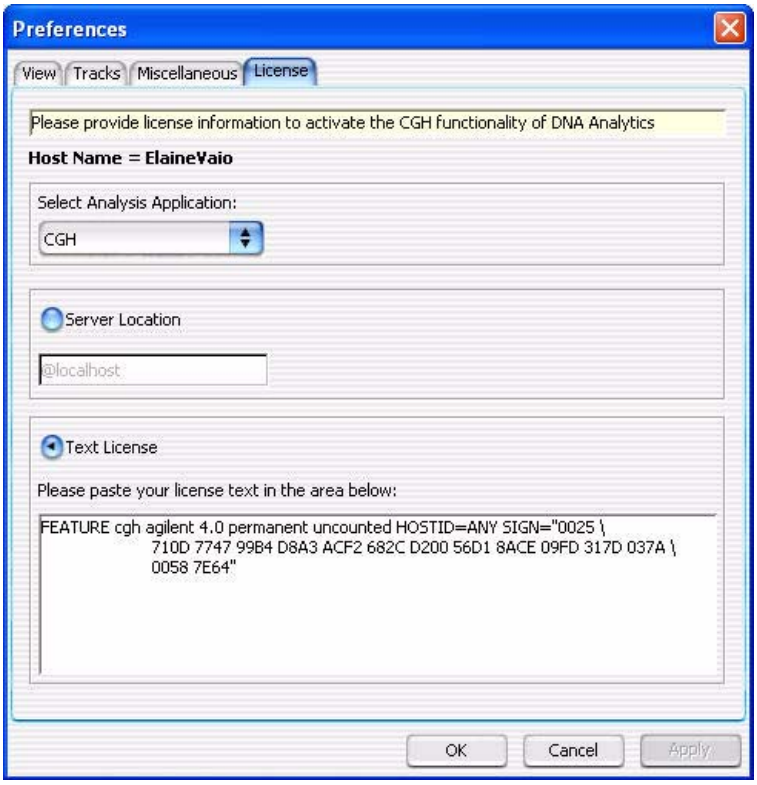


Figure 55 Preferences dialog box displaying License tab options

Purpose: To view and update the license for each module you have installed with the DNA Analytics software, if necessary.

To open: Right-click anywhere in the graphical interface, either in Genome View, Chromosome View or Gene View, click Preferences and click the License tab.

Host Name Displays the host name, automatically.

Select Analysis Application Select the DNA Analytics application for which you have a license.

Server Location Server location should be selected if you have a concurrent user license. If appropriate, click to enable and type in the name of your license server. The default is @localhost. Replace localhost with the name of the computer used as the license server. **Text License** is disabled (grayed) if **Server Location** is enabled.

Text License Text licenses are used if you have a workstation license. If you do have a workstation license, paste your license in the text box. If you have entered a license previously, it is displayed in the text box. **Server Location** is disabled (grayed) when **Text License** is enabled.

Apply Apply your changes to the parameters.

OK/Cancel Accept your changes and exit, or cancel all changes and return to the previously selected parameters.

Preferences – Miscellaneous Tab

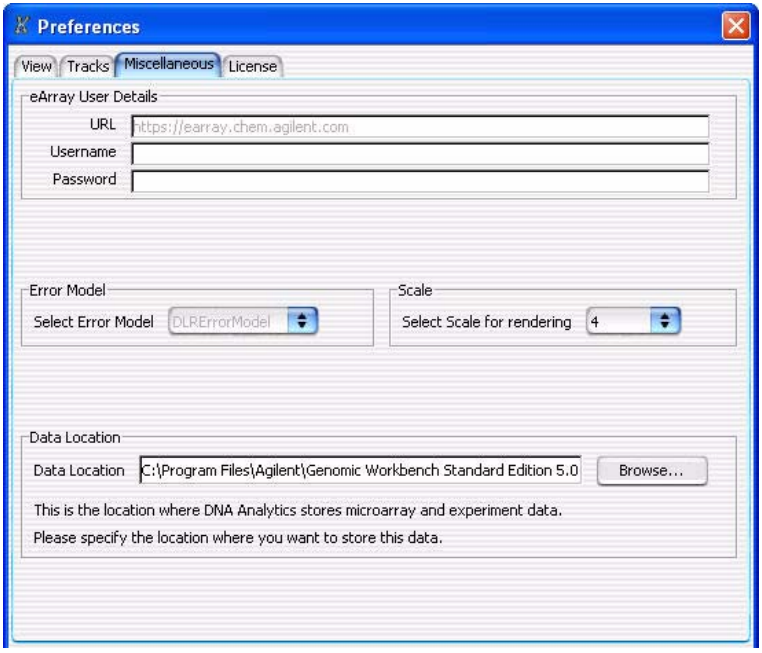


Figure 56 Preferences dialog box displaying Miscellaneous tab options

Purpose: In the Methylation (CH3) Analysis, to change the location for data.

To open: Right-click anywhere in the graphical interface, either in Genome View, Chromosome View or Gene View, click Preferences and click the Miscellaneous tab.

eArray User Details Sets login details for the Agilent eArray Web site.

- **URL** – At present, https://earray.chem.agilent.com and grayed out
- **Username** – The name registered on the eArray site.
- **Password** – The password registered on the eArray site.

Scale Select the scale for rendering the appearance of data in the display.

Data Location Type in the path to the folder where you want the ChIP module to store the microarray and experiment data.

Apply Apply your changes to the parameters.

OK Accept your changes and close the dialog box.

Cancel Cancel all changes and close the dialog box.

Preferences – Tracks Tab

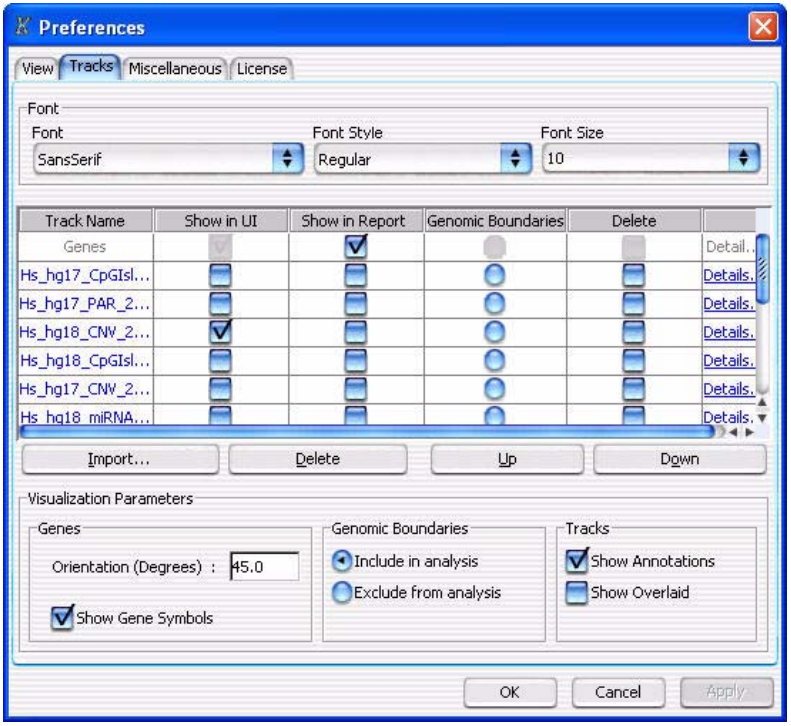


Figure 57 Preferences dialog box displaying Tracks Tab options

Purpose: To import and set up the appearance of tracks next to the Gene View. Tracks are additional graphic displays of genomic information loaded from an external file and that align with genomic coordinates in Gene View.

To open: Right-click anywhere in the graphical interface, either in Genome View, Chromosome View or Gene View, click **Preferences** and click the **Tracks** tab.

Font Options

Select the font type, style and size for the gene annotations that appear in the selected tracks.

Tracks List

Track Name Name of the track already loaded or imported

Show in UI Mark the check box to view the track next to Gene View.

Show in Report Mark the check box to view the track information in all the reports.

Genomic Boundaries Click to use the track to define only the regions that aberration detection algorithms will run. You can choose to do this for only one track.

Delete Click to delete the track from the list.

Details Click to view all the chromosome locations defined in the track.

Import Click to import new tracks.

Up Click to move a track up the list.

Down Click to move a track down the list.

Visualization Parameters

Genes These options affect the appearance of Gene View.

- Orientation – Type in a number in degrees to set the angle at which the Gene Symbols will appear in Gene View and the Track Annotations appear in the tracks.
- Show Gene Symbols – Mark to show them in Gene View, and clear the check box to hide them.

Genomic Boundaries These options allow you to include or exclude the Genomic Boundaries from the analysis.

Tracks These options affect the appearance of the Track Views.

- Show Annotations – Mark to show the names of the gene regions for the tracks, and clear to hide them.

- Show Overlaid – Mark to overlay all the tracks that appear next to Gene View, and clear the check box to view the information in separate tracks.

