

Affymetrix[®] Microarray Suite

User's Guide

Version 5.0

For Research Use Only.

Not for use in diagnostic procedures.

Affymetrix Confidential

All rights reserved.

701099 Rev 1



Trademarks

Affymetrix[®], GeneChip[®], EASI[™], [™], [®], [™], HuSNP[™], GenFlex[™], Jaguar[®], MicroDB[®], 417[™], 427[™], 418[™], 428[™], Pin-and-Ring[™], NetAffx[™], CustomExpress[™] and Flying Objective[™] are trademarks owned or used by Affymetrix, Inc.

Microsoft[®] is a registered trademark of Microsoft Corporation.

Oracle[®] is a registered trademark of Oracle Corporation.

GeneArray[®] is a registered trademark of Agilent Technologies.

Limited License

EXCEPT AS EXPRESSLY SET FORTH HEREIN, NO RIGHT TO COPY, MODIFY, DISTRIBUTE, MAKE DERIVATIVE WORKS OF, PUBLICLY DISPLAY, MAKE, HAVE MADE, OFFER TO SELL, SELL, USE OR IMPORT PROBE ARRAYS OR ANY OTHER PRODUCT IS CONVEYED OR IMPLIED WITH THE PROBE ARRAYS, INSTRUMENTS, SOFTWARE, REAGENTS OR ANY OTHER ITEMS PROVIDED HEREUNDER. EXCEPT FOR CERTAIN ARRAYS AND REAGENTS DESIGNATED AS "ANALYTE SPECIFIC REAGENTS" (SEE APPLICABLE PACKAGE INSERT) WHICH ARE LICENSED FOR USE AS ANALYTE SPECIFIC REAGENTS OR RESEARCH USE, ALL PRODUCTS (INCLUDING THE PROBE ARRAYS, INSTRUMENTS, SOFTWARE, AND REAGENTS) DELIVERED HEREUNDER ARE LICENSED TO BUYER FOR RESEARCH USE ONLY. THIS LIMITED LICENSE PERMITS ONLY THE USE BY BUYER OF THE PARTICULAR PRODUCT(S), IN ACCORDANCE WITH THE WRITTEN INSTRUCTIONS PROVIDED THEREWITH, THAT BUYER PURCHASES FROM AFFYMETRIX OR ITS AUTHORIZED REPRESENTATIVE. THE PURCHASE OF ANY PRODUCT(S) DOES NOT BY ITSELF CONVEY OR IMPLY THE RIGHT TO USE SUCH PRODUCT(S) IN COMBINATION WITH ANY OTHER PRODUCT(S). IN PARTICULAR, (i) NO RIGHT TO MAKE, HAVE MADE OR DISTRIBUTE OTHER PROBE ARRAYS IS CONVEYED OR IMPLIED BY THE PROBE ARRAYS, (ii) NO RIGHT TO MAKE, HAVE MADE, IMPORT, DISTRIBUTE, OR USE PROBE ARRAYS IS CONVEYED OR IMPLIED BY THE INSTRUMENTS OR SOFTWARE, AND (iii) NO RIGHT TO USE PROBE ARRAYS IN COMBINATION WITH INSTRUMENTS OR SOFTWARE IS CONVEYED UNLESS ALL COMPONENT PARTS HAVE BEEN PURCHASED FROM AFFYMETRIX OR ITS AUTHORIZED REPRESENTATIVE. FURTHERMORE, PROBE ARRAYS DELIVERED HEREUNDER ARE LICENSED FOR ONE (1) TIME USE ONLY AND MAY NOT BE REUSED. THE PRODUCTS DO NOT HAVE FDA APPROVAL. NO PATENT LICENSE IS CONVEYED TO BUYER TO USE, AND BUYER AGREES NOT TO USE, THE PRODUCTS IN ANY SETTING REQUIRING FDA OR SIMILAR REGULATORY APPROVAL OR EXPLOIT THE PRODUCTS IN ANY MANNER NOT EXPRESSLY AUTHORIZED IN WRITING BY AFFYMETRIX IN ADVANCE.

Patents

Software products may be covered by one or more of the following patents: U.S. Patent Nos. 5,733,729; 5,795,716; 5,974,164; 6,066,454; 6,090,555; 6,185,561 and 6,188,783; and other U.S. or foreign patents.

Scanner products may be covered by one or more of the following patents: U.S. Patent Nos. 5,578,832; 5,631,734; 5,834,758; 5,936,324; 5,981,956; 6,025,601; 6,141,096; 6,185,030; 6,201,639; 6,207,960; and 6,218,803; and other U.S. or foreign patents.

Arrayer products may be covered by one or more of the following patents: U.S. Patent No. 6,136,269; and other U.S. or foreign patents.

Copyright

©2000, 2001 Affymetrix, Inc. All rights reserved.



Contents

CHAPTER 1	<i>Welcome</i>	21
<hr/>		
What's New In Microarray Suite 5.0		22
Statistical Expression Analysis Algorithm		22
Genotyping (GT) Analysis Viewer		23
Export or Import Batch Files		23
Publish Database Security		23
Conventions Used in This Guide		24
<i>Font Styles</i>		24
<i>Screen Captures</i>		25
<i>Additional Comments</i>		25
Online Documentation		25
Technical Support		26
Your Feedback is Welcome		26
CHAPTER 2	<i>Installing Microarray Suite</i>	29
<hr/>		
Microarray Suite 5.0		29
CHAPTER 3	<i>Microarray Suite Overview</i>	43
<hr/>		
Microarray Suite Walkthrough		43

Setting Up an Experiment	45
Hybridizing, Washing, and Staining a Probe Array	45
Scanning a Probe Array	46
Data Analysis	46
LIMS	48
Workflow Based Tracking	49
Data Mining	49
CHAPTER 4 <i>Getting Started</i>	53
<hr/>	
Starting Microarray Suite	53
Microarray Suite User Interface	54
Shortcut Bar	54
<i>Microarray Suite Shortcut Bar</i>	55
<i>Instrument Control Shortcut Bar</i>	58
<i>Analysis Settings Shortcut Bar</i>	59
<i>Report Settings Shortcut Bar</i>	60
Data File Tree	61
<i>Opening Files and Viewing File Information</i>	62
<i>Data File Tree Display Options</i>	63
Status Log	63
Configuring Microarray Suite	64
Data Storage	64
File Locations Settings	65
<i>Microarray Suite Files</i>	67
LIMS Filters	67
CHAPTER 5 <i>Setting Up an Experiment</i>	73
<hr/>	
Setting Up an Experiment in Disk Files Mode	74

Setting Up an Experiment in LIMS Mode	76
Registering a Sample	76
Defining an Experiment	77
Viewing or Editing Experiment Information	81

CHAPTER 6 *LIMS Sample History & Workflow Monitor* **85**

Sample History	85
Displaying Sample History	85
Displaying Sample History Process View	86
Renaming a Sample	88
Workflow Monitor	89
Opening the Workflow Monitor	89
Workflow Monitor Tabs	90
<i>Hybridization Tab</i>	91
<i>Scan Tab</i>	91
<i>Grid Alignment Tab</i>	91
<i>Cell Intensity Analysis Tab</i>	92
<i>Probe Array Analysis Tab</i>	93
<i>Publish Tab</i>	93

CHAPTER 7 *Controlling the Instruments* **97**

The Fluidics Station 400	97
Configuring the Fluidics Station	98
Priming the Fluidics Station	99
Setting Up a Fluidics Protocol	101
<i>Disk Files Mode</i>	101
<i>LIMS Mode</i>	102
Running the Fluidics Protocol	103
Resuming a Fluidics Protocol	103

Bypassing Steps in a Fluidics Protocol	104
Editing a Fluidics Protocol	105
The Scanner	107
Scanning a Probe Array	107
Stopping a Scan	110
Shutting Down the Scanner	111
CHAPTER 8 <i>Working with Images</i>	115
The Image Window	115
Commands in the Image Window Task Bar	117
Opening Multiple Image Windows	117
Viewing the Grid & Manually Adjusting Grid Alignment	118
Computing Cell Intensities	120
Reanalyzing an Image Data File	120
Viewing Images (*.dat)	121
Calculating Average Intensity for User Specified Probe Cells	121
Adjusting the Image Settings	121
<i>Making Adjustments to the Current Image Only</i>	122
<i>Making Global Adjustments to Subsequently Opened Images</i>	123
<i>Setting the Intensity Range</i>	123
<i>Selecting Color Settings</i>	124
<i>Selecting Display Coordinates</i>	124
<i>Specifying Highlight Colors</i>	126
Image Display Options	127
Intensity Display	127
<i>Measured Image View</i>	127
<i>Cell Image View</i>	128
<i>Difference Image View</i>	128
Viewing Probe Cell Data	129
Probe Cell Data: Gene Expression Assay	130

Probe Cell Data: Genotyping Assay	131
Probe Cell Data: Polymorphism & Mutation Detection Assay	132
Probe Cell Data: Mapping Assay	133
Probe Cell Data: Hybridization Assay	134
Viewing Probe Set Information	135
Viewing Probe Array Tiling Information	135
Viewing Probe Cell Data for a Highlighted Probe Set(s)	136
Highlighting Tiles on the Probe Array	138
Bookmarks	139
Creating a Bookmark	140
Viewing Bookmarks	140
<i>Commands in the Edit Bookmarks Dialog Box</i>	141
Outliers	141
Viewing Outliers	142
Masking Outliers	142
Unmasking Outliers	142
CHAPTER 9 <i>Gene Expression Analysis</i>	145
Overview of Absolute & Comparison Expression Analysis	146
Absolute Expression Analysis	146
Comparison Expression Analysis of an Experiment & Baseline	148
Expression Analysis Settings	150
Absolute Expression Analysis	151
Comparison Expression Analysis	154
<i>Selecting the Baseline</i>	154
<i>Running the Comparison Analysis</i>	155
Expression Analysis Tabular Data	159
Analysis Information Table	159
<i>Exporting Analysis Information</i>	161

Metrics Table	162
<i>Exporting Metrics Data</i>	163
Pivot Table	165
<i>Pivot Table Display Options</i>	166
<i>Obtaining Further Information</i>	167
<i>Exporting the Pivot Table</i>	168
Finding Probe Set Information	169
Sorting Probe Set Information	169
<i>Single Column Sort</i>	169
<i>Multiple Column Sort</i>	170
Reordering or Removing Probe Set Information	170
<i>Reordering Columns in the Metrics or Pivot Table</i>	170
<i>Reordering Rows in the Metrics or Pivot Table</i>	171
Hiding Probe Set Information	171
Expression Analysis Measured Images	172
Displaying Measured Images	172
Clearing Measured Images	173
Expression Analysis Graphs	174
Scatter Graph	175
<i>Plotting a Scatter Graph</i>	176
<i>Working with the Scatter Graph</i>	178
<i>Scatter Graph Options</i>	183
Series Graph	186
<i>Plotting a Series Graph</i>	188
<i>Working with the Series Graph</i>	188
<i>Series Graph Options</i>	189
Intensity Bar Graph	191
<i>Plotting the Intensity Bar Graph</i>	191
<i>Intensity Bar Graph Options</i>	192
<i>Clearing Intensity Bar Graphs</i>	195
Expression Report	195
Generating an Expression Report	195
Expression Report Components	199

<i>Statistical Expression Algorithm</i>	199
Expression Report Settings	202
<i>Housekeeping Controls</i>	203
<i>Spike Controls</i>	203
<i>Other Controls</i>	203
Expression Report View Options	204
<i>Find Feature</i>	204
Editing an Expression Report	204
<i>Replace Feature</i>	204
Saving an Expression Report	205
CHAPTER 10 <i>Mutation & Polymorphism Analysis</i>	209
<hr/>	
Overview	209
Mutation & Polymorphism Analysis Algorithms	210
Mixture Detection Algorithm	210
<i>Viewing Mixture Detection Algorithm Settings</i>	211
Consensus Rules Algorithm	212
Running a Mutation & Polymorphism Analysis	212
Sequence Analysis Window (SAW)	214
Sequence Analysis Window View Options	215
<i>Viewing Expanded Information</i>	215
<i>Viewing Nucleotides, Amino Acids, or Both</i>	217
<i>Viewing the Complement of a Selected Tile Number</i>	218
Comparing Nucleotide Data	219
<i>Selecting a Nucleotide Comparison Option and Highlight Color</i>	219
Comparing Amino Acid Data	222
<i>Selecting an Amino Acid Comparison Option and Highlight Color</i>	223
Hiding, Reordering, or Removing Sequences	225
<i>Hiding Tiles, Bases, or Rows</i>	225
<i>Reordering a Probe Array Type or Experiment</i>	226
<i>Removing Data, Tiles, or Rows</i>	227

Alignment Bars	227
Mutation & Polymorphism Analysis Graphs	228
HIV PRT Plus Intensity Graph	228
<i>Plotting the HIV PRT Plus Intensity Graph</i>	229
<i>HIV PRT Plus Intensity Graph Options</i>	229
p53 Score Graph	232
<i>Plotting the p53 Score Graph</i>	233
<i>p53 Score Graph Options</i>	233
Navigating the Sequence Analysis Window (SAW)	235
Go To a Specific Nucleotide or Amino Acid	235
Finding Specific Nucleotides or Amino Acids	235
<i>Moving Between Highlighted Positions</i>	236
<i>Moving Between User-Specified Nucleotide Positions</i>	237
Sequence Analysis Window (SAW) Display Options	238
<i>Options</i>	239
<i>Alignment Bar Colors</i>	240
<i>Defaults</i>	240
Editing Consensus Base Calls	240
Importing Files	241
Mutation & Polymorphism Analysis Reports	242
p53 Report	242
<i>Generating the p53 Report</i>	242
<i>Viewing User Modifiable Report Settings</i>	246
<i>Selecting an Imported Reference File</i>	246
HIV PRT Plus Report	247
<i>Generating the HIV PRT Plus Report</i>	247
<i>Viewing User Modifiable Report Settings</i>	250
<i>Selecting an Imported Reference File</i>	251

CHAPTER 11	<i>Genotyping Analysis</i>	255
Overview		256
Genotyping Algorithms		256
Block Algorithm		257
Genotyping Algorithm		257
Viewing the Genotyping Algorithm Settings		258
Running a Genotype Analysis		258
Nucleotide Analysis Window (NAW)		259
Nucleotide Analysis Window (NAW) View Options		260
<i>Viewing Mutations Only</i>		260
<i>Locating a Polymorphism</i>		261
<i>Hiding Data Table Rows</i>		262
<i>Removing an Experiment from the NAW</i>		263
<i>Sorting Information</i>		263
Find Feature		264
Genotyping Tabular Data		265
CYP450 Tabular Data		265
HuSNP™ Tabular Data		266
Exporting Genotype Data		267
<i>Displaying Two Columns in the Text File</i>		268
Genotyping Graphs		269
CYP450 Intensity Graphs		269
<i>Plotting CYP450 Intensity Graphs</i>		271
<i>CYP450 Intensity Graph Options</i>		271
HuSNP™ Intensity Graphs		273
<i>Plotting HuSNP Intensity Graphs</i>		273
Measured Images		274
<i>Displaying Measured Images</i>		274
Removing Graphs and Measured Images		275
Removing All Graphs or Images		275

Removing Selected Graphs or Images	276
Genotyping Analysis Reports	276
CYP450 Report	277
Mapping Report	280
Genotyping Analysis Viewer Report	282
<i>Selecting Cell Intensity Files for the GT Viewer Report</i>	282
CHAPTER 12 <i>Hybridization Analysis</i>	289
<hr/>	
Running a Hybridization Analysis	289
Hybridization Analysis Window (HAW)	291
Exporting Data	291
Copying Data	291
Removing Data	292
Find Feature	292
CHAPTER 13 <i>Batch Analysis</i>	297
<hr/>	
Opening the Batch Analysis Window	297
Selecting Files for Batch Analysis (Disk Files Mode)	299
Selecting the Baseline for Comparison Expression Analysis	300
Selecting Files for Batch Analysis (LIMS Mode)	302
Selecting a User Set for Expression Analysis	305
Selecting a Baseline for a Comparison Expression Analysis	306
Renaming Results from Batch Analysis	308
Renaming Results Automatically	308
Renaming Results Individually	309
Removing Files from the Batch Analysis Window	310
Running a Batch Analysis	311
Resuming a Batch Analysis	311

Exporting a Batch File	311
Importing a Batch File	313

CHAPTER 14 *Publishing Data* 317

Selecting the Publish Database	317
Specifying the Task	319
Publishing Options	322
Publishing Intensity Data	322
Overwriting Data	322
Publishing, Canceling, or Restarting a Task	323
Publishing the Task	323
Canceling a Task	323
Restarting a Publishing Task	323

CHAPTER 15 *Reports* 327

Report Settings	328
Expression Report	328
Expression Report Components	332
<i>Statistical Expression Algorithm</i>	332
Mapping Report	335
Mapping Report Components	337
CYP 450 Report	337
CYP450 Report Components	339
HIV PRT Plus Report	340
HIV PRT Plus Report Components	341
p53 Report	342
P53 Report Components	344

Genotyping Viewer Analysis Report	345
GT Viewer Analysis Report Components	346

CHAPTER 16 *Defaults* **351**

Viewing Default Settings	351
Image Settings	352
<i>Intensity Range</i>	352
<i>Color</i>	353
<i>Coordinates</i>	353
<i>Highlight Colors</i>	355
<i>Grid Overlay</i>	356
Database	356
File Locations	357
Experiment Data Stored on the Workstation	358
<i>Modifying a File Location</i>	358
Experiment Data Stored on the LIMS Server	359
Configuration	361
Analysis Settings	362
Prompt for Output File Name	362
Display Settings When Analyzing Data	363

CHAPTER 17 *Printing* **367**

Printing an Image Data or Cell Intensity File	367
Printing an Analysis Output File (*.chp)	368
Expression Analysis	368
Mutation and Polymorphism Analysis	369
Genotype Analysis	370
Hybridization Analysis	371

Printing a Report (*.rpt) 372

APPENDIX A *Microarray Suite Instrument Installation* **375**

Fluidics Station and Agilent Scanner Installation 375

Fluidics Station Installation (only) 381

Agilent Scanner Installation (only) 386

APPENDIX B *Microarray Suite File Types* **393**

Probe Information (Library) Files 393

Fluidics Protocol Files 394

Experiment Data Files 394

APPENDIX C *Statistical Expression Algorithm* **399**

Notation 399

Algorithm Output Definitions 400

Background Subtraction 401

Absolute Expression Analysis 402

 Detection 402

 Discrimination Value 402

 Making the Call 403

 Calculating the Signal 404

Comparison Expression Analysis 404

 Signal Log Ratio 407

References 407

APPENDIX D *Expression Analysis Settings* **411**

Viewing Expression Analysis Settings	411
Scaling Tab	412
User Defined Scaling	413
<i>All Probe Sets Scaling</i>	414
<i>Selected Probe Sets Scaling</i>	415
Normalization Tab	422
<i>User Defined Normalization</i>	422
<i>All Probe Sets Normalization</i>	423
<i>Selected Probe Sets Normalization</i>	424
Probe Mask Tab	431
<i>Creating a Probe Mask</i>	432
<i>Selecting a Probe Mask</i>	434
<i>Editing a Probe Mask</i>	436
Baseline Tab	439
Parameters Tab	441
Changing a Parameter Value	443

APPENDIX E *Expression Analysis Metrics* **447**

Statistical Expression Algorithm	447
Absolute Results	447
<i>Signal</i>	447
<i>Detection</i>	447
<i>Detection p-value</i>	447
<i>Stat Pairs</i>	447
<i>Stat Pairs Used</i>	447
Comparison Results	448
<i>Signal Log Ratio</i>	448
<i>Signal Log Ratio Low</i>	448

<i>Signal Log Ratio High</i>	448
<i>Change</i>	448
<i>Change p-value</i>	448
<i>Stat Common Pairs</i>	448
Empirical Expression Algorithm	449
Absolute Analysis Results	449
<i>Positive & Negative</i>	449
<i>Pairs</i>	449
<i>Pair Used</i>	450
<i>Pairs in Avg</i>	450
<i>Pos Fraction</i>	450
<i>Log Avg</i>	450
<i>Pos/Neg</i>	451
<i>Avg Diff</i>	451
<i>Absolute Call</i>	451
Comparison Analysis Results	451
<i>Increase</i>	451
<i>Decrease</i>	452
<i>Inc Ratio</i>	452
<i>Dec Ratio</i>	452
<i>Pos Change</i>	452
<i>Neg Change</i>	452
<i>Inc/Dec</i>	452
<i>DPos-DNeg Ratio</i>	453
<i>Log Avg Ratio Change</i>	453
<i>Difference Call</i>	453
<i>Avg Diff Change</i>	453
<i>B=A</i>	453
<i>Fold Change</i>	454
<i>Sort Score</i>	454

APPENDIX F <i>Cross Hybridization, Hybridization, & Spike Probe Masks</i>	457
<hr/>	
Viewing the Probe Mask Tab	458
Cross Hybridization Probe Mask	459
Creating a Cross Hybridization Probe Mask	459
Editing a Cross Hybridization Probe Mask	464
Combining Cross Hybridization Probe Masks	466
Hybridization Probe Mask	470
Creating a Hybridization Probe Mask	471
Editing a Hybridization Probe Mask	475
Combining Hybridization Probe Masks	478
Spike Probe Mask	481
Creating a Spike Probe Mask	482
Editing a Spike Probe Mask	485
Combining Spike Probe Masks	487
APPENDIX G <i>Mixture Detection Analysis Settings</i>	493
<hr/>	
Viewing & Editing Analysis Settings	493
Intensity Cutoff	495
Score Cutoff	495
Discrimination Quality Filter	496
Reference Cell File	498
Selecting a Reference Cell Intensity File	498
APPENDIX H <i>Genotyping Analysis Settings</i>	503
<hr/>	
Viewing & Editing Analysis Settings	503

Number of Standard Deviations Cutoff	505
S Cutoff	506
Minimum Miniblock Number	506

APPENDIX I *Genotype Analysis Viewer Report* **509**

Overview	509
HuSNP.tab File	510
Making a Customized HuSNP.tab Table	512

APPENDIX J *Toolbars* **517**

Main Toolbar	517
Expression Analysis Window (EAW) Toolbar	519
Sequence Analysis Window (SAW) Toolbar	521
Nucleotide Analysis Window (NAW) Toolbar	522
GT Viewer Toolbar	523
Hybridization Analysis Window (HAW) Toolbar	524
Batch Analysis Toolbar	525
Report Toolbar	526
Publish Toolbar	527

APPENDIX K *Working With Windowpanes & Columns* **531**

Resizing Windowpanes	531
Resizing or Hiding Columns in the EAW or NAW	532

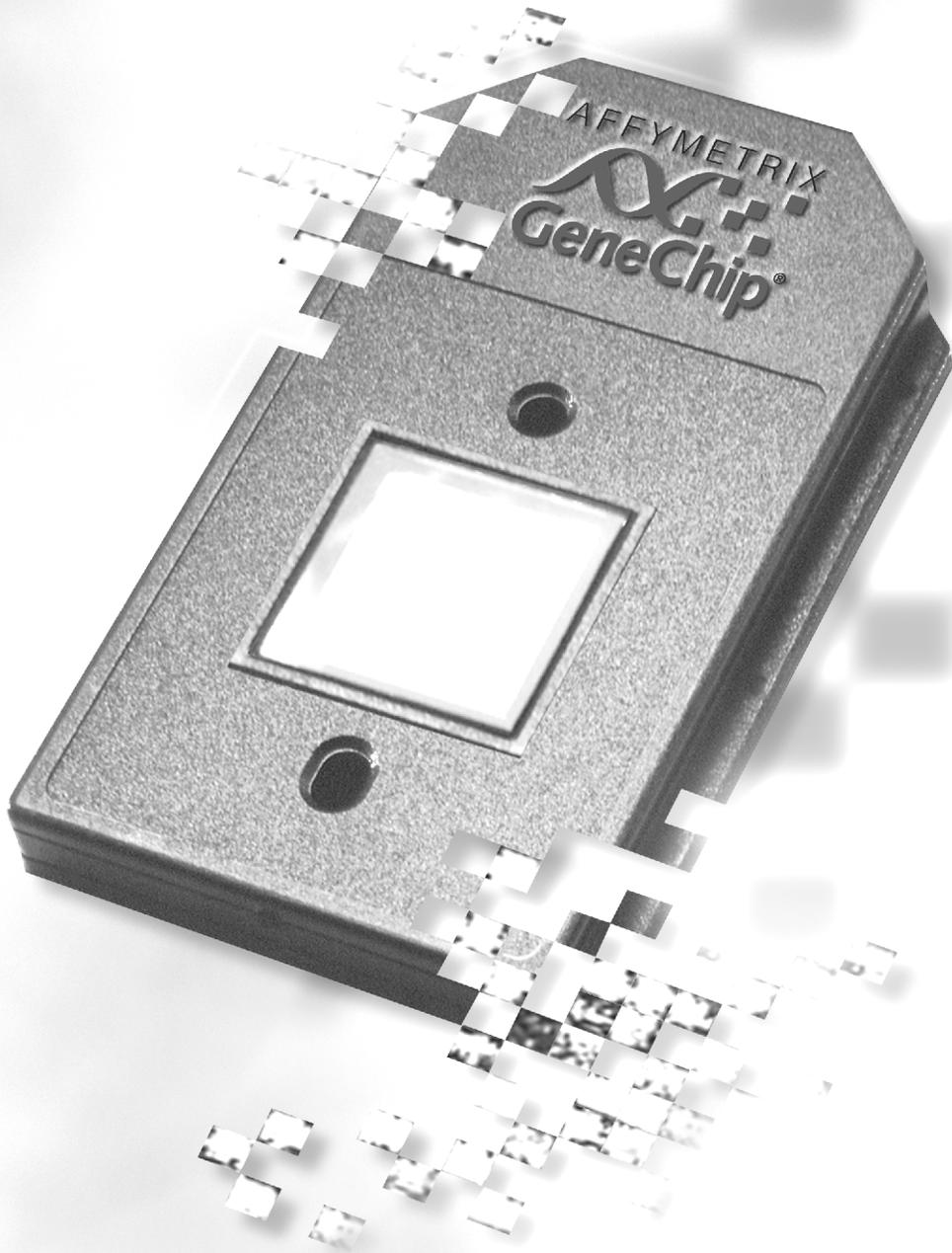
APPENDIX L *Base Codes and Amino Acid Abbreviations* **535**

IUPAC Base Codes 535

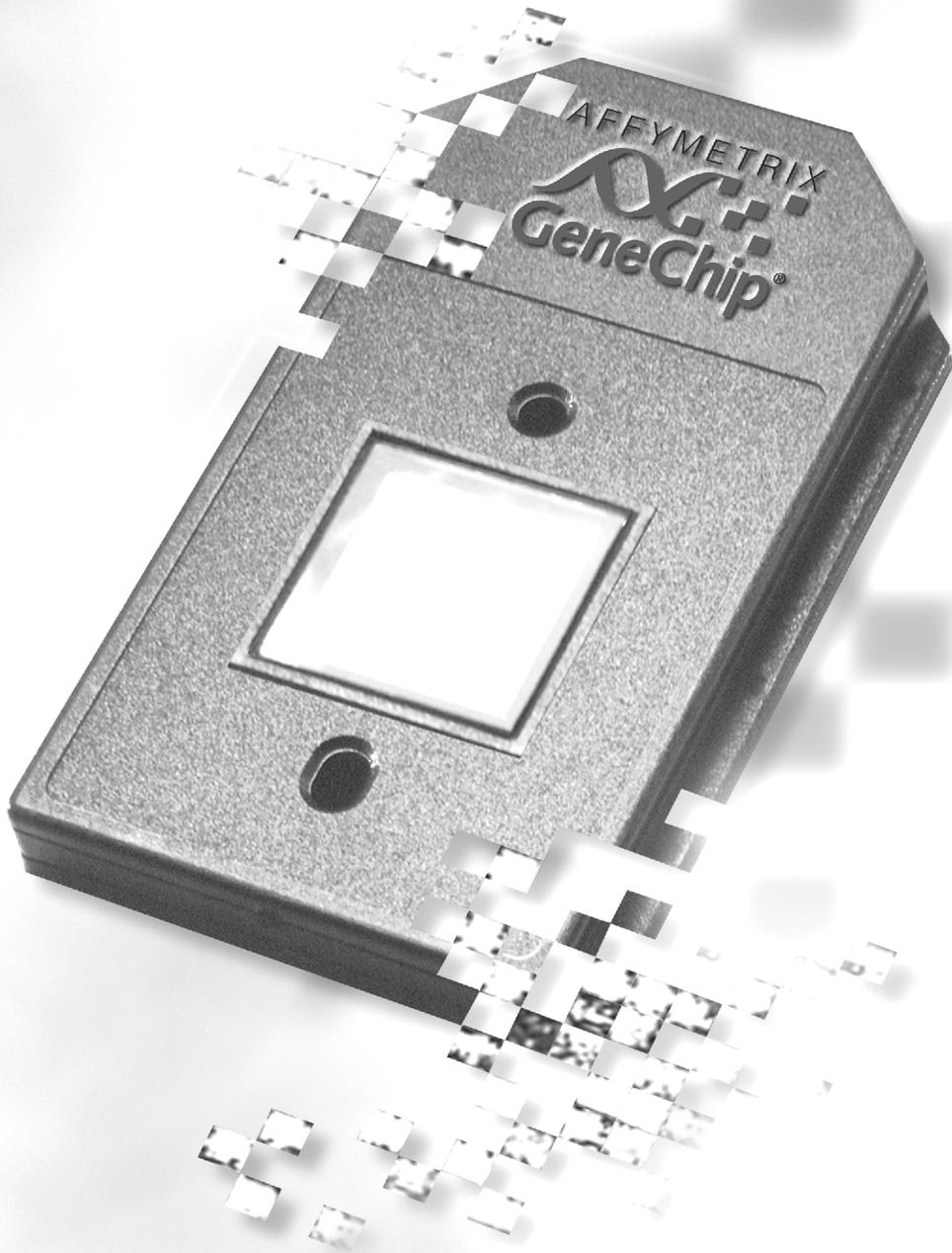
Amino Acid Abbreviations 536

APPENDIX M *Hot Key Descriptions* **539**

INDEX **543**



Chapter 1



Welcome to the Affymetrix® Microarray Suite User's Guide. Microarray Suite software is part of the Affymetrix® GeneChip® instrument and probe array platform. This manual explains how to use Microarray Suite to:

- control the GeneChip® Fluidics Station 400 that hybridizes, washes, and stains the GeneChip® probe and tag arrays
- control the Agilent GeneArray® Scanner that scans hybridized GeneChip probe or tag arrays
- analyze experimental data for GeneChip probe array assays
- generate reports that summarize intensity data, analysis output, and algorithm settings
- publish data (LIMS mode only) to a publish database that can be queried using the Affymetrix® Data Mining Tool

What's New In Microarray Suite 5.0

Statistical Expression Analysis Algorithm

Microarray Suite includes a new expression algorithm that generates new expression results (see [Table 1.1](#)).

Table 1.1
Statistical expression results

Statistical Expression Algorithm Metrics	
Absolute Analysis	Comparison Analysis
Detection call (present, absent, marginal, no call)	Change call (increase, decrease, no change, marginal increase, marginal decrease, no call)
Detection p-value	Change detection p-value
Signal	Signal log ratio
Stat Pairs	Signal log ratio low
Stat pairs used	Signal log ratio high
	Stat Common Pairs

In an absolute analysis, the detection call answers the question: Is the transcript of a particular probe set reliably detected by the probe array? The algorithm provides call information without reference to numerical values so you can easily filter and interpret results. The signal represents the amount of transcript in solution.

In a comparison analysis, the change call answers the question: Does the expression level of a transcript in an experiment change with respect to a baseline? Signal log ratio values provide accurate and sensitive measures of changes in expression level because they eliminate differences in probe hybridization efficiency (a source of noise).

The detection p-value and change p-value indicate the significance of the result. The Statistical algorithm has user-modifiable parameters that enable you to set the sensitivity and specificity of the analysis.

Genotyping (GT) Analysis Viewer

The GT Viewer provides the relative allele signal (RAS) for each marker used in the Affymetrix® HuSNP™ Mapping assay. RAS is a quantitative representation of the presence of the two possible alleles, A and B.

The GT Viewer algorithm detects variations between the genotype calls of a control and experimental sample. When the GT Viewer analyzes a control and an experimental sample, the algorithm computes the difference in RAS (Delta RAS) between the two samples. Delta RAS indicates the shift in signal for each SNP between a control and an experimental sample. The GT Viewer report displays Delta RAS for each SNP marker represented on the Affymetrix® HuSNP™ probe array.

The GT Viewer report also includes the number of heterozygote standard deviations that the experiment RAS value is from the control RAS value, assuming a normal distribution of RAS values. (See Appendix I for more information.)



Note

The GT Viewer is available when Microarray Suite is run in disk mode only.

Export or Import Batch Files

A batch analysis (specifies a group of *.cel files for analysis with unattended operation) can be exported to a tab-delimited text file (*.txt). The batch file provides a record of the batch analysis and may be imported for a subsequent batch analysis. A batch file (tab-delimited *.txt) may also be defined in Microsoft® Excel and imported into Microarray Suite.

If a batch analysis is interrupted, the recovery feature resumes the process for the remaining files.

Publish Database Security

Publishing (copying experimental data to a publish database) offers security that prevents unauthorized database access. Each publish database requires a login password before publishing can proceed.



Note

Publishing is available when Microarray Suite is run in LIMS mode only.

Conventions Used in This Guide

This manual provides a detailed outline for all tasks associated with Affymetrix® Microarray Suite software. Various conventions are used throughout the manual to help illustrate the procedures described. Explanations of these conventions are provided below.

Steps

Instructions for procedures are written in a step format. Immediately following the step number is the action to be performed. On the line below the step there may be the following symbol: ⇒. This symbol defines the system response or consequence that happens as a result of user action; what you see and what has happened.

Following the response additional information pertaining to the step may be found and is presented in paragraph format. For example:

9. Click **Yes** to continue.

⇒ The Delete task proceeds.

In the lower right pane the status is displayed.

To view more information pertaining to the delete task, right-click **Delete** and select **View Task Log** from the shortcut menu.

Font Styles

Bold fonts indicate names of commands, buttons, options or titles within a dialog box. When asked to enter specific information, such input appears in italics within the procedure being outlined.