Preferences – View Tab

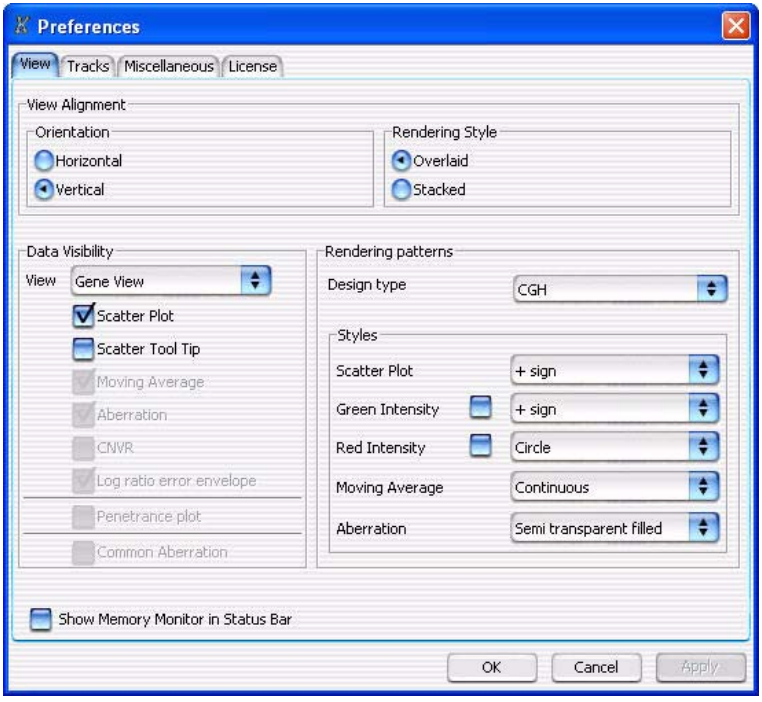


Figure 58 Preferences dialog box displaying View tab options for Methylation (CH3) Analysis

Purpose: To set up how the Views are aligned and settings for the scatter plot in Genome Viewer

To open: Right-click anywhere in the graphical interface, either in Genome View, Chromosome View or Gene View, click Preferences and click the View tab.

View Alignment

Orientation Selects the orientation of three views in the main window:

- Horizontal – Reorients three views to a horizontal aspect in the order of Gene, Chromosome, and Genome views, top to bottom. The Navigator and Tab view orientation remains unchanged.
- Vertical – Displays all views in a vertical aspect, left to right: Navigator, Genome, Chromosome, and Gene views. This is the default display. See "Genomic Workbench CH3 Application Main Window" on page 84.

Rendering Style Selects the way Chromosome and Gene data are rendered on your screen.

- Overlaid – Displays data from multiple arrays superimposed one on top of another (default).
- Stacked – Displays data from each array in a separate plot.

Data Visibility

View Choose what features you want to display for the Genome, Chromosome, and Gene views, either individually or together. Select one or more check boxes:

- Scatter Plot (Chromosome and Gene Views only)
- Scatter Tool Tip (Gene View only)
- Moving Average — Marking this check box enables the display of the moving average for both log ratio data and Z-score data. To see the moving average, you must also mark the Show check boxes for either or both in the Analysis ribbon.
- ZScores — Marking this checkbox displays the results from the methylation algorithm to produce the statistical likelihood of methylation or no methylation in CpG Island regions of the genome.
- Log ratio error envelope (Chromosome and Gene Views only)

Rendering patterns

Design type Specify the type of design to which you are applying these patterns: **ChIP, Expression, or Other.**

Styles Set up the parameters for displaying your data.

- Scatter Plot -- Specify how to display individual data points as: **Color filled circles (ellipses), circles, rectangles, filled rectangles, + signs,** or **x signs.** The latter two choices increase performance.
- Moving Average -- Specify how to display moving average plots as: **Continuous**, **Dashed**, **Dotted**, or **Do not show area**.

Apply Apply your changes to the parameters.

Probe Methylation Status Setup

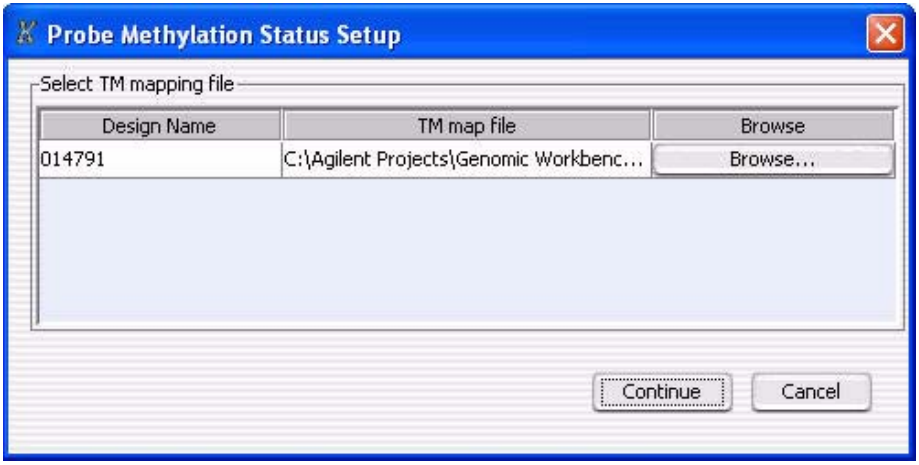


Figure 59 Probe Methylation Status Setup dialog box

Purpose: To associate Tm (melting temperature) map files with the design files for the arrays in the experiment or to make sure the Tm information is already in the design file.

To open: Click Analysis, then mark Apply for the Probe Methylation command

Design Name Shows the names of the design files that are already in the experiment

Tm Map File Click **Browse** to find the Tm map file to associate with the design file. This is necessary only if the design is a custom design or if Tm is still not available after an attempted update.

If the design is an Agilent Catalog array, the Tm information is available within the file as long as it has been updated. If the design file has not been updated, the Browse button is active. If it has been updated, the Tm Map File option says Available. If not, update it before continuing.Continue

After clicking, the methylation algorithm produces the Z-score results.

Cancel Closes the dialog box without generating a report.

Probe Methylation Report Dialog

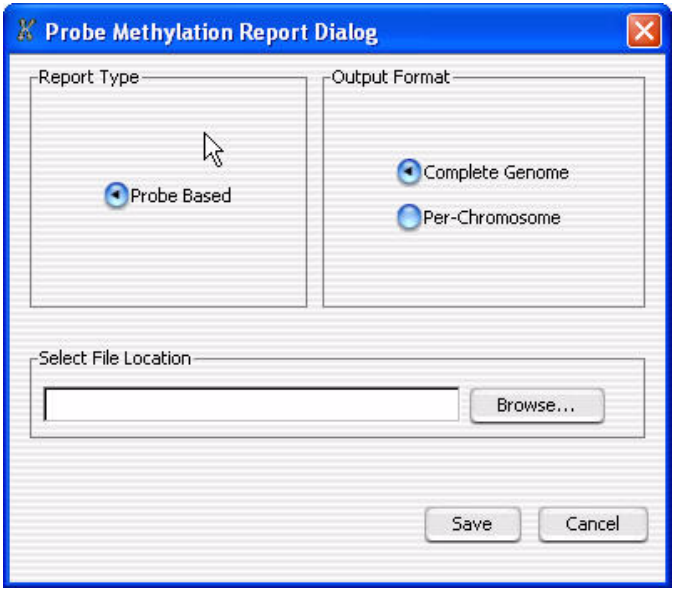


Figure 60 Probe Methylation Summary Setup dialog box

Purpose: To select whether the results are reported for the complete genome or for individual chromosomes. See "Report Format" on page 195.

To open: Click Reports, then click Generate Report.

Report Type All methylation reports are probe-based.

Output Format Complete Genome – Produces report as .xls file for entire genome

Per-Chromosome – Produces report for each chromosome as .txt file. Open in spreadsheet program to see headers properly aligned.

Select File Location Click **Browse** to select a folder to contain the file, which you must name.

Scroll to Column

Figure 61 Scroll to Column dialog box

Purpose: This dialog box allows you to select a column. The program then scrolls the tab so that you can see the selected column.

To open: Right-click a column heading in Tab View, then click Scroll To Column... in the shortcut menu.

Select column Lists the columns available in the currently selected tab. Select the one you want to view.

OK Scrolls the current tab so that you can see the selected column.

Cancel Closes the dialog box.

Search probes in eArray

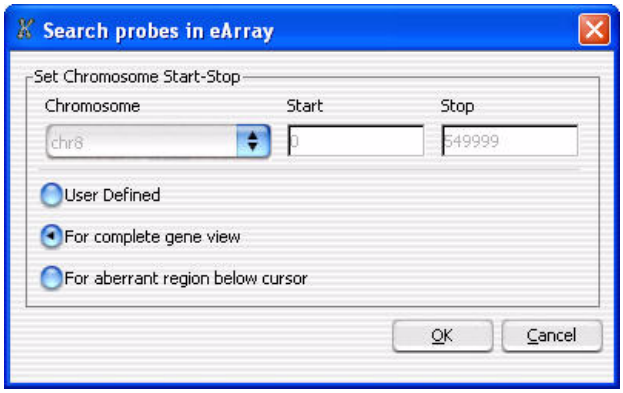


Figure 62 Search probes in eArray

Purpose: To select the probes you want to update in eArray

To open: Right-click Gene View, and click **Search probes in eArray**.