For example:

1. Click the **Find** toolbar button  or select **Edit** → **Find** from the menu bar.
⇒ The Find dialog box appears.
2. Enter *AFFX-BioB-5_at* in the **Find what** box, then click **Find Next** to view the first search result.
3. Continue to click **Find Next** to view each successive search result.

Screen Captures

The steps outlining procedures are frequently supplemented with screen captures to further illustrate the instructions given.

✓ Note

The screen captures depicted in this manual may not exactly match the windows displayed on your screen.

Additional Comments

Throughout the manual, text and procedures are occasionally accompanied by special notes. These additional comments and their meanings are described below.

💡 TIP

Information presented in Tips provide helpful advice or shortcuts for completing a task.

✓ Note

The Note format presents important information pertaining to the text or procedure being outlined.

! CAUTION

Caution notes advise you that the consequence(s) of an action may be irreversible and/or result in lost data.

▲ WARNING

Warnings alert you to situations where physical harm to person or damage to hardware is possible.

Online Documentation

The CD with Microarray Suite includes an electronic version of this user's guide. The online documentation is in Adobe Acrobat® format (a *.pdf file) and is readable with the Adobe Acrobat® Reader software, available at no charge from Adobe at <http://www.adobe.com>.

The electronic user's guide is printable, searchable, and fully indexed. You can have it open and minimized on screen while using the Affymetrix® GeneChip® instruments and Microarray Suite software.

Technical Support

Affymetrix provides technical support to all licensed users via phone or E-mail. To contact Affymetrix Technical Support:

AFFYMETRIX, INC.

3380 Central Expressway
Santa Clara, CA 95051 USA
Tel: 1-888-362-2447 (1-888-DNA-CHIP)
Fax: 1-408-731-5441

sales@affymetrix.com
support@affymetrix.com

AFFYMETRIX UK Ltd.,

Voyager, Mercury Park,
Wycombe Lane, Wooburn Green,
High Wycombe HP10 0HH
United Kingdom
Tel: +44 (0) 1628 552550
Fax: +44 (0) 1628 552585

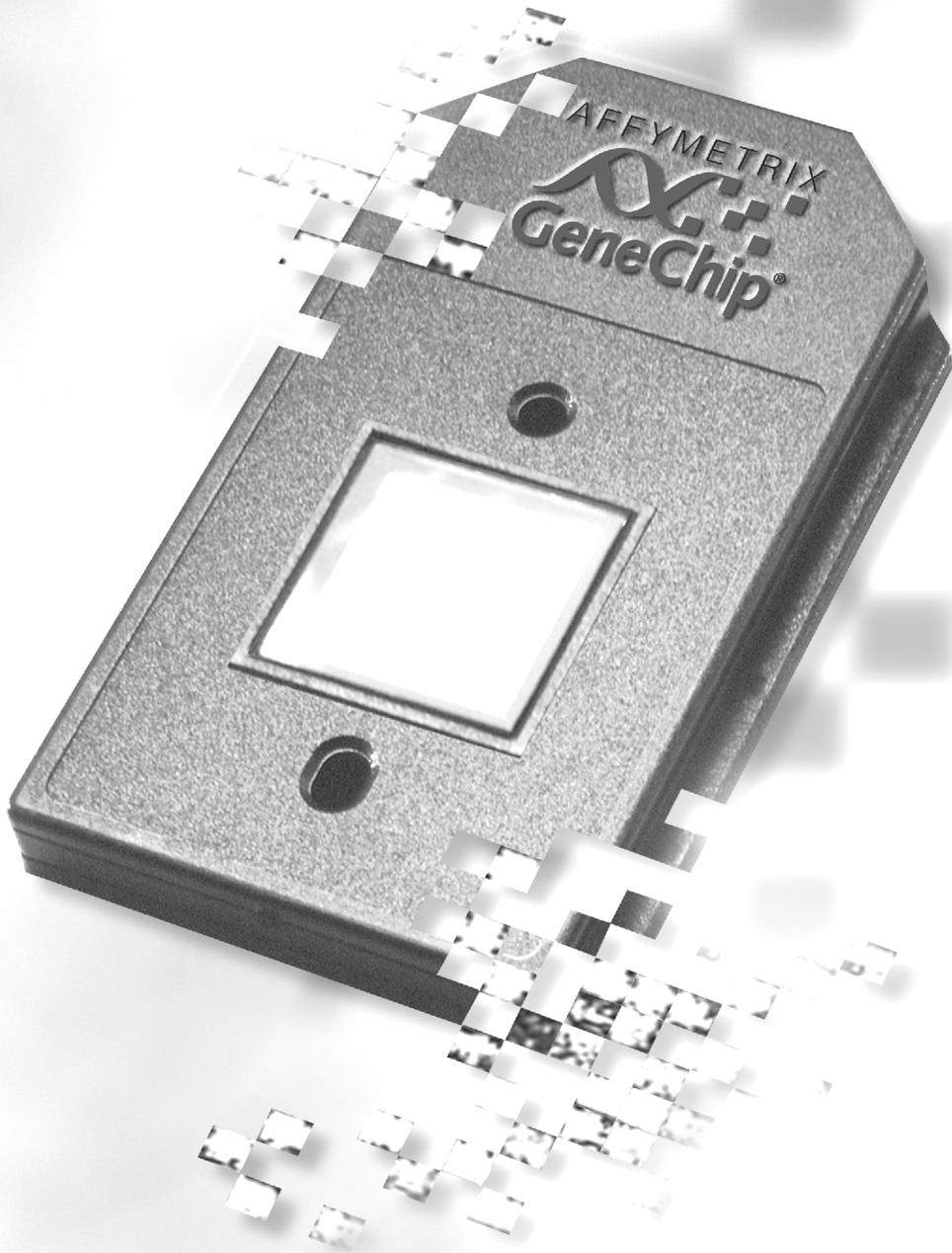
saleseurope@affymetrix.com
supporteurope@affymetrix.com

www.affymetrix.com

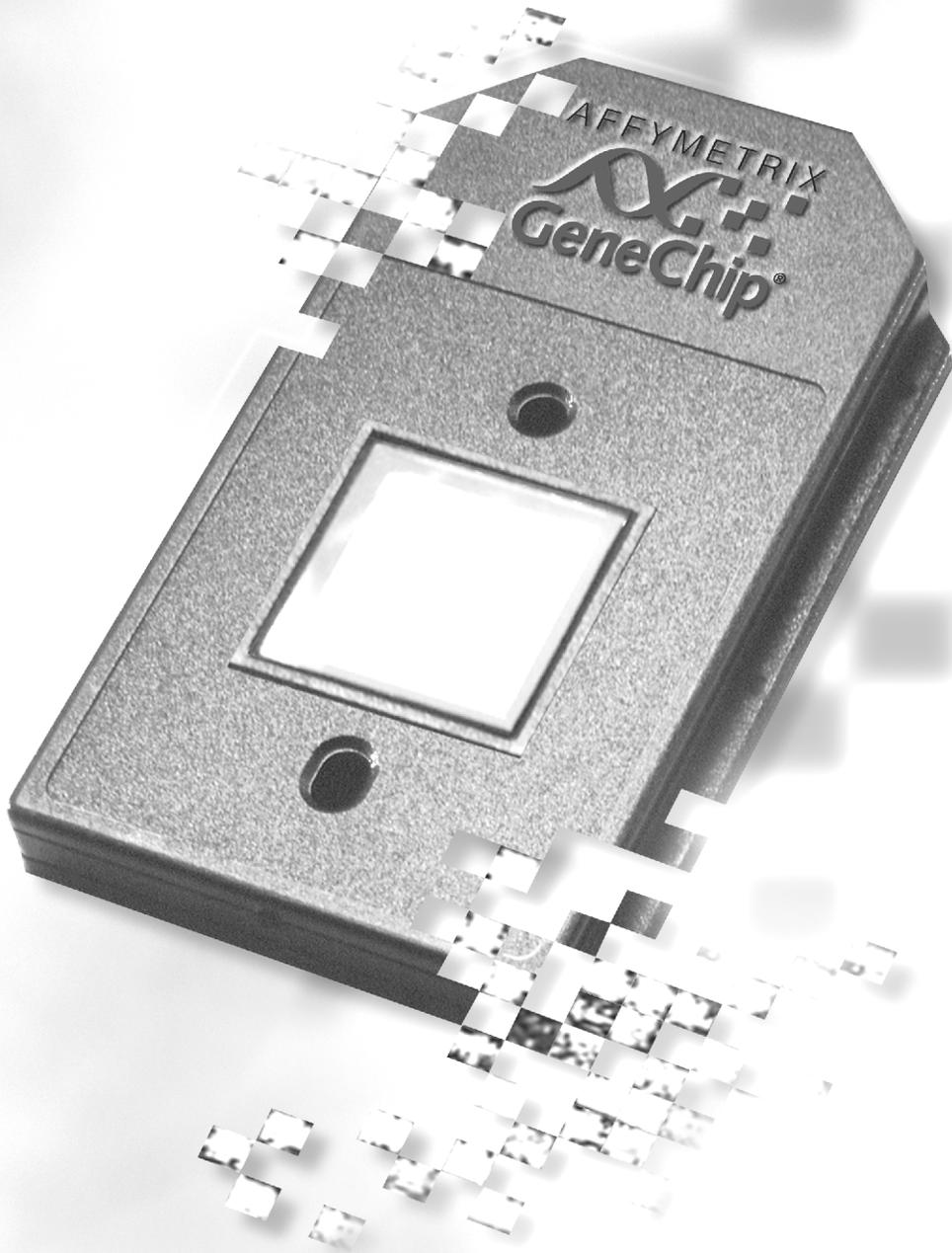
Your Feedback is Welcome

Affymetrix Technical Publications is dedicated to continually improving the quality of our documentation, helping you get the information that you need. We welcome any comments or suggestions you may have regarding this manual. Please contact us at:

Technical_Publications@Affymetrix.com



Chapter 2





Installing Microarray Suite

This chapter provides detailed instructions for installing Microarray Suite 5.0 or upgrading from a previous version of the software. See [Appendix A](#) for information on installing the instruments.

Microarray Suite 5.0

During the installation of Affymetrix® Microarray Suite 5.0, the installation program checks for the presence of Microsoft® Data Engine (MSDE). If not present, MSDE is installed prior to the installation of Microarray Suite 5.0 application. MSDE is used to store the gene descriptions and target sequences that are installed during the library files installation.

Following the installation of MSDE you are prompted for the location of GeneInfo. After GeneInfo is installed, the Microarray Suite installation begins.

Note

You must be logged in as an administrator to install Microarray Suite 5.0 software.

The screen captures depicted in this manual may not exactly match the windows displayed on your screen.

It is recommended that there is at least 500 MB of available disk space for the installation.

- 1.** Insert the Microarray Suite 5.0 CD-ROM.
- 2.** If the autorun feature does not start the program:
 - a.** Click **Start** → **Run**.
 - b.** Type `<cd drive letter>:\setup.exe`.
 - c.** Click **OK**.⇒ The Affymetrix Software Setup window appears ([Figure 2.1](#)).

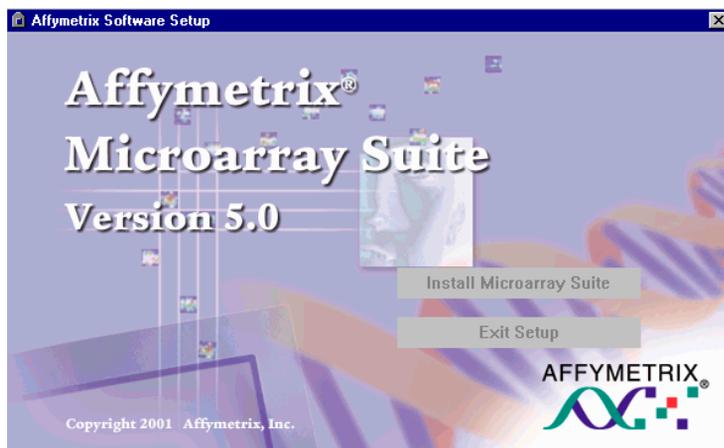


Figure 2.1
Install window

3. Click **Install Microarray Suite**.

⇒ If system files must be copied, the system prompts for a reboot before the install program continues (**Figure 2.2**).

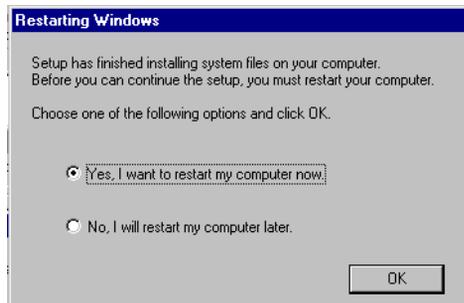


Figure 2.2
Restart Windows

4. Click **Yes I want to restart my computer now**.

If the “No” option is selected, the installation program will exit and Microarray Suite will not be installed.

5. Upon rebooting the computer, log in as the same installer with administrative privileges.
6. If you are upgrading from a previous version of Microarray Suite, [go to step 10](#). If this is a new installation, proceed from here.
⇒ The MSDE Welcome window appears ([Figure 2.3](#)).

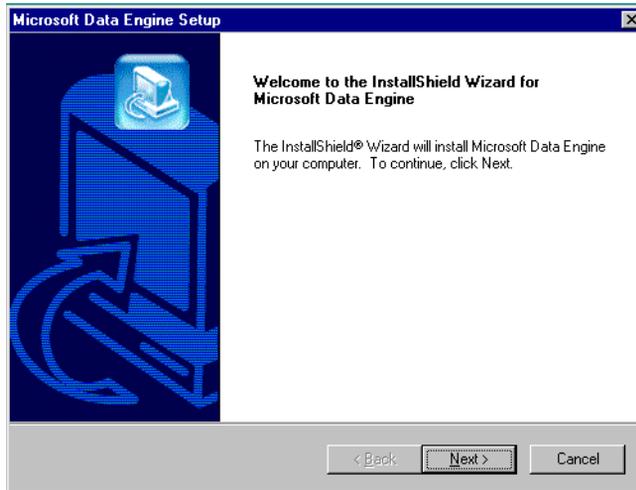


Figure 2.3
MSDE Welcome window

7. Click **Next**.
⇒ The MSDE Installation window appears ([Figure 2.4](#)).

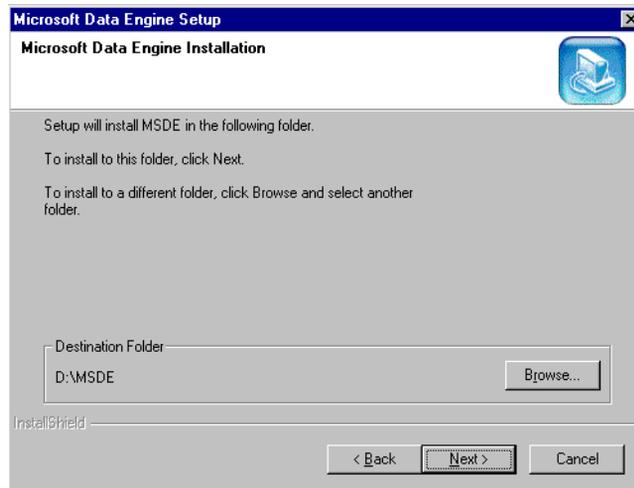


Figure 2.4
Select the location to install MSDE

8. Select the destination folder to install the MSDE application. The MSDE install will default to the drive with the most available free disk space. Click **Next**.
 - ⇒ The unpacking of MSDE begins (**Figure 2.5**).

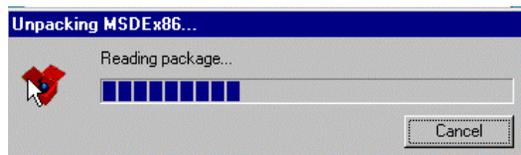


Figure 2.5
Unpacking MSDE Window

- ⇒ Installing MSDE window appears (**Figure 2.6**).
- ⇒ This takes a few minutes. Do not stop the installation. Be patient and wait for the installation of MSDE to complete.

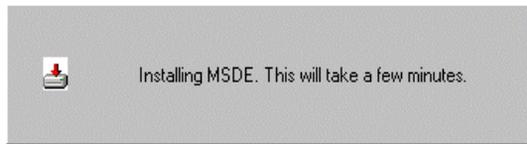


Figure 2.6
Installing MSDE Window

9. The Welcome window appears (Figure 2.7).

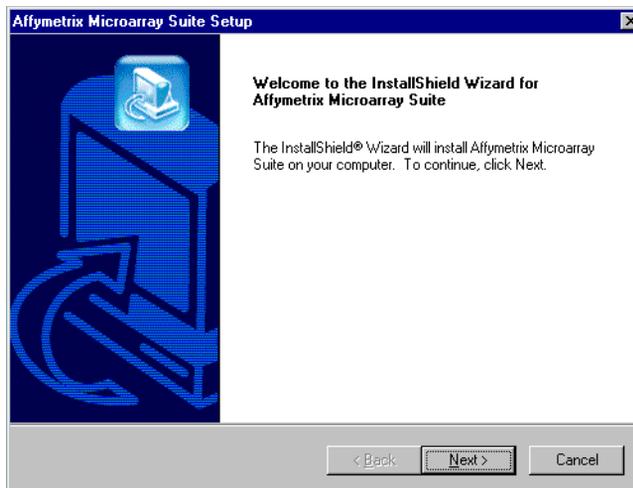


Figure 2.7
Welcome window

10. Click **Next**.
⇒ Several consecutive License Agreement windows appear (Figure 2.8).

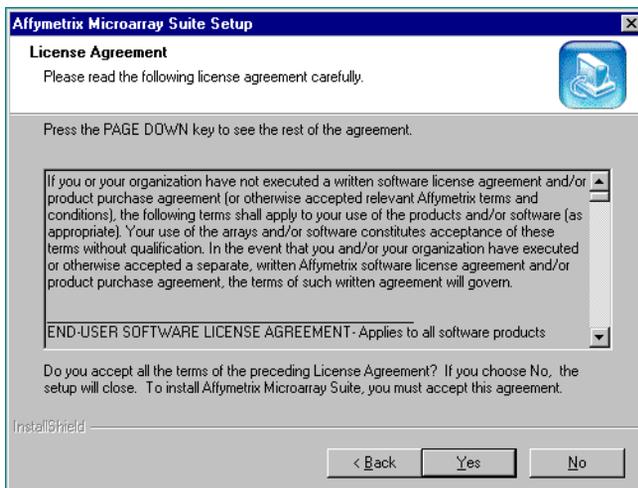


Figure 2.8
License Agreement Windows

11. Review the contents and click **Yes** in each window to accept the terms of the licensing agreement.

⇒ The Customer Information window appears (Figure 2.9).

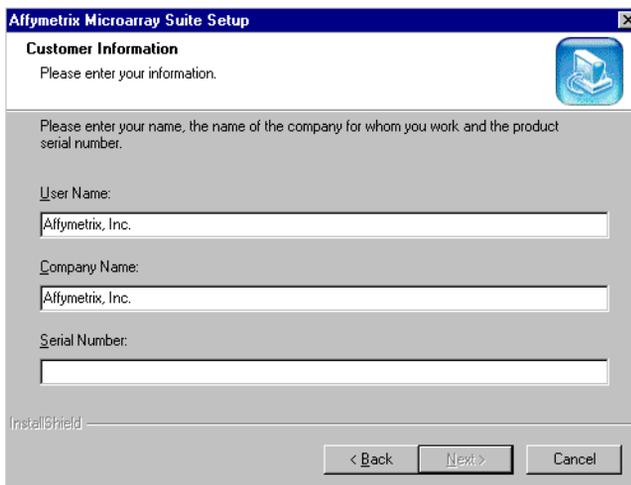


Figure 2.9
Customer Information window

12. Enter your Name, Company, and Serial Number.

The serial number is located on the *Affymetrix® Software Product Registration* card.

✓ Note

*If you do not have a serial number, contact Affymetrix Technical Support. If you are upgrading from a previous version, the **Serial** number field populates automatically.*

13. Click Next.

⇒ The **Choose Destination Location** window appears (Figure 2.10).

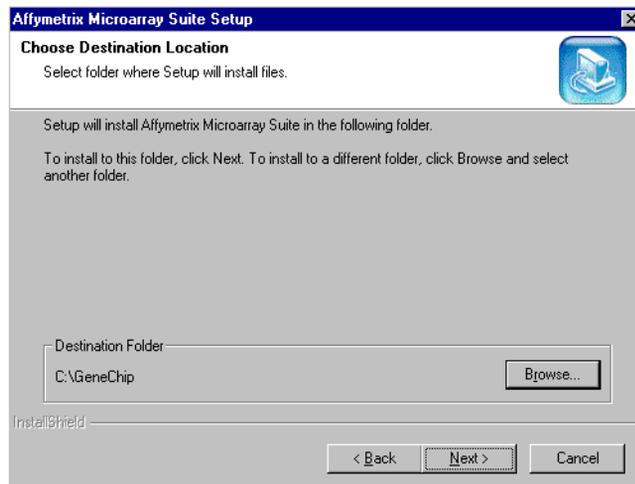


Figure 2.10
Choose Destination Location window

14. Select the destination where Microarray Suite 5.0 will be installed.

C:\GeneChip is the default location.

If you are upgrading from a previous version, we recommend installing in the same directory.

15. Click Next.

⇒ The **Select Components** window appears (Figure 2.11).

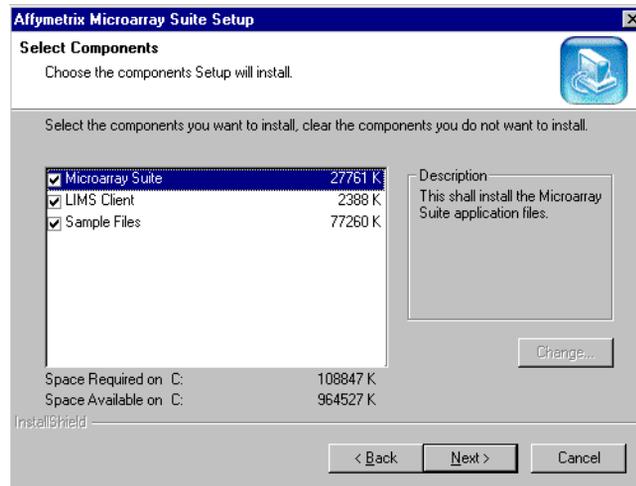


Figure 2.11
Select Components Window

16. Select the components you want to install. For documentation purposes, all components are selected.
 - **Microarray Suite** - This component contains all files required for the Microarray Suite and are required for any Microarray-related installation.
 - **LIMS Client** - This component includes files to connect to LIMS and are required if the workstation will be used in conjunction with LIMS. Select LIMS Client only if there is a LIMS Server; otherwise clear this option.

✓ **Note**

The **LIMS Client** option should only be selected if connecting to a LIMS server.

- **Sample Files** - This component includes example experiment files.
17. Click **Next**.
 - ⇒ The **Start Copying Files** window appears ([Figure 2.12](#)).
- This is a summary of the information that was selected by the installer.

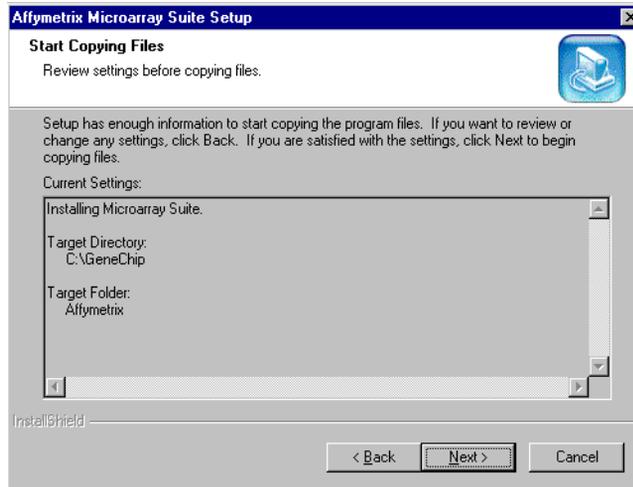


Figure 2.12
Setup Summary Information Window

18. Click Next.

⇒ The Installation Program Files window appears and installation begins (**Figure 2.13**).

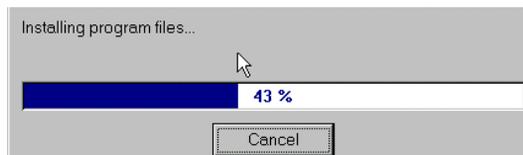


Figure 2.13
Installation of Program Files

19. If the LIMS client was selected, the LIMS configuration window appears (Figure 2.14**).**

- If you have the disk, click **Yes** and go to step a.
- If you do not have the disk, click **No** and go to step b.

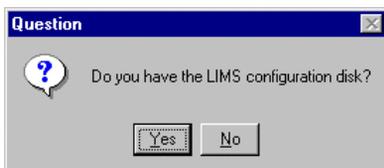


Figure 2.14
LIMS Configuration Disk window

- a. Browse to the location of LIMS configuration disk. Click **OK**. Go to step 20.

✓ Note

The LIMS configuration disk is a file named GeneChipDB.ini. It is located on the LIMS Server: <server>\gclims\library, where <server> is the name of the LIMS server.

- b. In the window that appears (**Figure 2.15**), enter the name of LIMS server. Click **Next**.

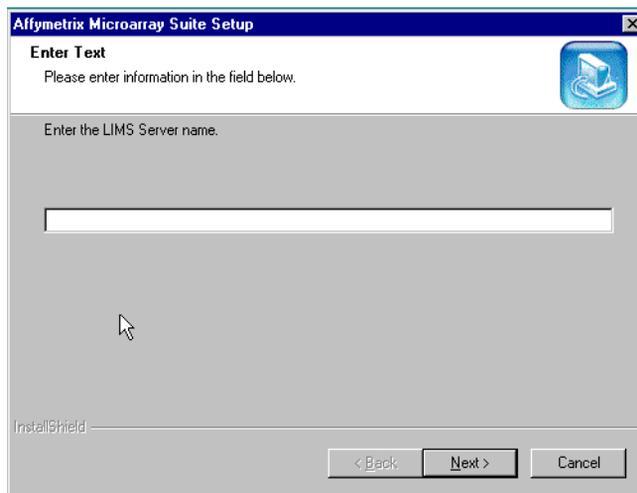


Figure 2.15
Enter Name of LIMS Server

⇒ The Global Group Name window appears (**Figure 2.16**).

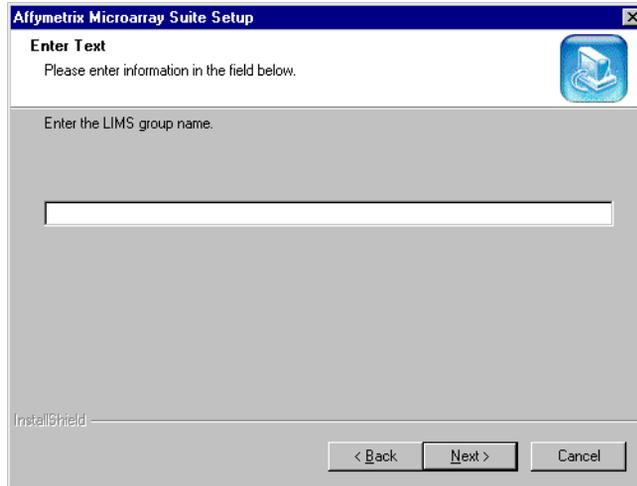


Figure 2.16
Enter Global Group Name window

- c. Enter the name of the Domain the server resides on with the Global Group name which contains the LIMS Users. (Domain Name\Global Group).
- d. Click **Next**.
 - ⇒ The install configures the system registry.
 - ⇒ The Install Shield Wizard Complete window appears ([Figure 2.17](#)).

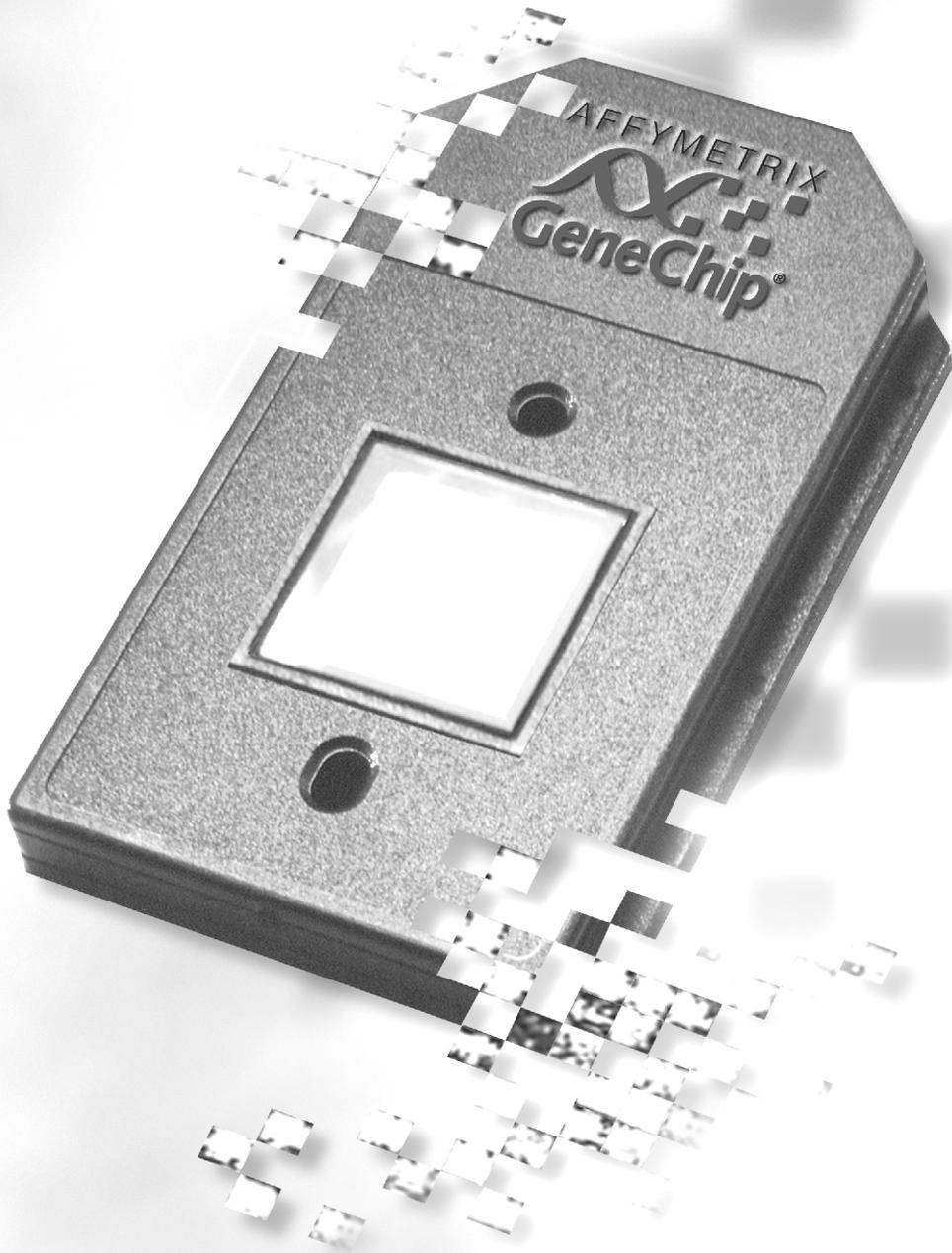


Figure 2.17
Restart Computer Window

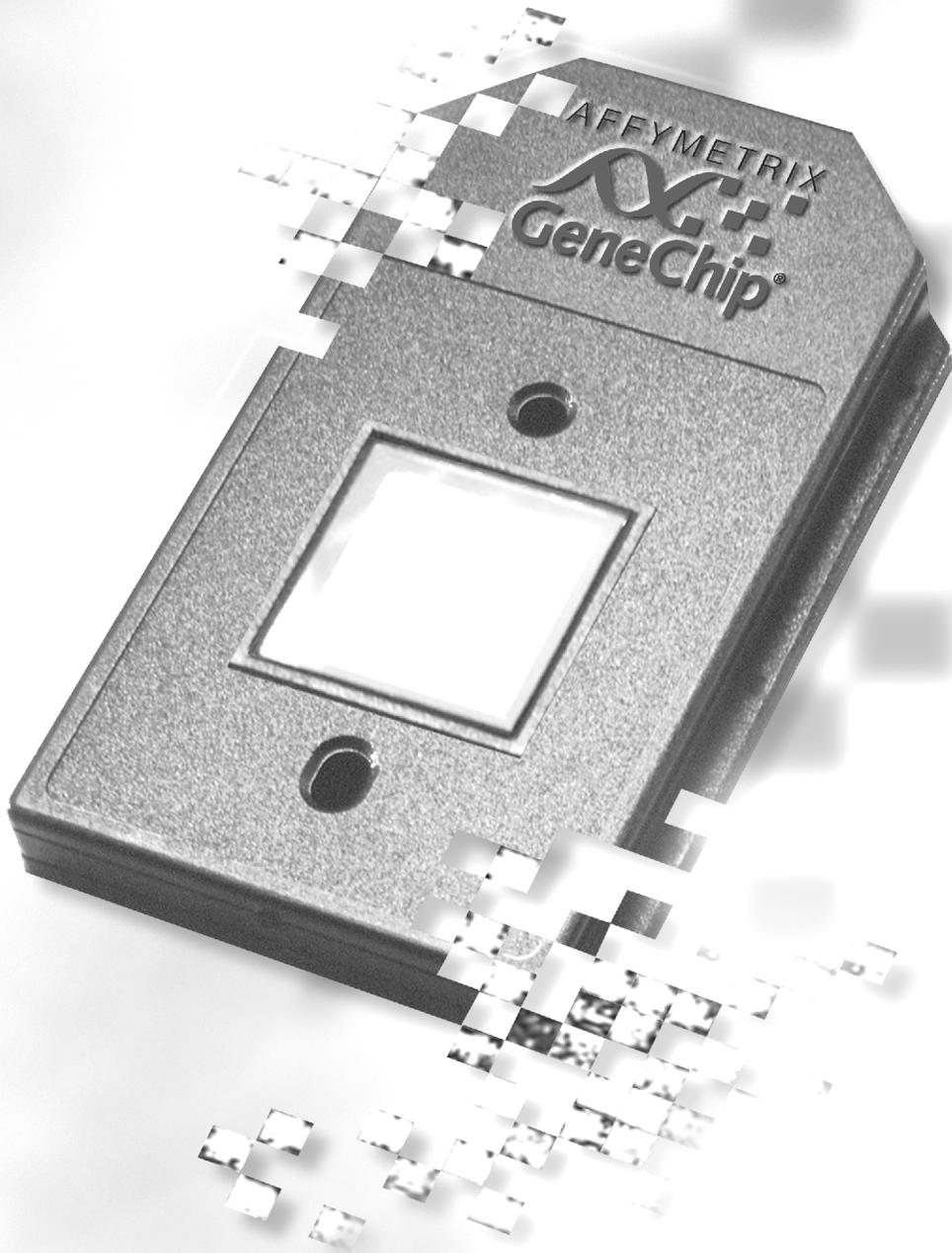
20. If the install prompts for a reboot, restart the computer: Select **Yes, I want to restart my computer now**, and click **Finish**.