Select a chromosome and a region in Chromosome View for selecting the probes related to the genes in this region.

User Defined Select to choose the region from which the probes to be searched in eArray will be selected. The chromosome selection list and the Start and Stop positions on the Y axis are activated when this option is selected.

For complete gene view All the probes related to the genes in Gene View will be searched.

For aberrant region below cursor Selects those probes for the genes that appear just below where the cursor sits in Gene View. Not operational in Methylation (CH3) Analysis.

Chromosome If you select User Defined, you can select a different chromosome than had been selected before opening this dialog box.

Start/Stop If you select User Defined, you can type in Start and Stop positions for defining the region contained the genes to be in the list.

Select Color (Edit Array Color) — Swatches Tab

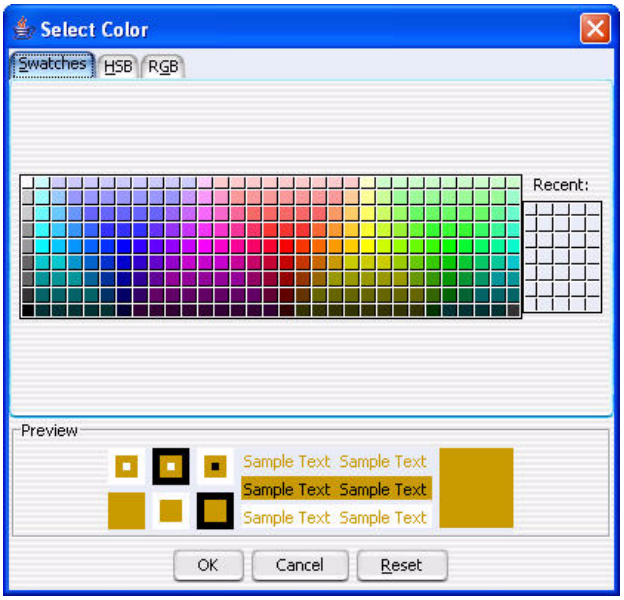


Figure 63 Select Color — Swatches Tab

Purpose: To select a color for each array based on color samples (swatches)

To open: Right-click on an array in an experiment, click Edit Array Color and click Swatches tab.

HSB Tab Choose colors based on an HSB schema (Hue, Saturation, and Brightness or Value). See

RGB Tab Choose colors based on an RGB schema (Red-Green-Blue).

Recent: Choose a recent color selection.

Reset Click to return HSB or RGB values back to default values.

Select Color (Edit Array Color) — HSB Tab

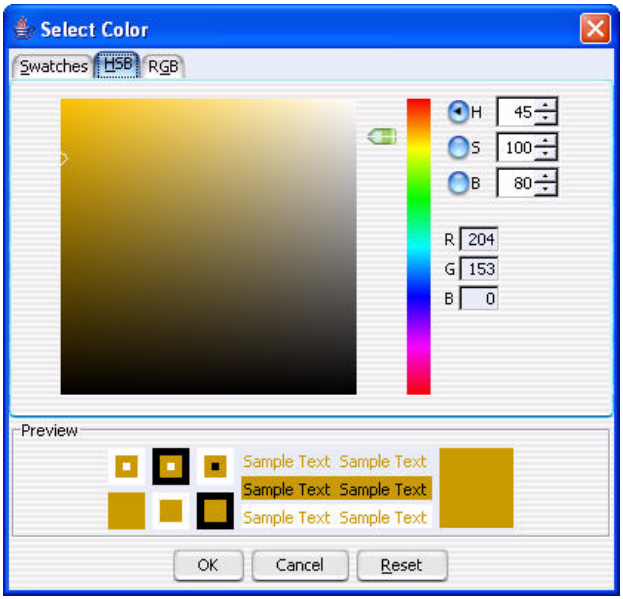


Figure 64 Select Color — HSB Tab

Purpose: To select a color for each array based on an HSB schema (Hue, Saturation, and Brightness)

To open: Right-click on an array in an experiment, click Edit Array Color and click HSB tab.

Hue Click the H button, and move the slider up and down, or go up and down the list of numbers, to select the hue or color of the array.

Saturation Click the S button, and move the slider up and down, or go up and down the list of numbers, to select the saturation level for the color.

Brightness Click the B button and move the slider up and down, or go up and down the list of numbers, to select the brightness level for the color.

RGB Numbers Reflect the amount of red, green and blue in the resulting color.

Reset Click to return HSB values back to default values.

Select Color (Edit Array Color) — RGB Tab

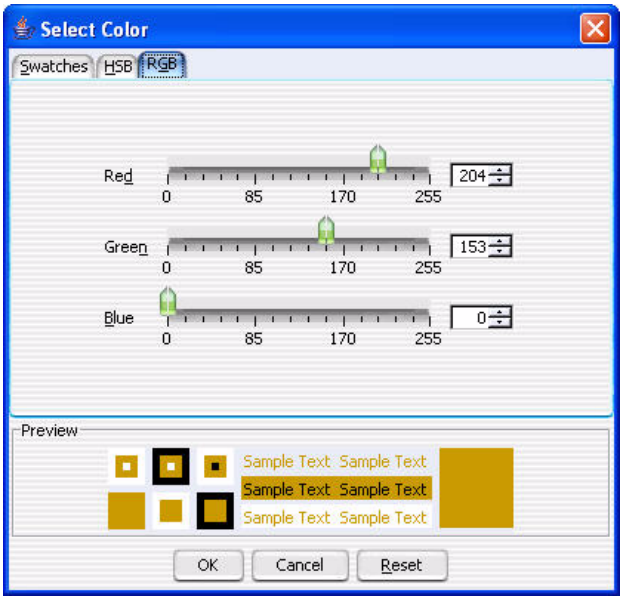


Figure 65 Select Color — RGB Tab

Purpose: To select a color for each array based on an RGB schema (Red- Green- Blue)

To open: Right- click on an array in an experiment, click Edit Array Color and click RGB tab.

Red Move the slider to change the amount of red in the color. Or, click the up or down arrow to select a number.

Green Move the slider to change the amount of green in the color. Or, click the up or down arrow to select a number.

Blue Move the slider to change the amount of blue in the color. Or, click the up or down arrow to select a number.

Reset Click to return RGB values back to default values.

Select data type for experiments

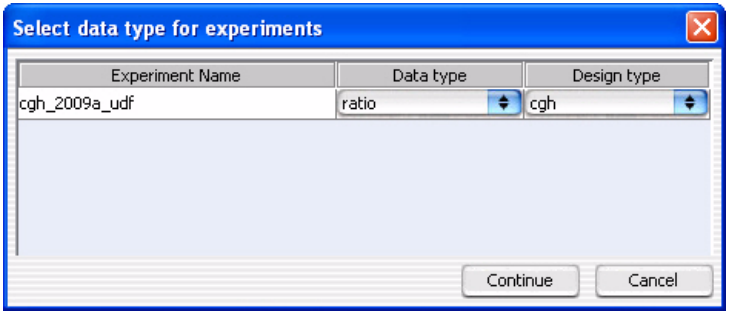


Figure 66 Select data type for experiments dialog box

Purpose: Allows you to specify the mathematical form of the data in an imported UDF file, and its associated application type. See "To import a UDF file" on page 16.

To open: In the **Home** tab, click **Import > Array Files > UDF File...** In the dialog box that appears, select the desired UDF file, then click **Open.**

Experiment Name By default, the experiment name is the name of the imported UDF file. To change the name, double-click it, then edit it as desired.

Data Type Select the mathematical form of the array data in the UDF file. The options are:

- **ratio**
- **$\log_2$ ratio**
- **$\log_{10}$ ratio**
- **ln ratio** (base e)

Design type Select the application type (CGH, ChIP, or expression, for example) associated with the array data in the UDF file.

Continue Accepts your selections, and goes to the next step in the UDF import process.

Cancel Cancels the UDF import.