⇒ The system reboots.

If you are not prompted to reboot your computer, the installation is complete.



Chapter 3





Microarray Suite Overview

This chapter provides background information on the Affymetrix® GeneChip® instrument and probe array platform including a walkthrough of Microarray Suite and an overview of data analysis.

Microarray Suite Walkthrough

The Affymetrix® GeneChip® probe array platform uses application-specific high-density GeneChip oligonucleotide probe arrays to efficiently acquire and analyze genetic information for gene expression, genotyping, and nucleotide variant detection. This integrated and complete system includes the:

- GeneChip® Hybridization Oven 640
- GeneChip® Fluidics Station 400
- Agilent GeneArray® Scanner
- Affymetrix® computer workstation
- Affymetrix® Microarray Suite software
- Affymetrix® MicroDB™ software
- Affymetrix® Data Mining Tool (DMT) software
- Affymetrix® Microarray Laboratory Information Management System (LIMS)

Microarray Suite automates control of the GeneChip® Fluidics Station 400 and Agilent GeneArray® Scanner, captures GeneChip probe and tag array data, and provides powerful algorithms that analyze hybridization intensity data.

The Microarray Suite algorithms determine absolute or comparison gene expression results, perform nucleotide variant detection, or identify genotype, depending on the type of probe array assay.

A GeneChip® probe array assay has five major steps (**Figure 3.1**):

1. Prepare the target (refer to the appropriate GeneChip probe array package insert, the *Expression Analysis Technical Manual*, or the *HuSNP™ Mapping Assay User Manual*).
2. Set up an experiment in Microarray Suite.
3. Hybridize, wash, and stain (if necessary) the probe array.
4. Scan the probe array.
5. Analyze the hybridization data.

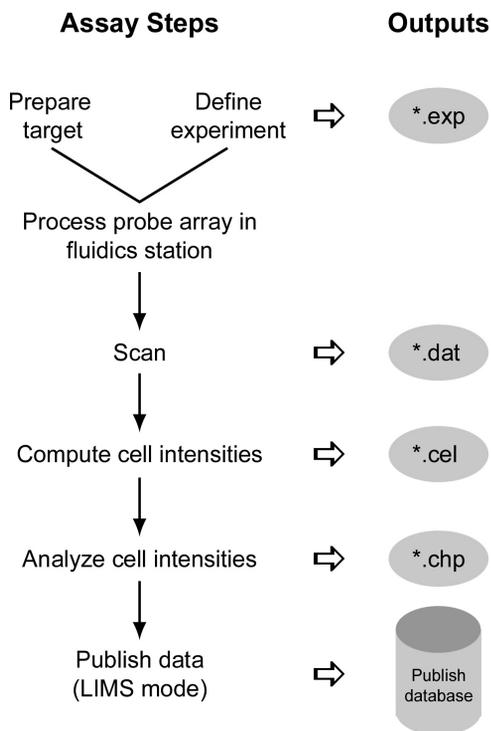


Figure 3.1
GeneChip® probe array assay and outputs

Setting Up an Experiment

Microarray Suite enables you to define an experiment, select the associated probe array type, and enter other relevant information. This process creates an experiment information file (*.exp).

Microarray Suite relies on the experiment information file to select the scanner specifications for a probe array and the appropriate algorithm for data analysis.

Hybridizing, Washing, and Staining a Probe Array

During hybridization, the probe array is incubated with a hybridization cocktail containing the labeled target and control oligonucleotides in a buffer optimized for each type of probe array (refer to the appropriate GeneChip® probe array package insert, *Expression Analysis Technical Manual*, or the *HuSNP™ Mapping Assay User Manual*).

The hybridization reaction occurs in the GeneChip Fluidics Station 400, except for assays requiring longer hybridization periods in which case the hybridization reaction occurs in the GeneChip Hybridization Oven 640.

During hybridization in the fluidics station, the probe array is repeatedly filled and drained with the hybridization cocktail. The strongest hybridization occurs between the array probes and sequences in the target that are most nearly complementary.

After hybridization, the probe array undergoes a series of stringent washes (specifically optimized for each type of probe array) in the fluidics station. If the target was labeled with a fluorescent tag, the array may be scanned at this point. In applications that use a biotin-labeled target, the probe array must be stained with a streptavidin-conjugated fluorescent stain and may require antibody amplification (depending on probe cell size) prior to scanning.

Microarray Suite controls the fluidics station using preprogrammed fluidics protocols for hybridizing, washing, and staining the probe arrays. The computer workstation running Microarray Suite can simultaneously control up to eight fluidics stations. A fluidics station contains four modules and each module can independently process a probe array using different fluidics protocols. Some of the fluidics protocols may be customized. Utilities for station maintenance are also included.

Scanning a Probe Array

The probe array is ready to be scanned after it has been hybridized, washed, and stained (if necessary). Microarray Suite controls the Agilent GeneArray® Scanner and enables you to start a scan and view the intensity data as they are collected during scanning.

The scanner is based on epifluorescence confocal microscopy and uses an argon-ion laser to excite the fluorophores at 488 nm. As the surface of the probe array is scanned, a photomultiplier tube collects and converts the fluorescence emissions into an electrical current. The scanner converts this current into a numeric value through an analog to digital converter. These digital intensity values are collected from discrete areas (pixels) on the array surface and are stored on the computer workstation as an image data file (*.dat) (see [Chapter 8](#)).

Microarray Suite manages the resolution settings for each type of probe array (resolution and other aspects of scanning are user-modifiable). A scan requires 5 to 25 minutes depending on the size of the scan region, the number of scans, and other parameters.

After the scan is completed, Microarray Suite displays a picture of the image data (*.dat) in the image window. The software represents the fluorescence intensity values from each pixel on the array in a grayscale or pseudocolor mode and superimposes a grid on the image to delineate the probe cells.

Microarray Suite analyzes the image (*.dat) and derives a single intensity value for each probe cell on an array. These data are automatically generated and saved to the cell intensity file (*.cel) (see [Chapter 8](#)).

Data Analysis

Microarray Suite refers to the experiment information file (*.exp) to select the appropriate analysis algorithm for a cell intensity file (*.cel) that generates the analysis output or *chip* file (*.chp) ([Figure 3.1](#)). The results contained in the analysis chip file are automatically displayed in a specific analysis window ([Table 3.1](#)).

Table 3.1

Microarray Suite algorithms analyze probe or tag array data to generate analysis output files (*.chp) that are viewed in specific analysis windows

Analysis	GeneChip® Probe or Tag Array	Microarray Suite Algorithm	View the analysis output (*.chp) in the...
Gene expression	GeneChip expression probe array	Expression	Expression Analysis window (EAW)
Mutation & polymorphism	GeneChip p53 probe array	Mixture Detection	Sequence Analysis window (SAW)
	GeneChip HIV PRT Plus probe array	Consensus Rules	
Genotype	GeneChip HuSNP™ probe array	Genotyping	Nucleotide Analysis window (NAW)
		Block	
Hybridization	GenFlex™ tag array	Hybridization	Hybridization Analysis window (HAW)

Microarray Suite generates reports (*.rpt) that summarize probe array and analysis output (*.chp) information:

- Expression Report
- Mapping Report
- p53 report
- CYP450 report
- HIV PRT *Plus* report.
- Genotyping Analysis (GT) Viewer Report

LIMS

The Affymetrix® LIMS (Laboratory Information Management System) is a separate database application that provides a central infrastructure to manage and track gene expression data generated by Microarray Suite.

LIMS moves the experimental data generated by Microarray Suite to a server database. LIMS manages several databases (see [Table 3.2](#)).

Table 3.2
LIMS databases

Database Name	Contents
Process	Data tracking information about samples and experiments, and information that relates the *.dat, *.cel, and *.chp files for a particular sample and experiment.
Publish	Expression and mapping data in the Affymetrix® Analysis Data Model database schema that can be accessed by the Affymetrix Data Mining Tool (DMT) and other software.
Gene Information	Accession numbers, probe set descriptions, target sequence, and user annotations accessible by DMT.
Data Mining Info	Queries, user settings, and probe lists for DMT.

Together Microarray Suite and LIMS enable users to register samples and set up experiments, and access and analyze data from any workstation on the network.

Workflow Based Tracking

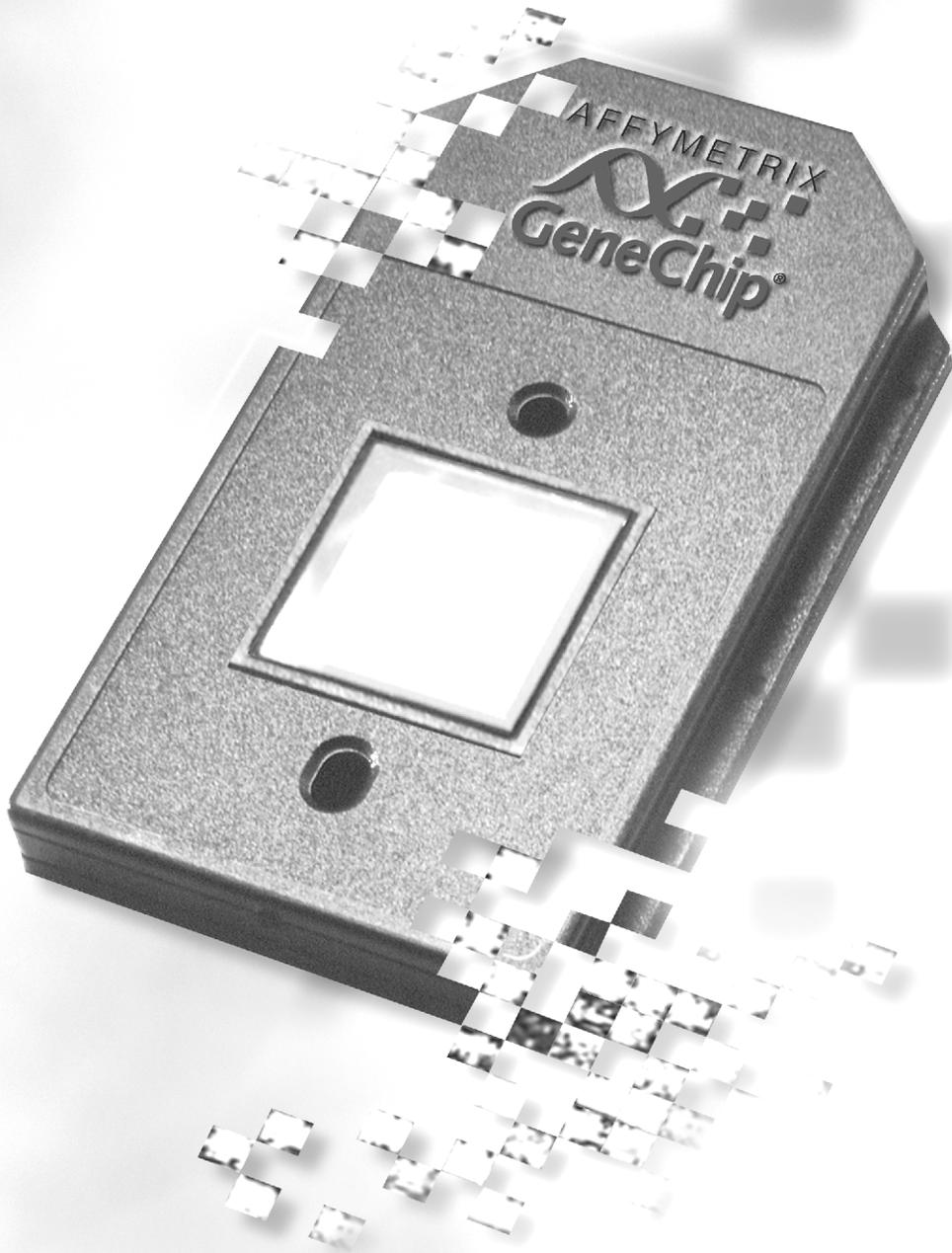
LIMS uses a workflow-based tracking system to process data from workstations on the network. The workflow monitor enables you to track sample or experiment status through a series of steps or *queue* that includes:

- hybridization
- scan
- grid alignment (if required)
- cell intensity analysis (for image data files requiring manual grid alignment)
- probe array analysis
- publish

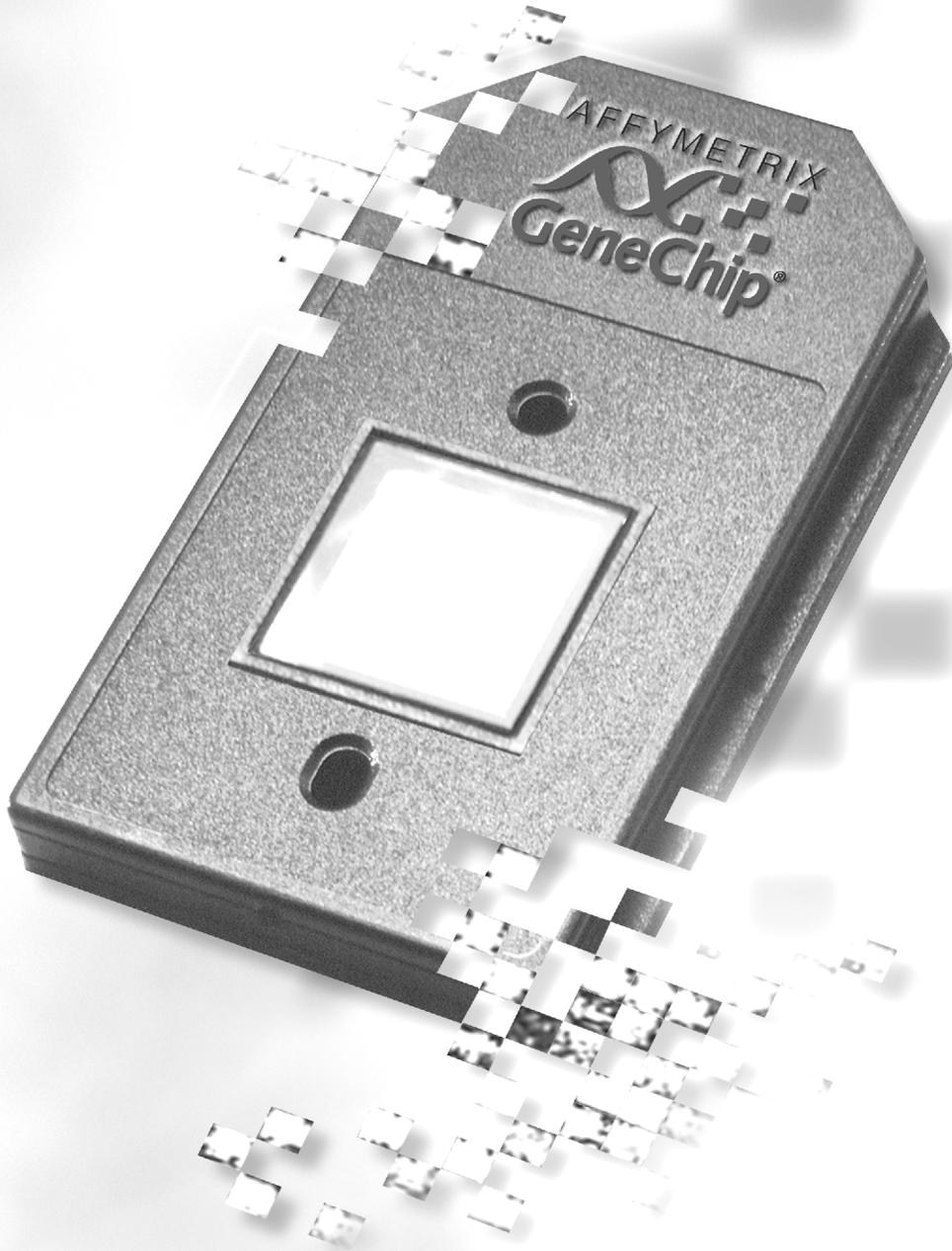
Data Mining

You can create publish databases populated with GeneChip® expression data. The publish database is accessible to Affymetrix Data Mining Tool (DMT) or several other third party analysis tools.

In LIMS mode, Microarray Suite publishes the data entered during sample registration and experiment setup as well as all analysis output (*.chp) to the database. In disk files mode, the separate application Affymetrix® MicroDB publishes the experiment and analysis output.



Chapter 4



This chapter explains how to:

- start the Microarray Suite software
- configure the software for use as a stand alone application (disk files mode) or with Affymetrix® LIMS (LIMS mode)
- specify file locations
- specify LIMS filters that determine the data displayed (if running Microarray Suite in LIMS mode)

Starting Microarray Suite

1. Turn on the power for the computer workstation. If the workstation is connected to the instruments, also turn on the power for the GeneChip® Fluidics Station 400 and the Agilent GeneArray® Scanner.

Note

The scanner laser should be turned on and warmed up at least 15 minutes before use. Refer to the Agilent GeneArray Scanner User's Guide for a description of the instrument, its components, and set up.

2. After the computer startup is complete, press **Ctrl+Alt+Delete** to open the logon dialog box. Enter your user name, password, and domain.
3. Click the Windows **Start** menu button  **Start**, then select **Programs** → **Affymetrix** → **Microarray Suite**.

When Microarray Suite starts, the program displays the shortcut bar, data file tree, and status log (Figure 4.1).

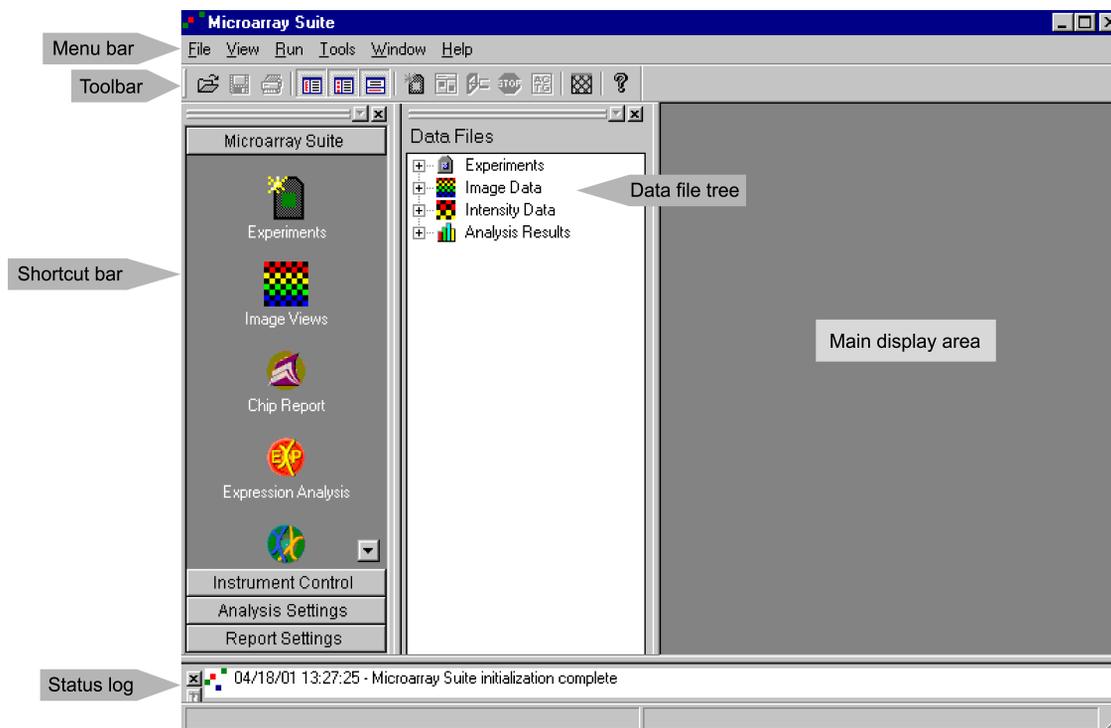


Figure 4.1
Microarray Suite user interface

Microarray Suite User Interface

Shortcut Bar

The shortcut bar provides quick alternatives to menu bar commands. The shortcut bar has four sections: Microarray Suite, Instrument Control, Analysis Settings, and Report Settings ([Figure 4.1](#)).

- To display a different part of the shortcut bar, click the section name ([Figure 4.3](#) through [Figure 4.7](#)).

In LIMS mode, the Microarray Suite shortcut bar displays additional buttons: Publish, Sample History, and Workflow Monitor.

Microarray Suite Shortcut Bar

Figure 4.2 and **Figure 4.3** show the Microarray Suite shortcut bar at start up. The shortcut bar may display additional buttons, depending on the file types that are open. For example, if a report file (*.rpt) is open, the shortcut bar includes the **Chip Report** button . Click the report button to view the report in the main display area.

Click a button to perform the following functions:



Set up an experiment or display open experiment information files (*.exp).

Open the Batch Analysis window.

Open the GT Viewer.

Figure 4.2
Microarray Suite shortcut bar, disk files mode

Click a button to perform the following functions:

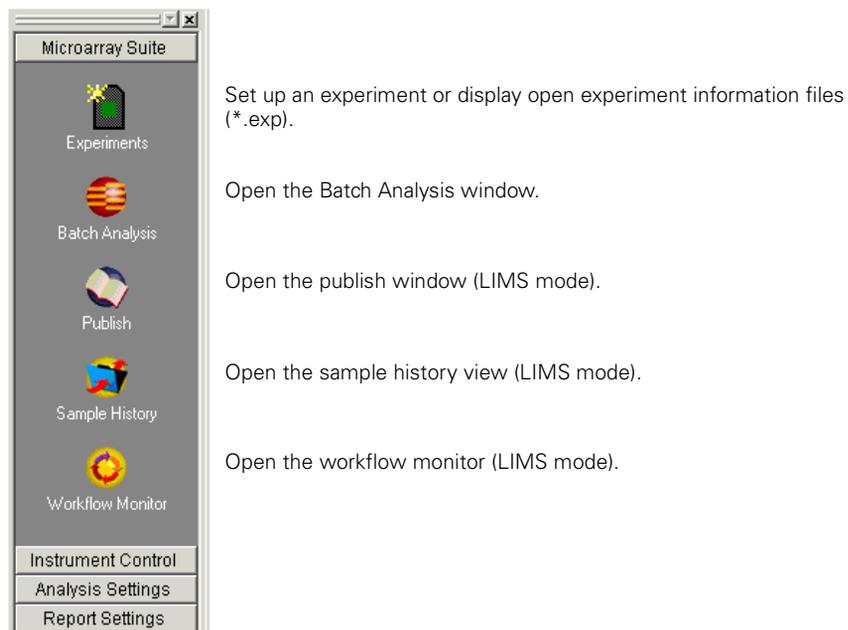


Figure 4.3
Microarray Suite shortcut bar, LIMS mode

Figure 4.4 shows additional buttons that may be included in the shortcut bar, depending on the file types open.

Click a button to perform the following functions:

 Image Views	Display open image data files (*.dat) or cell intensity files (*.cel).
 Chip Report	Display an open report file (*.rpt).
 Expression Analysis	Display open expression analysis output files (*.chp) in the Expression Analysis window (EAW).
 Nucleotide Analysis	Display open genotype analysis output files (*.chp) in the Nucleotide Analysis window (NAW) (disk files mode only).
 Sequence Analysis	Display open polymorphism and mutation analysis output files (*.chp) in the Sequence Analysis window (SAW) (disk files mode only).
 Hybridization Analysis	Display open hybridization analysis output files (*.chp) in the Hybridization Analysis window (HAW) (disk files mode only).

Figure 4.4
Other Microarray Suite shortcut buttons

Instrument Control Shortcut Bar

Click a button to perform the following functions:

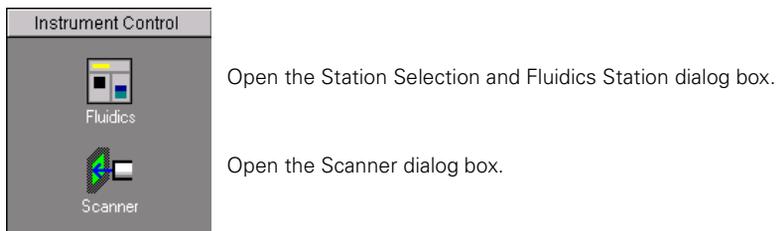


Figure 4.5
Instrument control shortcut bar

✓ Note

*The instrument control shortcut bar does not display instrument buttons if the workstation is not connected to the instruments, or if the default configuration settings specify no fluidics station or scanner installed (select **Tools** → **Defaults** from the menu bar to view the default settings).*

Analysis Settings Shortcut Bar

Click a button to view user-modifiable parameters of the probe array analysis algorithms.

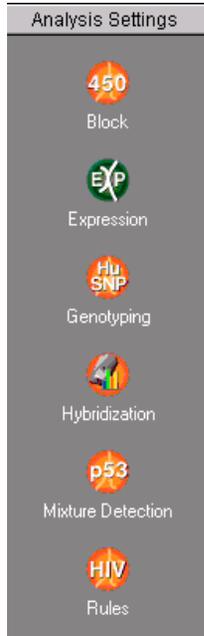


Figure 4.6
Analysis settings shortcut bar

✓ Note

Some Microarray Suite analysis algorithms have no user-modifiable parameters.

Report Settings Shortcut Bar

Click a button to view any user-modifiable report settings.



Figure 4.7
Report settings shortcut bar



Note

Some reports have no user-modifiable parameters.

Data File Tree

The data file tree displays:

- experiment information files (*.exp)
- image data files (*.dat)
- cell intensity files (*.cel)
- probe array analysis output or *chip* files (*.chp) (Figure 4.8).

(See Appendix B for information about the different types of Microarray Suite files.)

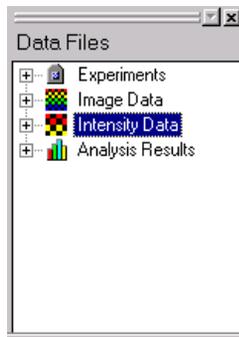


Figure 4.8
Data file tree

- Right-click a file in the tree to view a shortcut menu of commands (specific to the file type). From the data file tree you can:
 - open or analyze a file
 - display file information
 - generate a report

In LIMS mode, you can update the data file tree to show new information in the system.

- Right-click the data file tree and select **Refresh** from the shortcut menu.

Opening Files and Viewing File Information

1. In the data file tree, double-click the file name. Alternatively, right-click the file name in the tree and select **Open** from the shortcut menu.

⇒ This displays the file in the main display area.

Many files may be open at the same time, but you can view only one in the main display area (except for a cascade or tile view of image data (*.dat) or cell intensity (*.cel) files).

2. Click a Microarray Suite shortcut bar button (Figure 4.2 to Figure 4.4) to toggle the view between different types of open files.
3. To view file information, right-click the file name in the data file tree, then click **Information** in the shortcut menu.

⇒ An Information dialog box appears (Figure 4.9 and Figure 4.10).

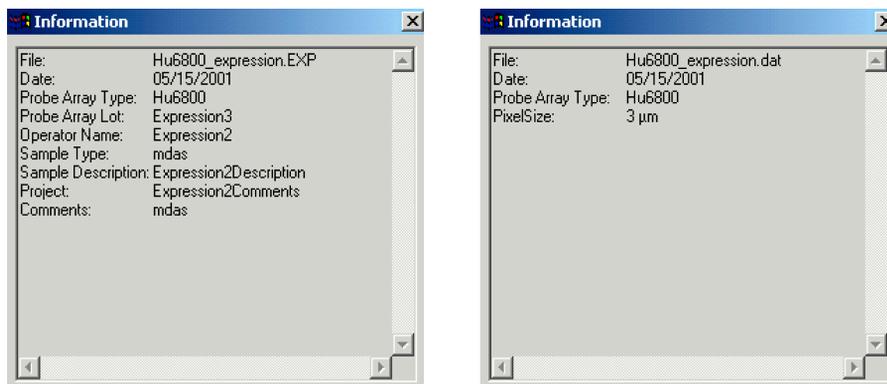


Figure 4.9

Information for an experiment information file (*.exp) (left) and image data file (*.dat) (right)

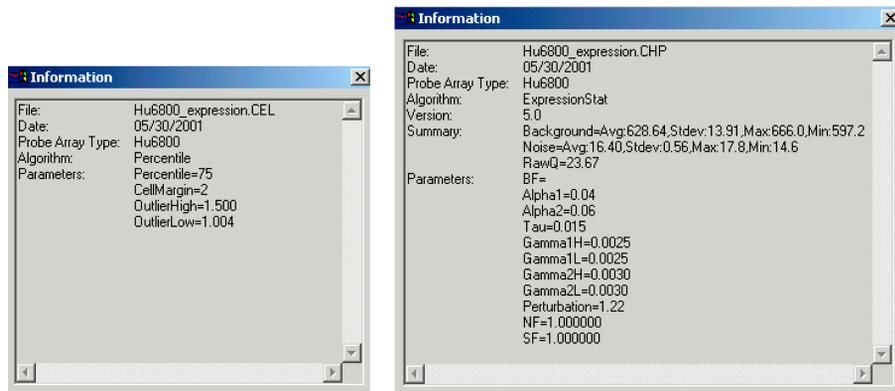


Figure 4.10
Information for a cell intensity file (*.cel) (left) and analysis output file (*.chp) (right)

Data File Tree Display Options

1. Click the minus [-] or plus [+] sign next to each file type icon to collapse or expand the list of files.
2. Click the **Data File Tree** toolbar button  to hide (or display) the data file tree.

Status Log

The status log displays system status messages ([Figure 4.1](#)).

1. To clear the messages, right-click the status log and click **Clear Messages** in the shortcut menu.
2. To mute the error message sound, right-click the status log and remove the check mark from the **Play Error Sound** option.
3. Click the **Status Log** button  in the main toolbar to hide (or display) the status log.

Configuring Microarray Suite

The Microarray Suite software may be run as a stand alone application (disk files mode) or with the Affymetrix® Laboratory Information Management System (LIMS) (LIMS mode).

Check the data storage and file locations settings to make sure Microarray Suite is properly configured for your system (disk files or LIMS mode).

Data Storage

1. Select **Tools** → **Defaults** from the menu bar.
⇒ The Defaults dialog box appears.
2. Click the Database tab (**Figure 4.11**)

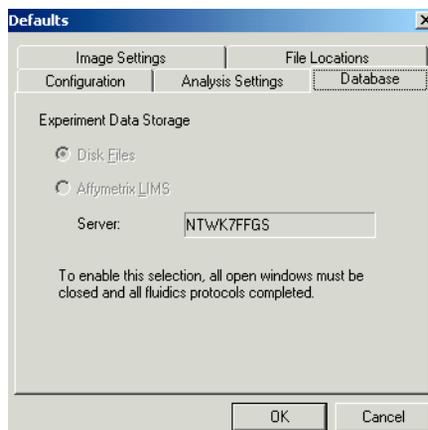


Figure 4.11
Defaults dialog box, Database tab

✓ Note

The Experiment Data Storage options are only available if all windows are closed and no instrument is active.

3. Choose the **Disk Files** option (default) to run Microarray Suite as a stand alone application.

In disk files mode, experiment data are stored on the local hard drive or a network drive.

4. Choose the **Affymetrix LIMS** option to run Microarray Suite with the LIMS.

In LIMS mode, experiment data are stored on the LIMS server.



Note

LIMS supports gene expression assay data only.

File Locations Settings

When you run Microarray Suite in disk files mode, experiment data are stored on disk files and there are three file locations. When you run the software in LIMS mode, experiment data are stored on a network drive and there are four file locations,

The file location paths determine where Microarray Suite looks for the files needed to run a fluidics protocol or analyze data from a particular probe array type, and specify where to save experiment data.

You may change some of the default file locations. Any modifications made by one user (identified by the logon name) do not affect the file locations of other users.

1. Select **Tools** → **Defaults** from the menu bar.
⇒ The Defaults dialog box appears.
2. Click the File Locations tab (**Figure 4.12**).

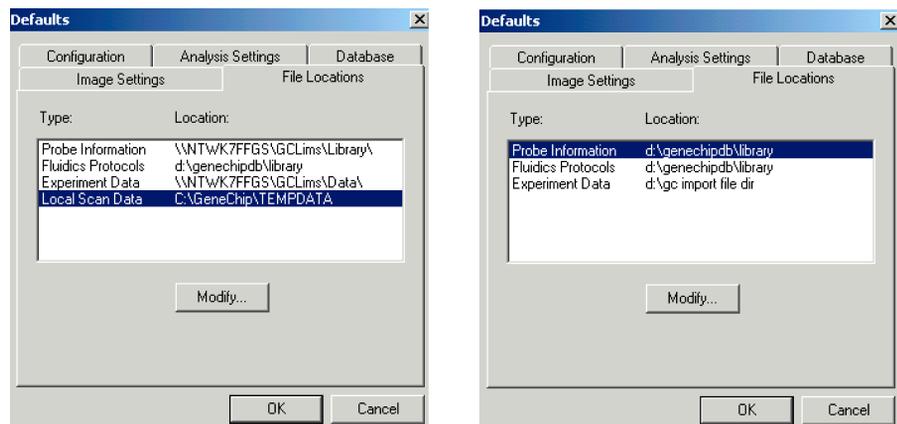


Figure 4.12

Defaults dialog box, File Locations tab, LIMS mode (left) and disk files mode (right)

3. To modify a path, highlight the file type of interest, then click **Modify** (or double-click the file type).

⇒ The Modify Location dialog box appears (Figure 4.13).

✓ Note

In LIMS mode, the paths on the network server for Probe Information and Experiment Data cannot be modified.