Set genome build and species for Axon design files

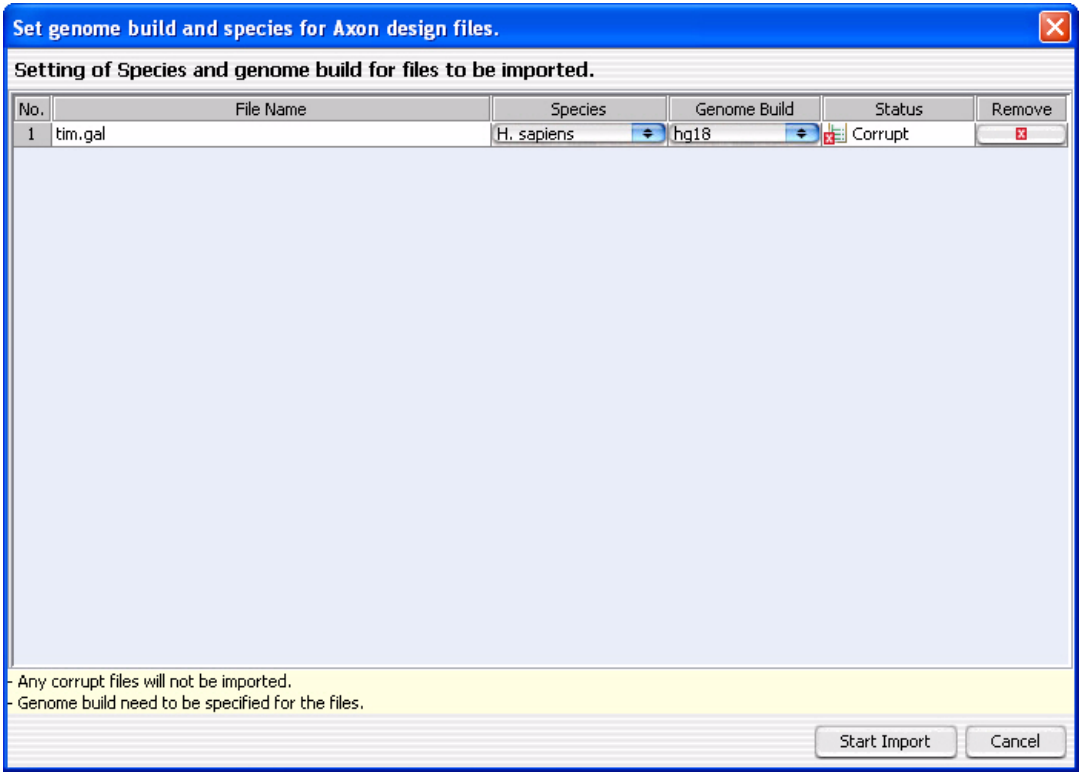


Figure 67 Set genome build and species for Axon design files dialog box

Purpose: Allows you to set the species and genome builds associated with imported Axon design file(s), and to remove specific designs files from the import, if necessary. See "To import GenePix/Axon design files" on page 20.

To open: In the **Home** tab, click **Import > Design Files > Axon File...** In the dialog box that appears, select at least one Axon design file, then click **Import.**

No. An index number within the dialog box for each Axon file.

File Name The names of each Axon design file selected for import.

Species The species associated with each design file. If a species is incorrect, select the correct one from the appropriate list.

Genome Build The genome build associated with each of the design files. If a genome build is incorrect, select the correct one from the appropriate list.

Status The status of the file is one of the following:

- **Valid** – The file is a new file that can be imported.
- **Overwrite** – The file is a valid design file, but when you import it, it will replace an existing design that has the same name.
- **Corrupt** – The file failed validation. When you start the import process, the program ignores the file.

Remove Click  to remove a specific design file from the list. This can be useful if you select a design file in error, or if you do not want to overwrite an existing one.

Start Import Imports the file(s) and closes the dialog box.

Cancel Cancels the import and closes the dialog box.

Show/Hide Columns

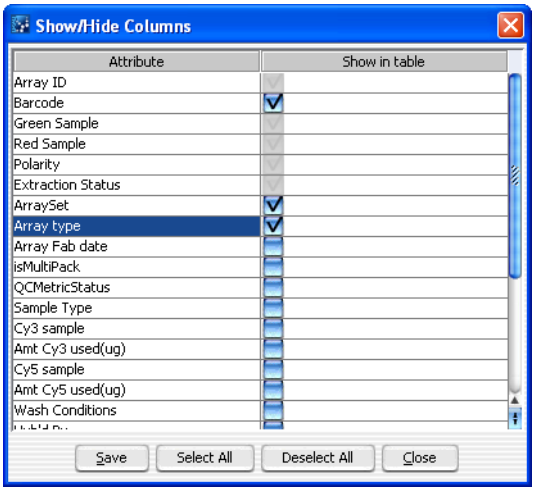


Figure 68 Show/Hide Columns dialog box

Purpose: Used to select the attributes that will be displayed in the Sample Utilities tab.

To open: This dialog box appears when you press the Show/Hide Attributes button at the bottom of the Sample Utilities tab.

All available attributes are shown in the Attributes column. Attributes with a check-mark next to them will be displayed in the Sample Utilities tab for each sample. To select an attribute for display, click the Show in Table box next to it. To de-select a currently-selected attribute, click the Show in Table box again.

Save Saves the current list of selected attributes and updates the Sample Utilities table based on the selections.

Select All Selects all the attributes in the list.

Deselect All Removes all check marks from attributes in the list.

Close Closes the dialog box. If changes have been made, the program asks if you want to save your changes before closing.

Track

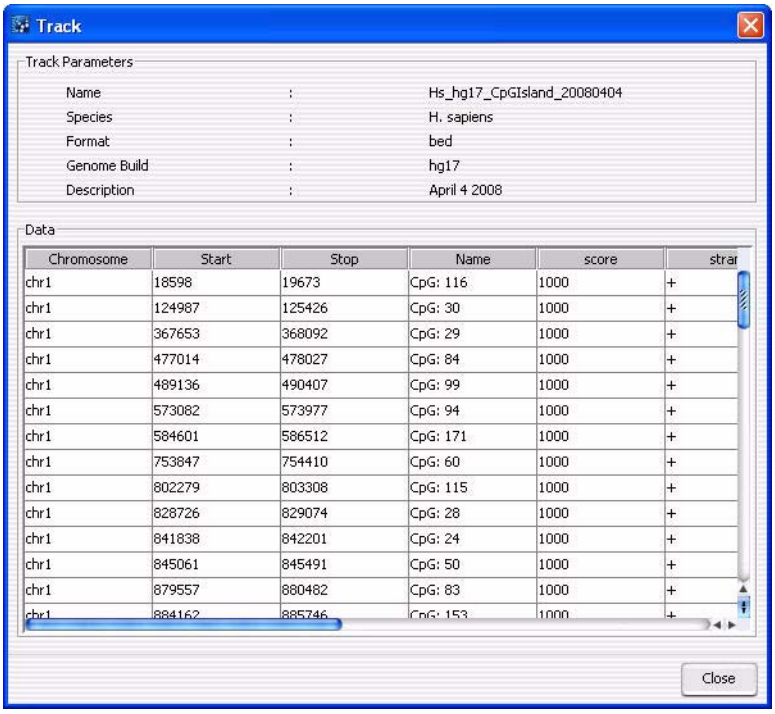


Figure 69 Track details

Purpose: This dialog box allows you to view the chromosome locations in the track.

To open: Click the **Details** link next to the desired track in the **Tracks** tab of the Preferences dialog box. See "Preferences – Tracks Tab" on page 174.

Track Parameters These parameters appear:

Parameter	Description
Name	The name of the track.
Species	The species to which the track applies.

Parameter	Description
Format	The format of the track data. DNA Analytics supports the BED format.
Genome Build	The specific genome build of the species to which the track applies.
Description	Descriptive text saved with the track.

Data Tracks must contain entries for at least these four columns in the table:

Column	Description
Chromosome	The name of the chromosome
Start	The first base pair of the particular feature in the chromosome.
Stop	The last base pair of the particular feature in the chromosome.
Name	The name of the feature. This name appears alongside the defined region for the feature.

The other columns are additional BED track file columns that can appear for some tracks. DNA Analytics does not render these.

Close Closes the Track dialog box.

UDF Import Summary

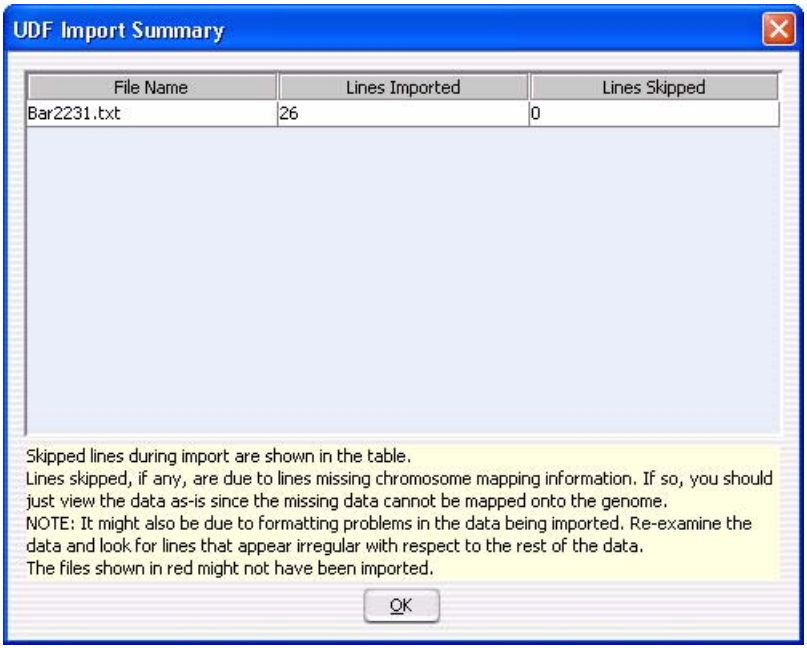


Figure 70 UDF Import Summary dialog box

Purpose: Reports how many lines of data were successfully imported from a UDF file, and how many lines were skipped. Skipped lines can be caused by missing chromosome mapping information, or improper formatting of the UDF file.