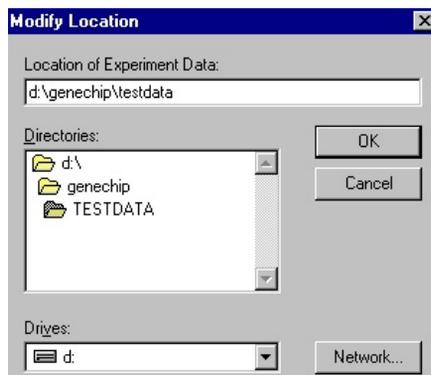


Figure 4.13
Modify Location dialog box

4. Select a drive from the **Drives** drop-down list, then select the desired directory path in the **Directories** box.

✓ Note

Microarray Suite can detect a network drive and writes a local temporary image data file (.dat). This file is deleted after the scan image has been transferred to the network drive.*

5. Click **Network** to map a network drive that is not already mapped. Network drives may also be mapped using Windows Explorer.
6. Click **OK** to close the Modify Location dialog box. Click **OK** to close the Defaults dialog box.

Microarray Suite Files

There are three types of Microarray Suite files (see Appendix B for more information about the file types):

- Probe Information or *library* files are unique for each probe array type and contain information about the probe array design characteristics, scanning parameters, and default analysis parameters.
- Fluidics Protocol files define the hybridization, wash, or stain protocols run by the GeneChip® Fluidics Station 400.
- Experiment Data files include the *.exp created by the user when an experiment is defined and other file types Microarray Suite generates during an analysis (for example, *.dat, *.cel, *.chp).

LIMS Filters

In LIMS mode, you may apply filters to the experimental data. The filters determine the data that Microarray Suite displays (for example, the files in the data file tree, the sample history view, workflow monitor, and the instrument control dialog boxes).

Filters are applied on a per user basis (identified by the logon name). The filters you specify do not affect the filters specified by other users.

✓ Note

Filters are not available when Microarray Suite is run in disk files mode.

1. Select **Tools** → **Filters** from the menu bar.
⇒ The Filters dialog box appears (Figure 4.14).

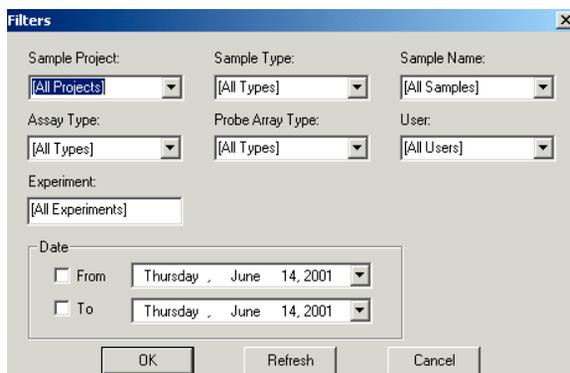


Figure 4.14
Filters dialog box, LIMS mode

2. Make selections from the drop-down lists to specify the filters. Choose date options if desired.
3. Click **OK** when finished to close the Filters dialog box and apply the filters.
⇒ The status bar indicates that filters are applied ([Figure 14.5](#)).

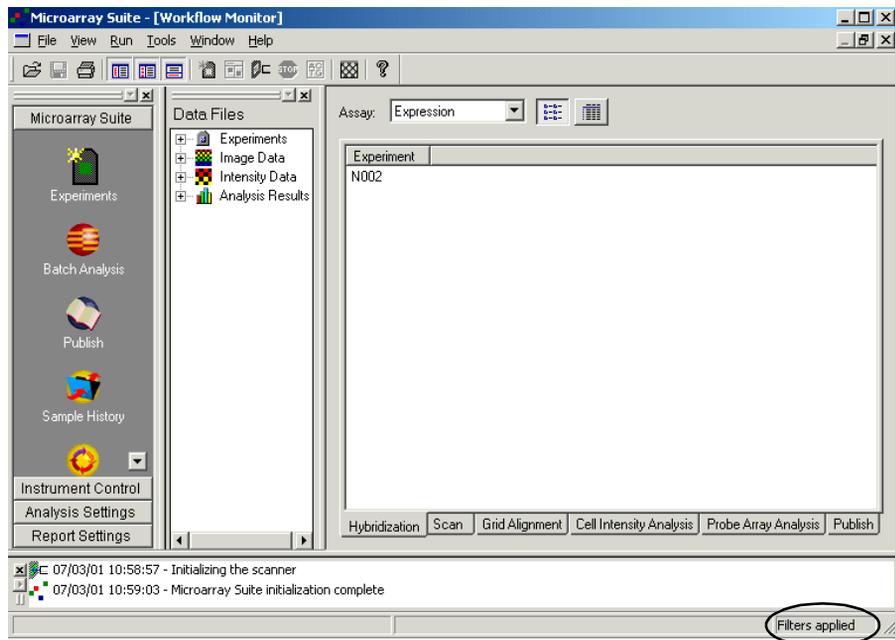
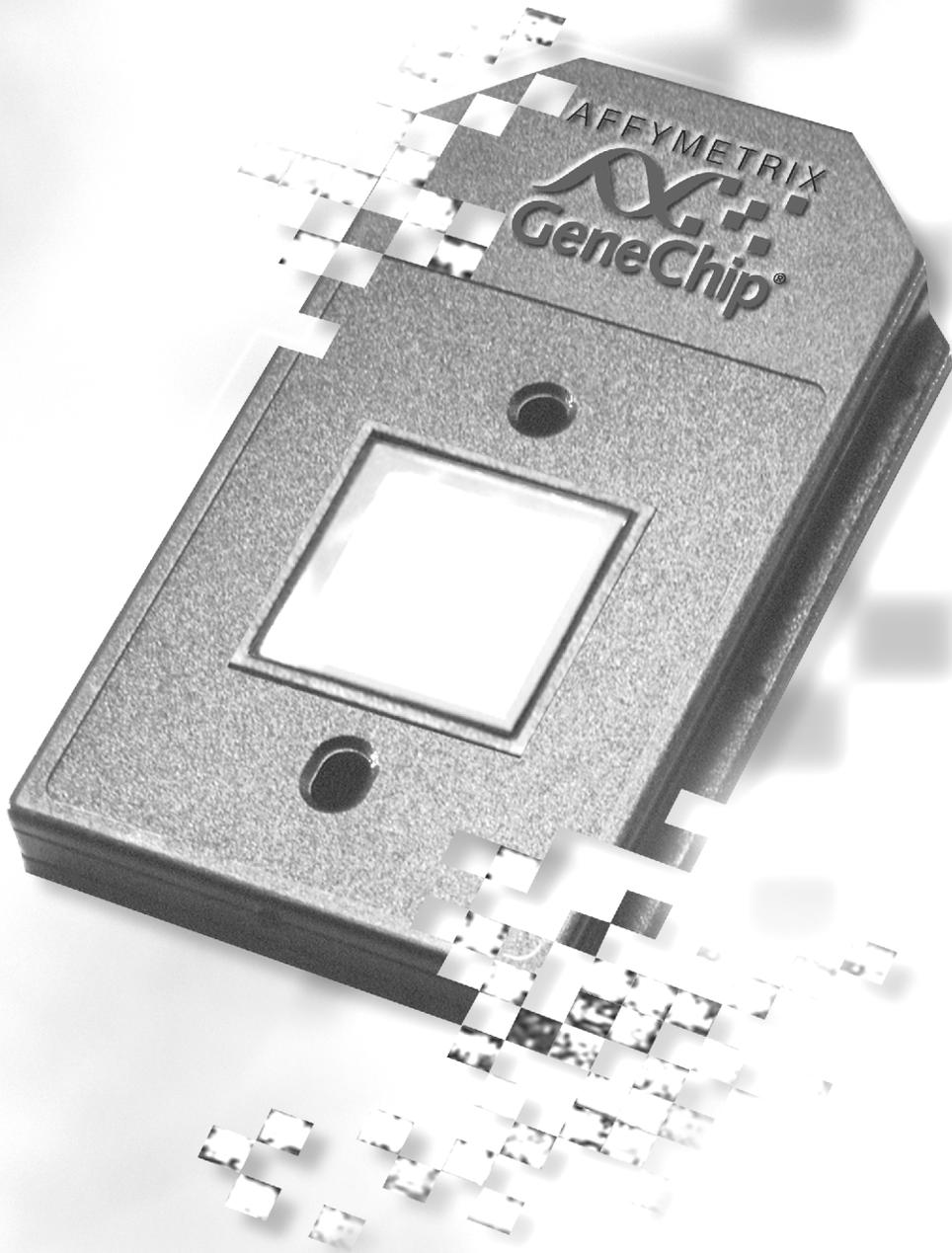


Figure 4.15
Status bar indicates LIMS filters applied



Chapter 5



Setting Up an Experiment

This chapter explains how to set up an experiment. You must define an experiment in Microarray Suite before processing a probe array in the fluidics station or scanning a probe array (Figure 5.1). Microarray Suite relies on the experiment information to direct the scanner and apply the correct cell and data analysis algorithms.

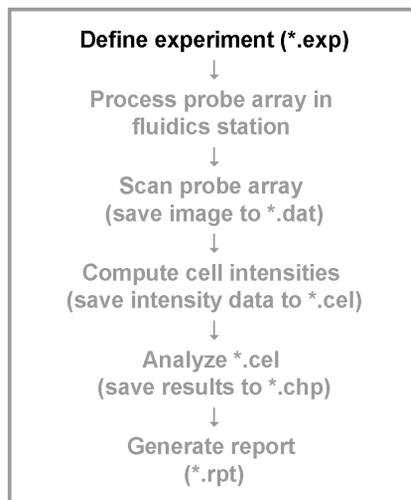
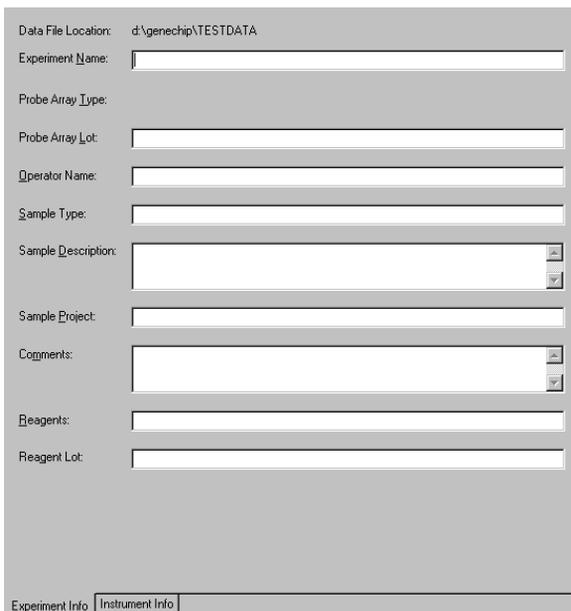


Figure 5.1
Assay & analysis flow chart

Setting Up an Experiment in Disk Files Mode

1. Click **Experiments**  in the Microarray Suite shortcut bar or main toolbar. Alternatively, select **File** → **New Experiment** from the menu bar.
⇒ The Experiment Information dialog box appears ([Figure 5.2](#)).
2. In the Experiment Info tab, enter a unique experiment (file) name for the probe array (up to 64 alphanumeric characters). (Required by Microarray Suite.)
3. Make a selection from the **Probe Array Type** drop-down list. (Required by Microarray Suite.)

The experiment name also serves as the file name for subsequent data files generated during the analysis. (See [Appendix B](#) for a description of Microarray Suite file types).



Data File Location: d:\genechip\TESTDATA

Experiment Name:

Probe Array Type:

Probe Array Lot:

Operator Name:

Sample Type:

Sample Description:

Sample Project:

Comments:

Reagents:

Reagent Lot:

Experiment Info | Instrument Info

Figure 5.2
Experiment Information (disk files mode)

4. Enter the remaining information as desired to help track pertinent information about the experiment.
5. When finished entering the information, click the **Save** toolbar button  or select **File** → **Save** from the menu bar.
 - ⇒ This creates an experiment information file (*.exp) in the current experiment data directory and displays the *.exp name in the data file tree.

Microarray Suite uses the information in the *.exp file to identify the probe array type. It records the hybridization and scanning parameters in the *.exp file.

In the Experiment Information dialog box, the **Data File Location** displays the directory location for the *.exp file and other types of experiment data files generated during the analysis. (For example, d:\genechip\TESTDATA in [Figure 5.2](#).)

6. Click the Instrument Info tab to view information about the fluidics protocol and scanning parameters captured by Microarray Suite after the probe array is processed in the fluidics station and scanned ([Figure 5.3](#)).

This page is blank until the probe array has been processed in the fluidics station.

The screenshot shows a dialog box titled "Experiment Information" with two tabs: "Experiment Info" (selected) and "Instrument Info". The "Experiment Info" tab contains the following fields and values:

Data File Location:	d:\genechip\TESTDATA		
Experiment Name:	113099		
Probe Array Type:	Hu6800		
Protocol:	EukGE-wS2		
Wash A1 Temperature (C):	25		
Number of Wash A1 Cycles:	10		
Mixes per Wash A1 Cycle:	2		
Wash B Temperature (C):	50		
Number of Wash B Cycles:	4		
Mixes per Wash B Cycle:	15		
Stain Temperature (C):	25		
First Stain Time (seconds):	600		
Wash A2 Temperature (C):	25		
Number of Wash A2 Cycles:	10		
Mixes per Wash A2 Cycle:	4		
Second Stain Time (seconds):	600		
Third Stain Time (seconds):	600		
Wash A3 Temperature (C):	30		
Number of Wash A3 Cycles:	15		
Mixes per Wash A3 Cycle:	4		
Holding Temperature (C):	25		
Fluidics Station:	1	Module:	1
		Date:	11/30/99 14:49
Pixel Size:	3 µm	Filter:	570 nm
		Date:	11/30/99 15:09
Scanner ID:		# of Scans:	2

At the bottom of the dialog box, there are two tabs: "Experiment Info" and "Instrument Info".

Figure 5.3
Hybridization and scan information

- To close the Experiment Information dialog box, click the **Close** button  or select **File** → **Close** from the menu bar.

Setting Up an Experiment in LIMS Mode

Registering a Sample

You identify or *register* a sample by entering sample information or *attributes* and associating the sample with a project (which may include several different samples) (Figure 5.4). The registration process adds the sample information to the process database on the LIMS server.

Defining an Experiment

To define an experiment, specify the type of probe array to which the target (sample) will be hybridized as well as other information relevant to the experiment. You may define one or more experiments per sample (for example, the different probe arrays of a multiple probe array set hybridized to the same sample).

After you define an experiment, the Hybridization tab of the workflow monitor displays the experiment and associated sample. The Hybridization tab displays experiments that have not yet been processed in the fluidics station. (See [Chapter 6](#) for more information about the workflow monitor.)

Microarray Suite automatically adds the experiment information to the process database. LIMS uses this information to maintain relationships between the experimental data files, manage the workflow queue, and identify the user-modifiable algorithm parameters for data analysis.

**Note**

LIMS supports expression data only.

1. In the Microarray Suite shortcut bar, click **Experiment Info** , or select **File** → **New Experiment** from the menu bar.
⇒ The Experiment Information dialog box appears ([Figure 5.4](#)).

The screenshot shows the 'Experiment Information' dialog box in LIMS mode. It is divided into two main sections: 'Sample Template' and 'Experiment Template'. Each section has a dropdown menu for selecting a template (currently set to '[No template]') and a table for entering information. The 'Sample Template' table has columns 'Name' and 'Value', with rows for 'Sample Name', 'Sample Type', 'Project', and 'User' (pre-filled with 'Administrator'). The 'Experiment Template' table has similar columns and rows for 'Experiment Name', 'Probe Array Type', and 'User' (pre-filled with 'Administrator'). Below these tables are dropdowns for 'User Set' (set to '[No analysis]') and 'Publish Database' (set to '[No publishing]'), along with a checkbox for 'Publish [ntensities]'. At the bottom, there are tabs for 'Experiment Info' and 'Instrument Info'.

Two callout boxes point to the tables:

- The first callout points to the 'Sample Template' table and contains the text: "Enter information here to register the sample".
- The second callout points to the 'Experiment Template' table and contains the text: "Enter information here to define the experiment (*.exp)".

Figure 5.4
Experiment Information dialog box, Experiment Info tab (LIMS mode)

2. Enter a sample name (*.exp file name, up to 64 alphanumeric characters). (Required by Microarray Suite.) Or, if you wish to use a sample template, make a selection from the **Sample Template** drop-down list and enter the sample information.

You may enter sample and experiment information using a template defined by the LIMS administrator. A template specifies the information fields available in the Experiment Information dialog box. See the *Affymetrix® LIMS Manager User Manual* for more information about templates.

3. Enter a unique experiment name (*.exp file name, up to 64 alphanumeric characters). (Required by Microarray Suite.) Or, if you wish to use an experiment template, make a selection from the

Experiment Template drop-down list and enter the experiment information.

The experiment name also serves as the file name for subsequent data files generated during the analysis. (See Appendix B for a description of Microarray Suite file types.)

✓ Note

In the Experiment Info tab, items in bold are required entries.

4. Make a selection from the **Probe Array Type** drop-down list. (Required by Microarray Suite.)
5. To automatically publish the analysis output (*.chp):
 - a. Make a selection from the **User Set** drop-down list.