To open: Import a UDF file (see "To import a UDF file" on page 16). This dialog box appears after you map the columns of the UDF file.

Table Displays the file name of the imported UDF file, the number of lines that were successfully imported, and the number of lines, if any, that were skipped during import. If many lines were skipped, re-examine the data for improper formatting or missing chromosome mapping information.

OK Closes the dialog box.

Universal Data Importer - Map Column Headers

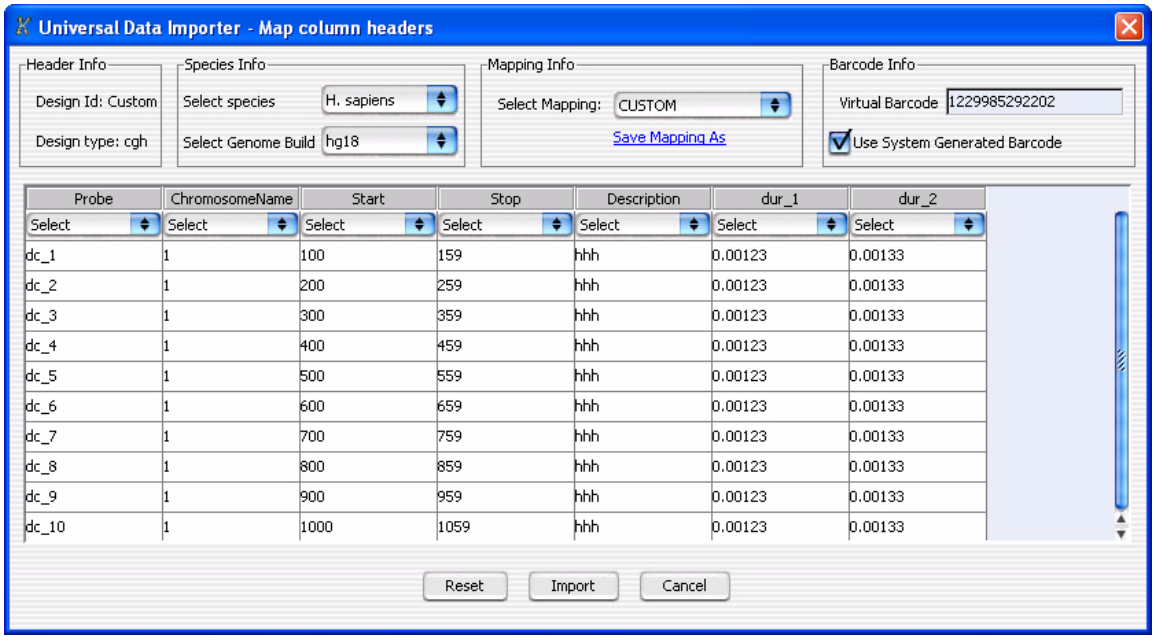


Figure 71 Universal Data Importer - Map Column Headers dialog box

Purpose: Allows you to set up a universal data file (UDF) for import. You specify several properties associated with the UDF, and identify the contents of each column of data in the file. You can also save column mappings for re-use.

To open: As you go through the UDF import process (see "To import a UDF file" on page 16), in the Select data type for experiments dialog box, click **Continue.** See "Select data type for experiments" on page 185.

Species Info

Select Species Select the species associated with the array data in the UDF. The program supports these species:

- *M. musculus*
- *H. sapiens*

- *R. norvegicus*

Select Genome Build Sets the species-specific build to use.

Mapping Info

Select Mapping Applies a previously saved column map to the current UDF. A column map identifies the contents of each column of data. To create a new column map for the current UDF, select **CUSTOM.**

Save Mapping As Saves the column map under a new name. Opens an Input dialog box, where you can type a name for the new map.

Barcode Info

Virtual Barcode A number that uniquely identifies the data in the UDF. Typically, an Agilent microarray slide has a physical barcode that enables Genomic Workbench to track the data from the slide as it goes through the steps of an analysis workflow. A "virtual" barcode is, by default, a system-generated ID that serves the same purpose for data from UDFs. You can also specify a virtual barcode of your own choosing.

Use System Generated Barcode By default, the virtual barcode assigned to the array data in a UDF is a number that is generated internally by the program. To specify a virtual barcode of your own choosing, clear **Use System Generated Barcode,** then type a new number in **Virtual Barcode.**

Table

This table allows you to identify the contents of the columns of data in the UDF. The first row of the table lists the column heading information from the UDF. The second row contains lists of labels that you apply to each column, and the rest of the table displays lines of data from the UDF. If the UDF contains data from Agilent CGH arrays, the column headings will exactly match the labels in the lists.

In the list below each column heading, select the applicable label. You must use each of the labels exactly once, except LogRatio, which you can use more than once. These labels are available:

Label	Description
ProbeName	The column contains names of probes.
ChrName	The column contains names of chromosomes.
Start	The column contains the first chromosomal location to which each probe is designed.
Stop	The column contains the last chromosomal location to which each probe is designed.
Description	The column contains text annotation related to the probe.
LogRatio	The column contains array data values that correspond to each probe. You can use this label more than once.

NOTE If you select a saved column mapping, then change or reset the column labels in the table, the program changes or resets the saved column map as well.

Reset Clears all the column labels in the second row of the table. If you have selected a saved column mapping, this command also clears the labels in the saved map.

Import Imports the UDF file with the specified parameters, and opens the UDF Import Summary dialog box (see "UDF Import Summary" on page 191).

Cancel Cancels the import and closes the dialog box.

Report Format

The Methylation Report contains all the statistical results to help you assess if there are methylated or unmethylated regions in the genome. The report is produced in .xls format for the complete genome or in .txt format for each individual chromosome. Use a spreadsheet program to open the report(s).

To learn how to interpret the values in the ZScore columns see "Methylation Detection and Visualization Algorithms Overview" on page 198.

Methylation Reports contain the following columns, in this order:

Column	Description
CytoBand	Accepted name of the cytoband for each section of chromosome
ChrName	Name of the chromosome (for example, chr17)
ProbeName	Name of the probe (for example, P)
Start	The first base pair of the chromosomal location to which the probe binds
Stop	The last base pair of the chromosomal location to which the probe binds
Description	Name or phrase for the type of probe (for example, promoter)
GeneNames	Names of each gene in whose region probes are located
"Name of CpG Island Track"	Name of each CpG Island in whose region probes are located
Combined ZScore	Combination of both the methylated and unmethylated ZScores. The higher the positive combined ZScore, the more likely the probe is methylated and vice versa.
ZScore_Methylated	Probe ZScore derived from right Gaussian. All of these values will be positive or small negative values.
ZScore_Unmethylated	Probe ZScore derived from left Gaussian. All of these values will be negative or small positive values.

Column	Description
logOdds	This number reflects the likelihood of a probe being methylated rather than unmethylated. The higher the positive value, the more likely it is methylated.
logRatio	Log ratio from the extracted FE image file for the probe
Tm	Melting temperature associated with the probe region

5
Statistical Algorithms

Methylation Detection and Visualization Algorithms Overview 198
Methylation Status Detection Algorithm 199
Visualization Algorithms 209
    Moving Average 209
    Triangular Smoothing 209
References 213

This section provides implementation details of the algorithms used in the CH3 application of Genomic Workbench 5.0. The CH3 application algorithms facilitate the statistical analysis of methylated genomic regions. The first section presents the main methylation detection algorithm. This is followed by a second section which details the visualization options available for methylation analysis.

Methylation Detection and Visualization Algorithms Overview

The CH3 module of Genomic Workbench provides algorithms for genomic methylation detection and visualization. The main methylation detection algorithm is described below. Information regarding algorithms that visualize the data follows.

Methylation Detection

Methylation of cytosines in DNA is an epigenetic modification that can play a role in the regulation of gene expression. Generally, methylation is associated with repression of gene expression. Much of the aberrant DNA methylation associated with disease is found in CG rich regions termed CpG islands. Agilent catalog CH3 arrays are specifically designed to assess the methylation pattern within CpG islands and gene promoter regions. The methylation detection algorithm is designed to be compatible with the Agilent protocol for DNA methylation analysis and affinity based methylated DNA enrichment methods only[1]. It is designed for two color assays wherein the green (Cy3) channel is comprised of input DNA and the red (Cy5) channel is comprised of the affinity enriched DNA. As a result of constraints imposed on the array design, the probes targeting genomic regions with varying GC content do not share a uniform melting temperature (T_m), which can result in compression of the log-ratios. The methylation detection algorithm allows the user to normalize the log-ratios for each probe based on its T_m and returns the methylation status of a probe. This sharpens the bimodal distribution observed with log-ratios alone. See 'Methylation Status Detection Algorithm' on page 199 for more information

Moving Average - Linear

The moving average algorithm sets a fixed window size around every point of interest. The points can be either the probe log-ratio scores determined prior to the methylation detection algorithm or the combined scores resulting from application of the methylation detection algorithm. The average of that point and neighboring points within the window boundaries are reported as the value for that point. See 'Visualization Algorithms' on page 209 for more information

Moving Average -Triangular Smoothing

The triangular smoothing algorithm is a shaped smoothing algorithm based on a moving average. As the moving average centers on a point of interest, a maximum weight is applied and that weight falls off with increasing distance. The point of interest is then adjusted as the weighted mean of itself and neighboring points. See 'Triangular Smoothing' on page 209 for more information

Methylation Status Detection Algorithm

Methylation of cytosines in DNA is an epigenetic modification that can play a role in the regulation of gene expression. Epigenetic methylation is found predominantly in genomic areas of increased GC content, known as CpG islands. Agilent human methylation arrays specifically query approximately 25,000 CpG islands in the human genome in addition to gene promoter regions. The Agilent methylation protocol uses antibody enrichment for 5-methylcytosine, resulting in a log-ratio of the relative abundance of genomic fragments with methylation to those without.

Although the log-ratio scores give information about the relative abundance of the genomic fragments, individual specificities of the probes in each log-ratio score vary by melting temperature. Differences in melting temperature can therefore decrease probe specificity. Normally, the probes on Agilent arrays are chosen in such a way as to minimize probe-to-probe melting temperature differences.

Within these bins, direct probe comparisons can be made. When comparing probes by log-ratio scores within each bin, a bimodal distribution is observed; a mode comprised of those probes which were not methylated, and a mode comprised of methylated probes. A bi-modal distribution is also observed in per-T_m *Z*-normalized scores of the data. The determination of methylation status of a probe is therefore accomplished in a T_m-dependent manner. The binning process is a way to compare log-ratios between probes of similar melting temperatures to determine which probes are methylated.

In order to apply parametric statistical techniques, the bimodal distribution is fitted to a set of Gaussian curves, each of which requires only a limited number of parameters to accurately fit the data. Normalized *Z*-scores from the Gaussian distributions are used to effectively judge the methylation status of a given probe on the array.

Purpose

Methylation status detection is a visualization and report analysis which determines the confidence score for probes which query known CpG islands and gene promoter regions. The analysis uses a combination of log-ratio scores from competitive hybridization methods and the range in melting temperatures between probes located in CpG islands to determine the probe methylation status.

Use Methylation status detection is used to identify probes likely to query methylated genomic regions in order to establish a link between genomic methylation and gene expression. It is generally observed that CpG islands, measured on the Agilent catalogue CH3 design array, are either fully methylated or fully unmethylated. This algorithm is used to identify methylation status indicators for probes. These probes can then be visualized in a genomic context - for example, CpG islands within the genome.

Algorithm Probes are first binned by their melting temperature. For each bin, Gaussian fits are applied using one of three models. The probe log-ratios are fitted to Gaussians using a local searching algorithm called hill climbing[2]. *Z*-scores and *p*-values derived from the Gaussian data gives probabilities and confidence values for methylated and non-methylated probe populations. A methylation logOdds is then calculated which gives the relative probability that a probe is more likely methylated than unmethylated.

NOTE The logOdds score indicates how likely a probe is to be methylated by comparing the *p*-values from the methylated and unmethylated populations. If it is as likely to be unmethylated as it is to be methylated, then this value is zero.

Step 1: Create a binned distribution of probes The algorithm orders probes into discrete melting temperatures in order to compare log-ratio scores between genomic fragments enriched by methylation and those without methylation.

1. Probes are binned according to their melting temperature. Each bin is 1ºC wide.
2. The Gaussian distributions are then fitted to the probe log-ratios within each T_m bin. This is done by further dividing the probe log-ratios into sub-bins (the default size of these bins are 0.1ºC) to generate the observed distribution of signals as shown in Figure 5-1.

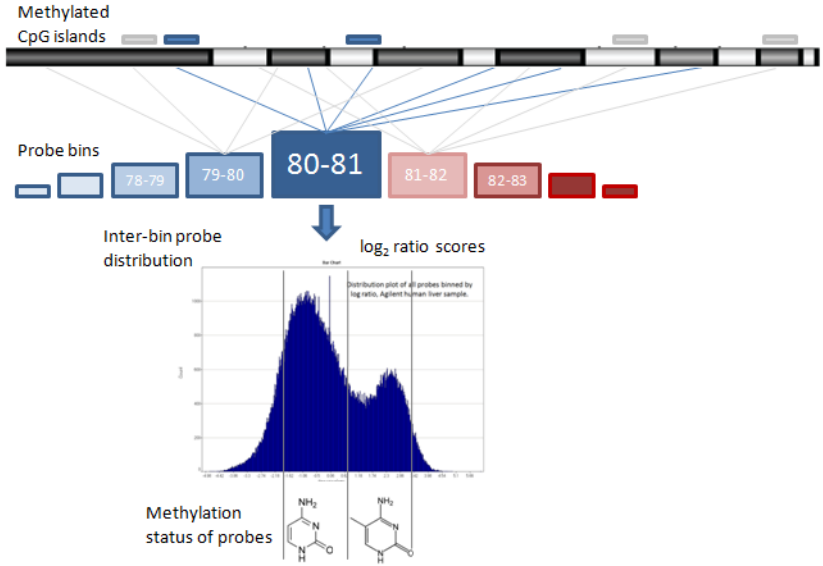


Figure 5-1 Overview of the methylation status algorithm procedures.

Step 2: Fit Gaussian curves to the binned data

For each bin of 1C, Gaussian fitting is applied in one of three models depending on whether the data is observed to be unimodal or multimodal:

- If the data is unimodal, then 1 Gaussian distribution is fit to the data.
- If the data is bimodal (the most common fit for methylation data), then 2 overlapping Gaussian distributions are fit to the data.
- If the data is trimodal, then 3 overlapping Gaussian distributions are fit to the data.

NOTE

One Gaussian goodness of fit can exceed two Gaussians if, for example, no methylated probes were pulled down during the antibody enrichment process or if the number of methylation events is exceedingly small. If this occurs, Genomic Workbench will automatically fit one Gaussian to the data and a message of "SINGLE GAUSS" will appear in the report. The *Z*-score in the visualization and report will then be calculated from the single Gaussian (*i.e.* how likely the probe is to be methylated) instead of a combined *Z*-score (which is reported from models using either two or three Gaussians).

The process of fitting a Gaussian curve to the data consists of two steps:

1 Initial parameter approximation

2 Improving the model parameters using a local search procedure called random hill climbing[2]

Initial parameter approximation

For each Gaussian curve the parameters to be fit are the mean (μ), the standard deviation (σ), and the mixture coefficient (the total Gauss integral, α). For each of the possible models, the initial parameters are approximated in the following ways:

- First Gaussian

  For any of the three models, the parameter estimations are done in the following ways:

1 The mean is estimated by the maximum bin.

2 The standard deviation is estimated using the IQR of the distribution as inferred from the data left of the mean.

3 The mixture coefficient is estimated from the max bin height.

- Second Gaussian

  If the data is bimodal or trimodal, the additional parameter estimations are done by the following:

1 The first Gaussian is subtracted from the distribution. The mean, standard deviation, and mixture coefficient is then estimated in the same way as the first Gaussian.

- Third Gaussian

  If the data is not well explained by either a unimodal or bimodal distribution, then a third Gaussian curve can be fit in order to allow better estimation of the first and second Gaussian parameters from the data. In this case, the following parameter estimations are:

1 The mean is estimated from the mean from the entire bin.

2 The standard deviation is estimated from the entire bin.

3 The mixture coefficient is estimated from the height of the bin used to estimate the third Gaussian mean.

Parameter improvement procedure

Because many probes are generally present in each T_m bin (in the order of thousands), a straight-forward local search procedure called hill climbing is used for parameter optimization. This procedure is shown to be highly accurate when fitting the Gaussian models to the data and is extremely fast.

In each step a random parameter is altered. This alteration is accepted if there is an improvement in the target fitness function. In order to converge to a (locally) optimal result, each random step gets shorter with the number of iterations. This reduction in step is made logarithmically by n, the number of iterations, and a default starting value of 100 (termed LB):

$$\frac{\log(LB+n)}{\log(LB)} \quad (1)$$

Convergence is achieved in general when a new step does not yield an improvement of the target fitness function. The procedure converges if there is no change in the fitness value for 1000 iterations or alternatively if a maximum of 50000 iterations have been performed.

Target fitness function

The default target function used is the Chi-squared goodness of fit. The chi-square fitness function is described using an observed distribution $OD_{i=1...N}$ and an estimated distribution function EF (for example, the convex combination of 2 Gaussians) as follows:

$$\chi^2 = \sum_{i=1}^{N} \frac{OD_i - EF_i^2}{EF_i} \quad (2)$$

Step 3: Calculating *Z*-normalization scores

After the Gaussian curve(s) have been fitted to the data, parametric statistical analysis can be calculated. The log-ratios from the T_m-binned probes may differ in value. Genomic Workbench therefore normalizes the scores using a measure from the Gaussian distributions themselves called a *Z*-score:

$$Z = \frac{log-ratio_{probe} - \mu_{bin}}{\sigma_{bin}} \quad (3)$$

Where μ_{bin} is the mean and σ_{bin} is the standard deviation of the Gaussian distribution. The *Z*-score is a measure of how far a given probe log-ratio score is from the mean of any of the Gaussian curves, given the standard

deviation of that curve. It therefore normalizes comparisons of probes from different bins by taking into account the Gaussian-fit curves of the different bins. Since the methylation model typically contains two Gaussians, there are *Z*-scores calculated for each one. The left Gaussian represents those probes with a lower log-ratio score and consists predominantly of unmethylated probes. The left Gaussian is also the major mode, since the majority of probes are generally unmethylated. The right Gaussian represents probes with a high log-ratio score and consists predominantly of probes enriched for methylation. The algorithm calculates the following *Z*-scores as statistical measures of whether or not a probe is methylated:

1. The *Z*-score derived from the left Gaussian. A negative or small positive value means that the probe is likely to be unmethylated.
2. The *Z*-score derived from the right Gaussian. A positive or a small negative value means that the probe is likely to be methylated.
3. The combined *Z*-score. This is the summation of the left and right Gaussian *Z*-score and reflects the location of a probe log-ratio value in relation to the Gaussian distribution(s) of probes with similar T_m. A strong positive value of the combined score means it is methylated, while a strongly negative value indicates the probe is unmethylated. There are varying degrees of positive, negative, and zero combined scores. See Figure 5-3 for more information.

In addition to the *Z*-score calculation, Genomic Workbench calculates a *p*-value that takes into account how much of a tail continues on past the log-ratio point on the Gaussian curve. Such a calculation is important because the tails of the Gaussian distributions overlap; hence it is often difficult to clearly assign a given probe log-ratio score to the left or the right Gaussian if there is a pronounced valley between the two distributions. The following *p*-values are calculated for the usual case of two Gaussians:

1. The *p*-value derived from left Gaussian. This is denoted p_M in the report, and is the confidence at which the non-methylation null hypothesis is rejected (*i.e.* the confidence at which we can call the probe methylated).
2. The *p*-value derived from the right Gaussian. This is denoted p_U in the report, and is the confidence at which the methylation null hypothesis is rejected (*i.e.* the confidence at which we can call the probe unmethylated).

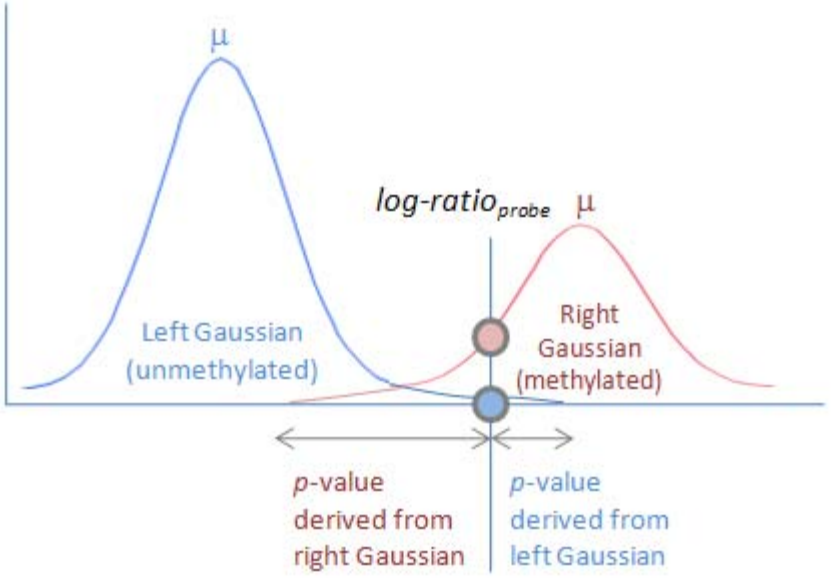


Figure 5-2 Calculating *Z*-scores and *p*-values from two overlapping Gaussian distributions.

Finally, by utilizing the bimodal distribution of log-ratio scores between probes of similar T_m, a logOdds score of the probe is calculated. This final logOdds reflects the likelihood of a probe to be methylated rather than unmethylated and is calculated as ω:

$$\omega = -\log\left(\frac{p_M}{p_U}\right) \quad (4)$$

Interpreting the results

As shown in Figure 5-2 above, the decision boundary for deciding whether a probe is methylated or unmethylated often comes from log-ratio scores that fall in a valley between two Gaussians. Although a ratio of the contribution of each Gaussian can be established by calculating the logOdds score ω, there is a region where ω approaches a value of 1 and a methylation status decision cannot be made. In other words, there is a threshold on the logOdds score which decides the methylation status call. This threshold is not fixed and can be determined by the user after inspecting the output in the context of a project.

Visualization Genomic Workbench can display a moving average for both the probe log-ratio scores and the *Z*-normalization scores from the methylation status algorithm. The *Z*-scores used in the moving average visualization are the combined *Z*-scores and reflect the position of the probe log-ratio score on both Gaussian curves as shown in Figure 5-3.

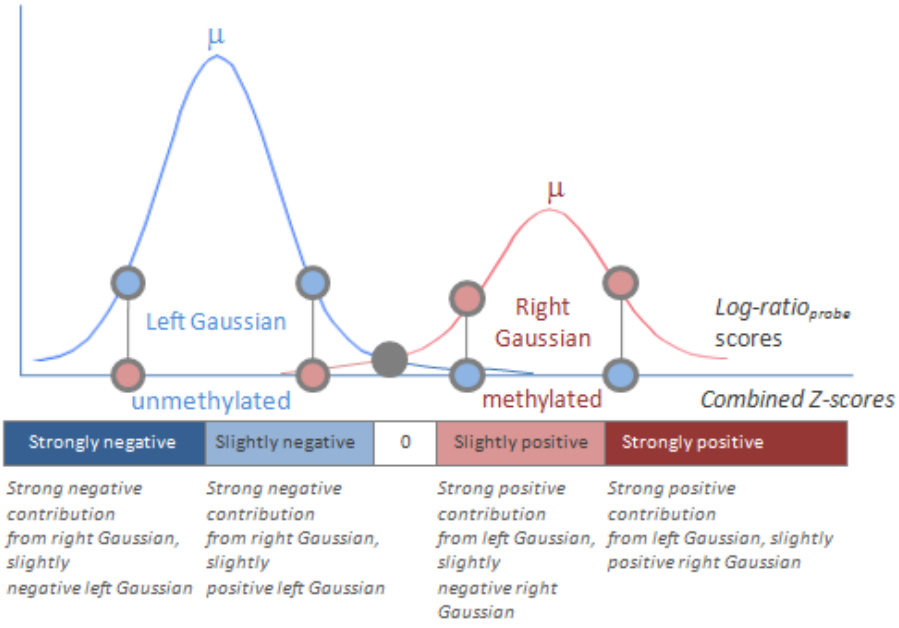


Figure 5-3 Effect of probe log-ratio score placement on a mixture of Gaussians - *Z*-normalization score result ranges.

The following screenshot displays the Gene View results of the methylation status algorithm for Chromosome 21. In the left panel, the Z-score moving average is displayed. There is no defined cut-off value for determining whether a probe or probe region is methylated. The middle panel shows the moving average for the probe log-odds ratios with genomic tracks annotated. The right panel displays the CpG island track. Note that the probes fall specifically within CpG islands and upstream of genomic tracks such as coding regions.

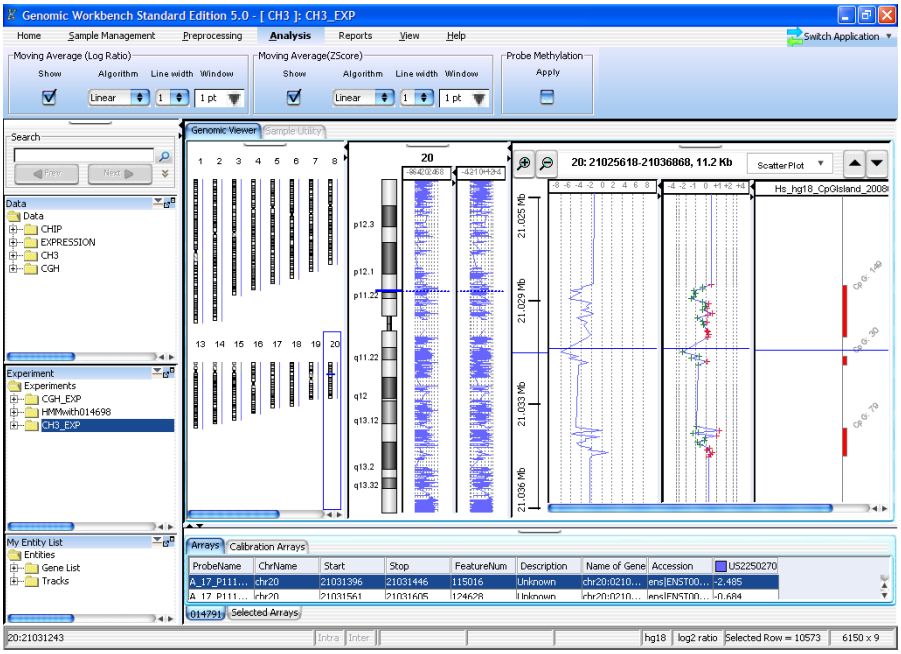


Figure 5-4 Output of methylation status detection algorithm. On the right side of Genomic Workbench, Gene View displays the moving average for both the *Z*-score (left panel of Gene View) and the probe log-ratios (middle panel of Gene View). Human CpG islands are displayed as a track along the right side of the moving averages. Note that the probe log-ratios which are only slightly positive in this region (colored red) may take on negative combined *Z*-scores (colored by sample).

Output Report An output report is generated per-chromosome. The report contains the following statistical measures:

Table 1 Methylation Output Report fields

Reported parameter	Meaning
Z-score unmethylated	*Z*-score derived from the left Gaussian. This is the probability that the observed value for a probe comes from the unmethylated population.
Z-score methylated	*Z*-score derived from the right Gaussian. This is the probability that the observed value for a probe comes from the methylated population.
Combined Z-score	Summation of the methylated and unmethylated *Z*-scores.
p_M	The *p*-value derived from the left Gaussian. This is the confidence at which the non-methylated null hypothesis is rejected (*i.e.* the confidence at which we can call the probe methylated).
p_U	The *p*-value derived from the right Gaussian. This is the confidence at which the methylated null hypothesis is rejected (*i.e.* the confidence at which we can call the probe unmethylated).
logOdds	The logOdds score ω; the likelihood that a probe is methylated rather than unmethylated.

Visualization Algorithms

After applying the methylation algorithm, it is useful to visualize the general trend of the pre-algorithm probe log-ratios and post-algorithm combined *Z*-scores by using a moving average line plot. In order to facilitate the visualization of large datasets, Genomic Workbench employs a sliding window across the genome to smooth the data points for detailed examination of the data points. These functions are available independently from the main methylation algorithm. Changing the window size in the visualization algorithms will therefore not affect the methylation calls.

In order to reduce the complexity of the data from probe-to-probe noise, two smoothing algorithms can be applied to the data. One algorithm applies linear smoothing (the moving average algorithm), and the other applies a triangular smoothing function. These algorithms use a sliding window of fixed size, set in the user interface (UI) and can be applied to the log ratios, the *Z*-score output from the methylation algorithm, or both.

The moving average visualization algorithm sets a fixed window size around every point of interest. The average of that point and neighboring points within the window boundaries are reported as the value for that point. The moving average algorithm is available in two modes: Linear and Triangular Smoothing.

Moving Average

To compute a moving average, probe log-ratios or combined *Z*-scores are averaged over a small subset of points in the genome. This moving average window, w, may be simply a number of adjacent measurements or it may be over a positional window (such as every megabase).

Triangular Smoothing

The triangular smoothing algorithm is a shaped smoothing algorithm based on a moving average. As the moving average centers on a point of interest, a maximum weight is applied and that weight falls off with increasing distance. The point of interest is then adjusted as the weighted mean of itself and neighboring points.

Purpose When visualizing or analyzing array data, it is common to smooth the data using a moving average. However, the moving average approach is not the optimal means of reducing the noise associated with each independent point, since it can minimize log-ratio changes or methylation *Z*-scores and obscure individual points. A good compromise is achieved between reducing the noise of the individual points, and still remain sensitive to true localized, or small-scale, variations in the data by using triangular smoothing.

Use The smoothing functions are used for visualization purposes only and does not affect the methylation algorithm. However, application of the Moving Average (*Z*-score) is applied after application of the methylation algorithm, and will therefore smooth the methylation *Z*-scores visualized in the user interface.

Algorithm The number of neighboring points to be used for smoothing depends on the type of moving average. If we have point input (Pt input) from UI, the number of points that are averaged is kept fixed, say at 3, 5, 7, 9, or 11 points, and each point is given equal weight. In other cases, a window of constant width (specified from UI in Mb or Kb) moves across the data and centers on the point of interest. All points within its range are averaged to yield the moving average value for the point.

NOTE In linear smoothing, the user defined point input (Pt input) uses the same number of probes as in triangular smoothing. However, an equal number of probes are taken in each direction (to the left and right of the center probe) to compute the moving average. This might not be ideal if the probes are placed at varying distances.

There are two potential problems in linear smoothing:

- First, using a fixed window width (by choosing a base pair from the UI) causes a variable number of probes to be averaged at each smoothed point. Therefore the degree of averaging varies from probe to probe, depending on how many probes happen to be in the fixed-width window. Since varying numbers of measurements contribute to each smoothed point, the degree of statistical noise reduction also varies for each point. This can complicate the error analysis.

- The second problem with linear smoothing arises on so-called "zoom-in" arrays, in which some genomic regions of interest are covered more densely than neighboring regions. For such arrays, the appropriate window size can vary greatly between the different genomic regions. Smoothing windows that are appropriate for sparsely tiled regions will obliterate all structure in densely tiled regions, whereas windows appropriate to densely tiled regions will perform practically no averaging at all in sparsely tiled regions.

These problems with linear smoothing can be avoided in triangular smoothing by using smoothing windows containing a fixed number of probes, regardless of the total range of sequences those probes span. This respects the fact that nearby probes are more relevant than distant probes to the average at any point.

The concept is illustrated in Figure 5-5. Fixed count smoothing includes the same number of points in each average, but weights probes far from the averaged point as much as points near to it. In triangular smoothing with Pt input, a symmetric window around the averaged point is enlarged until it contains the number of points chosen for the fixed size window. These points may be on one side of the averaged point or on both, depending on the probe density around the averaged point. These points are then weighted appropriately for triangular smoothing function, depending on their distance from the averaged point.

When user inputs number of points from the user interface, it uses a variable window width, which is chosen to be the smallest window, symmetrical about the averaged point, which includes the specified number of points. Figure 5-5 illustrates the application of this method for Pt input.

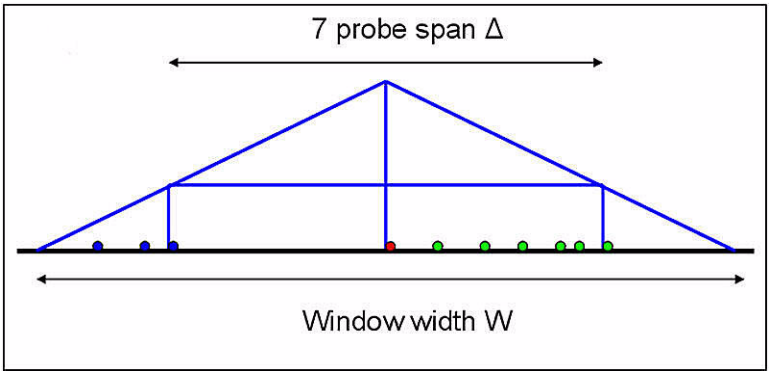


Figure 5-5 Triangular smoothing with Pt input.

Smoothing is applied to a region of varying probe density. The effective width of the smoothing window, W, depends on the length of the smallest symmetrical region (Δ) that includes the specified number of points. The weight given to each point is proportional to the height of the triangle at that point.

The weights assigned to the log-ratio or combined *Z*-score values of these probes are given by the equation:

$$w(x) = (W - |x|)/W^2 \quad \textbf{(1)}$$

where the effective window width, W, is determined by the span, Δ, of the symmetrical region spanning the specified number of points by the equation:

$$W = \Delta/(2 - \sqrt{2}) \quad \textbf{(2)}$$

Interpretation The smoothing algorithms affect the scatter plot in the chromosome and gene view. See Figure 5-4 for more information.

References

1 "Agilent Microarray Analysis of Methylated DNA Immunoprecipitation - Protocol, version 1.0" *Agilent Life Sciences and Chemical Analysis Protocols*. © 2008. <http://www.chem.agilent.com/Library/usermanuals/Public/G4170-90012_Methylation_Protocol.pdf>.

2 Russell, Stuart J., Norvig, Peter (2003). *Artificial Intelligence: A Modern Approach* (2nd ed.), Upper Saddle River, NJ: Prentice Hall, pp. 111-114, ISBN 0-13-790395-2

www.agilent.com

In This Book

This guide describes how to use the Methylation (CH3) Analysis application of Genomic Workbench to apply an algorithm that helps you assess the regions that are methylated or unmethylated.

First Edition, April 2009

Agilent Technologies