The LIMS administrator defines the user set and specifies the values for the user-modifiable expression algorithm parameters. See the *Affymetrix LIMS Manager User Manual* for more information about user sets.

- b. Make a selection from the **Publish Database** drop-down list. (See [Chapter 14](#) for more information about publishing data.)

✓ Note

*You must specify both a user set and a publish database, otherwise the *.chp will not be published. If you specify only a user set, the analysis proceeds using the analysis settings in the Expression Analysis Settings dialog box, but the data are not published. (To view the Expression Analysis Settings dialog box, click **Expression**  in the Analysis Settings shortcut bar).*

- c. To include the cell intensity data (*.cel) in the publish database, choose the **Publish Intensities** option.

✓ Note

Publishing cell intensity data (.cel) uses large amounts of computer memory. For example, publishing the analysis output (*.chp) for a high density probe array requires approximately 2 MB of disk space compared to 30 MB for both the *.chp and *.cel data.*

6. Click the **Save** toolbar button .
⇒ This saves the experiment and displays it in the data tree.

✓ Note

If you specified a user set and publish database during experiment setup, the software prompts you for the publish database password before the experiment is saved.

After the probe array is processed in the fluidics station and scanned, click the Instrument Info tab to view information about the fluidics protocol and scanning parameters captured by Microarray Suite (Figure 5.5).

The instrument information page is blank until the probe array has been processed in the fluidics station.

Data File Location: \\ntwk98z63\gclims\data

Experiment Name: 113099

Probe Array Type: Hu6800

Protocol: Wash A1 Temperature (C): 25
EukGE-WS2 Number of Wash A1 Cycles: 10
Mixes per Wash A1 Cycle: 2
Wash B Temperature (C): 50
Number of Wash B Cycles: 4
Mixes per Wash B Cycle: 15
Stain Temperature (C): 25
First Stain Time (seconds): 600
Wash A2 Temperature (C): 25
Number of Wash A2 Cycles: 10
Mixes per Wash A2 Cycle: 4
Second Stain Time (seconds): 600
Third Stain Time (seconds): 600
Wash A3 Temperature (C): 30
Number of Wash A3 Cycles: 15
Mixes per Wash A3 Cycle: 4
Holding Temperature (C): 25

Fluidics Station: 1 Module: 1 Date: 11/30/99 14:49
Pixel Size: 3 µm Filter: 570 nm Date: 11/30/99 15:09
Scanner ID: # of Scans: 2

Experiment Info Instrument Info

Figure 5.5
Experiment Information dialog box, Instrument Info tab (LIMS mode)

Viewing or Editing Experiment Information

1. To display the experiment information, double-click the experiment name in the data file tree.
2. To edit an entry, highlight the entry and enter the new information. Alternatively, right-click to open a shortcut menu of edit commands (Figure 5.6).

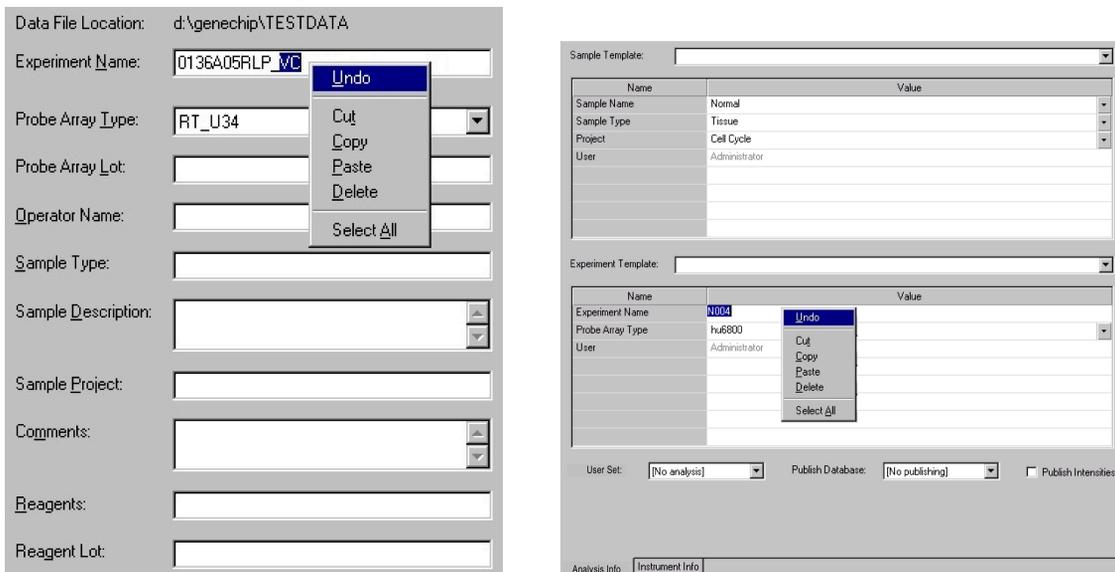


Figure 5.6
Experiment information file (*.exp) shortcut menu of edit commands (disk files mode left, LIMS mode right)

3. To clear all fields, select **Edit** → **Clear** from the menu bar (equivalent to closing the *.exp file).
4. To permanently delete the experiment from the system, select **Edit** → **Delete** from the menu bar.

✓ Note

In LIMS mode, only users authorized by the System Administrator may delete data from the system using the LIMS Manager software.

5. When finished, click the **Close** button  in the upper right corner of the window or select **File** → **Close** from the menu bar.

When the experiment information file is closed, Microarray Suite prompts you to save changes.



Chapter 6



LIMS Sample History & Workflow Monitor

The sample history feature and workflow monitor are available in LIMS mode. These are tools that help you track samples and experimental data.

Sample History

The sample history feature provides two different views of the samples in the system:

- **File view** shows the experimental data files derived from a particular sample
- **Process view** shows all stages (sample registration, experiment setup, hybridization, scan, grid alignment, cell intensity analysis, probe array analysis, and publishing) that are completed or pending for a sample

Displaying Sample History

1. In the Microarray Suite shortcut bar, click **Sample History** .
⇒ The sample history file view appears ([Figure 6.1](#)).

The file view is the default. If the file view is not displayed, select **View** → **Files** from the menu bar.

The expression folder in the sample tree (left pane) contains all registered expression analysis samples.



Note

If filters have been applied, not all registered samples may be displayed.

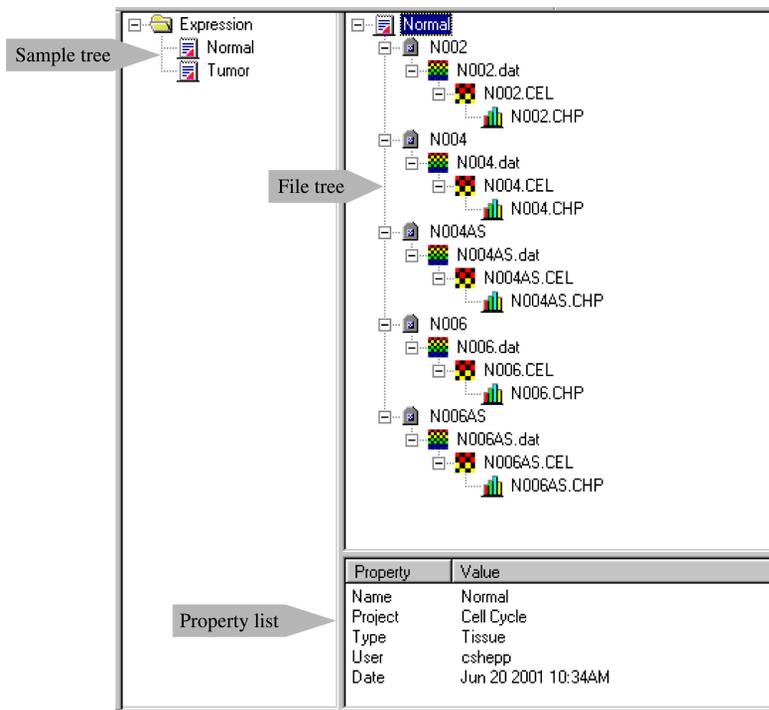


Figure 6.1
Sample history, file view

2. To display a file tree of the experimental data files derived from a sample, click the sample name in the sample tree.
3. To display a list of file properties and their values (bottom right pane), click a sample or a file name in the file tree.

Displaying Sample History Process View

The process view shows all completed or pending stages for a particular sample.

1. In the Microarray Suite shortcut bar, click **Sample History** .
⇒ The sample history view appears.
2. To display the process view (**Figure 6.2**), select **View** → **Processes** from the menu bar.

The expression folder in the sample tree (left pane) contains registered expression analysis samples.

✓ Note

If filters have been applied, not all registered samples may be displayed. (See [Chapter 4](#) for more information about LIMS filters.)

The screenshot displays the software interface for viewing sample history. On the left, the 'Sample tree' shows a hierarchy with 'Expression' as the parent folder, containing sub-folders for 'Normal' and 'Tumor'. The main area on the right, labeled 'File tree', shows a detailed view of a selected sample's workflow. The workflow steps include: Sample Registration (with sub-items for Normal <sample> and Normal <vessel>), Experiment Setup (with sub-items for Inputs, Normal <sample>, and Outputs), Hybridization (with sub-items for Inputs, Normal <vessel>, and N002 <chip>), Scan (with sub-items for Inputs, N002 <chip>, and Outputs, N002.dat), Grid Alignment (with sub-items for Inputs and N002.dat), Cell Intensity Analysis (with sub-items for Inputs, N002.dat, and Outputs, N002.CEL), Probe Array Analysis (with sub-items for Inputs, N002.CEL, and Outputs, N002.CHP), and finally Publish. At the bottom, a 'Property list' table provides details for the selected sample.

Property	Value
Name	Normal
Project	Cell Cycle
Type	Tissue
User	cshepp
Date	Jun 20 2001 10:34AM

Figure 6.2
Sample history, process view

3. To display the process tree, click a sample name in the sample tree. Gray bullets indicate completed processes and green bullets indicate those that are pending.

4. Click the plus sign [+] to the left of a process name to expand the information in the tree (Figure 6.2). Use the scroll bars to the right or below the process tree to view the expanded information.

The tree displays the inputs and outputs for each process. For example, a cell intensity file (*.cel) is the input to the Probe Array Analysis process that generates an analysis output file (*.chp).

5. To display a list of file properties (bottom right pane), click one of the following in the process tree (Figure 6.2):

- sample name
- experiment name
- a process input or output

Renaming a Sample

1. In the sample tree, right-click the sample and click **Rename Sample** in the shortcut menu.

⇒ The Rename Sample dialog box appears (Figure 6.3).

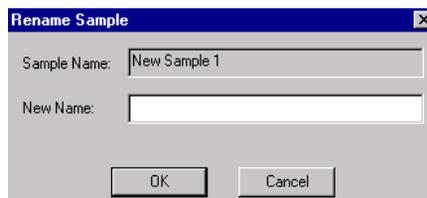


Figure 6.3

Rename Sample dialog box

2. Enter a new sample name and click **OK**.

Workflow Monitor

When you define an experiment, it automatically enters the workflow and may be tracked through the sequential tabs in the workflow monitor (see [Table 6.1](#)).

Table 6.1
Workflow monitor tabs

Workflow Monitor Tab	Displays...
Hybridization	Experiments that have not been processed in the fluidics station.
Scan	Experiments that have been processed in the fluidics station, but have not been scanned.
Grid Alignment	Image data files (*.dat) that require manual grid alignment.
Cell Intensity Analysis	Image data files (*.dat) that have been manually aligned, but have not been analyzed by the Cell analysis algorithm (no cell intensity file (*.cel) exists)
Probe Array Analysis	Cell intensity files (*.cel) that have not been analyzed by the Expression analysis algorithm.
Publish	analysis output files (*.chp) that have not been published.

Opening the Workflow Monitor

1. In the Microarray Suite shortcut bar, click **Workflow Monitor** .
⇒ The workflow monitor appears ([Figure 6.4](#)).

Note

If filters have been applied, not all registered samples or experimental data may be displayed. (See [Chapter 4](#) for more information about LIMS filters.)

Assay: Expression

Experiment	Sample Name	Sample Project	Type	User	Queue Date
N004	Normal	Cell Cycle	Tissue	cshepp	Jun 20 2001 10:34AM
N004AS	Normal	Cell Cycle	Tissue	cshepp	Jun 20 2001 10:34AM
N006	Normal	Cell Cycle	Tissue	cshepp	Jun 20 2001 10:34AM
N006AS	Normal	Cell Cycle	Tissue	cshepp	Jun 20 2001 10:35AM
T000	Tumor	Cell Cycle	Tissue	cshepp	Jun 20 2001 10:39AM
T000AS	Tumor	Cell Cycle	Tissue	cshepp	Jun 20 2001 10:39AM
T002	Tumor	Cell Cycle	Tissue	cshepp	Jun 20 2001 10:39AM
T002AS	Tumor	Cell Cycle	Tissue	cshepp	Jun 20 2001 10:39AM

Hybridization | Scan | Grid Alignment | Cell Intensity Analysis | Probe Array Analysis | Publish

Figure 6.4
Workflow monitor, details view

- To update the display, select **View** → **Refresh** from the menu bar.
This refreshes the workflow monitor so that it displays any new items added to the system since it was last opened.
- To view a list of experiments, click the **List** button . To view sample and project information for each experiment, click the **Details** button .
- To view experiment or data file information, right-click the item and select **Open Item** from the shortcut menu. Alternatively, select **View** → **Open Item** from the menu bar.

Workflow Monitor Tabs

- Click a tab to view the experiments in that part of the workflow queue.
- To update the workflow monitor, select **View** → **Refresh** from the menu bar.

Hybridization Tab

The Hybridization tab displays experiments that have not been processed in the fluidics station. An experiment is moved to the Scan tab after it has been processed in the fluidics station.

If a probe array is not processed (hybridized, or washed and stained) in the fluidics station, the experiment must be manually advanced to the Scan tab.

To advance an experiment to the Scan tab:

- Right-click the experiment and click **Advance to Scan** in the shortcut menu. Alternatively, click the experiment and select **View** → **Advance to Scan** from the menu bar.

Scan Tab

The Scan tab displays experiments that have been processed in the fluidics station, but have not been scanned. After the scan and grid alignment processes are completed, the cell intensity analysis is automatically run and the experiment is moved to the Probe Array Analysis tab. If the grid was not successfully aligned by the Alignment algorithm, the Grid Alignment tab displays the experiment.

Grid Alignment Tab

The Grid Alignment tab displays image data files (*.dat) that require manual grid alignment.

To manually align the grid:

1. Right-click the *.dat file in the Grid Alignment tab and click **Open Item** in the shortcut menu.
⇒ This displays the image data file (*.dat).
2. Place the mouse arrow over the grid perimeter (the arrow becomes a double arrow, ) (Figure 6.5).
3. Use the click-and-drag method or the keyboard arrow keys to adjust the horizontal and vertical position of the grid

The diagonal orientation of the double arrow along the perimeter of each corner probe cell indicates horizontal and vertical adjustments can be made simultaneously using the click-and-drag method or the keyboard arrow keys.

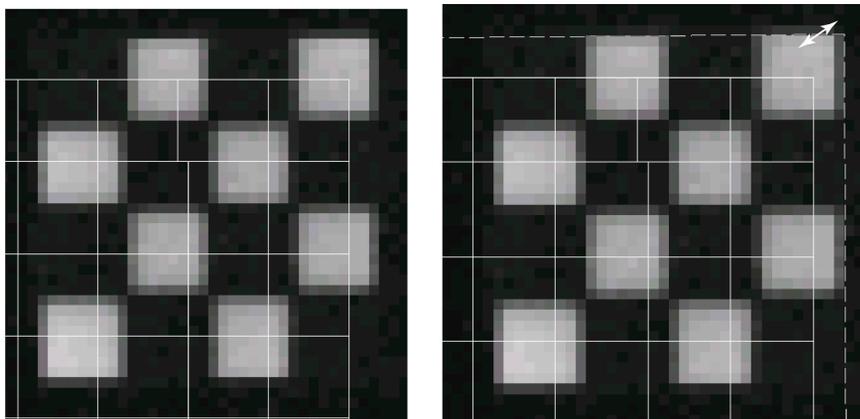


Figure 6.5
Misaligned grid (left) and adjusting grid alignment (right)

Cell Intensity Analysis Tab

The Cell Analysis tab displays image data files (*.dat) after the grid has been manually aligned. Microarray Suite does not automatically generate a cell intensity file (*.cel) from an image data file (*.dat) that requires manual alignment. As a result, no *.cel files exist for the *.dat files in the Cell Analysis tab.

To generate the cell intensity file:

- In the Cell Analysis tab, right-click the *.dat and click **Analyze Item** in the shortcut menu.
 - ⇒ This generates the *.cel and places it in the Probe Array Analysis tab.

To generate the cell intensity (*.cel) and analysis output file (*.chp):

1. In the Cell Analysis tab, right-click the *.dat and click **Open Item** in the shortcut menu.
 - ⇒ This displays the *.dat.
2. Select **Run Analysis** from the menu bar.
 - ⇒ This generates the *.cel and the *.chp.

The Publish tab displays the *.chp.

Probe Array Analysis Tab

The Probe Array Analysis tab displays cell intensity files (*.cel) that have not been analyzed by the expression analysis algorithm (no analysis output file (*.chp) exists).

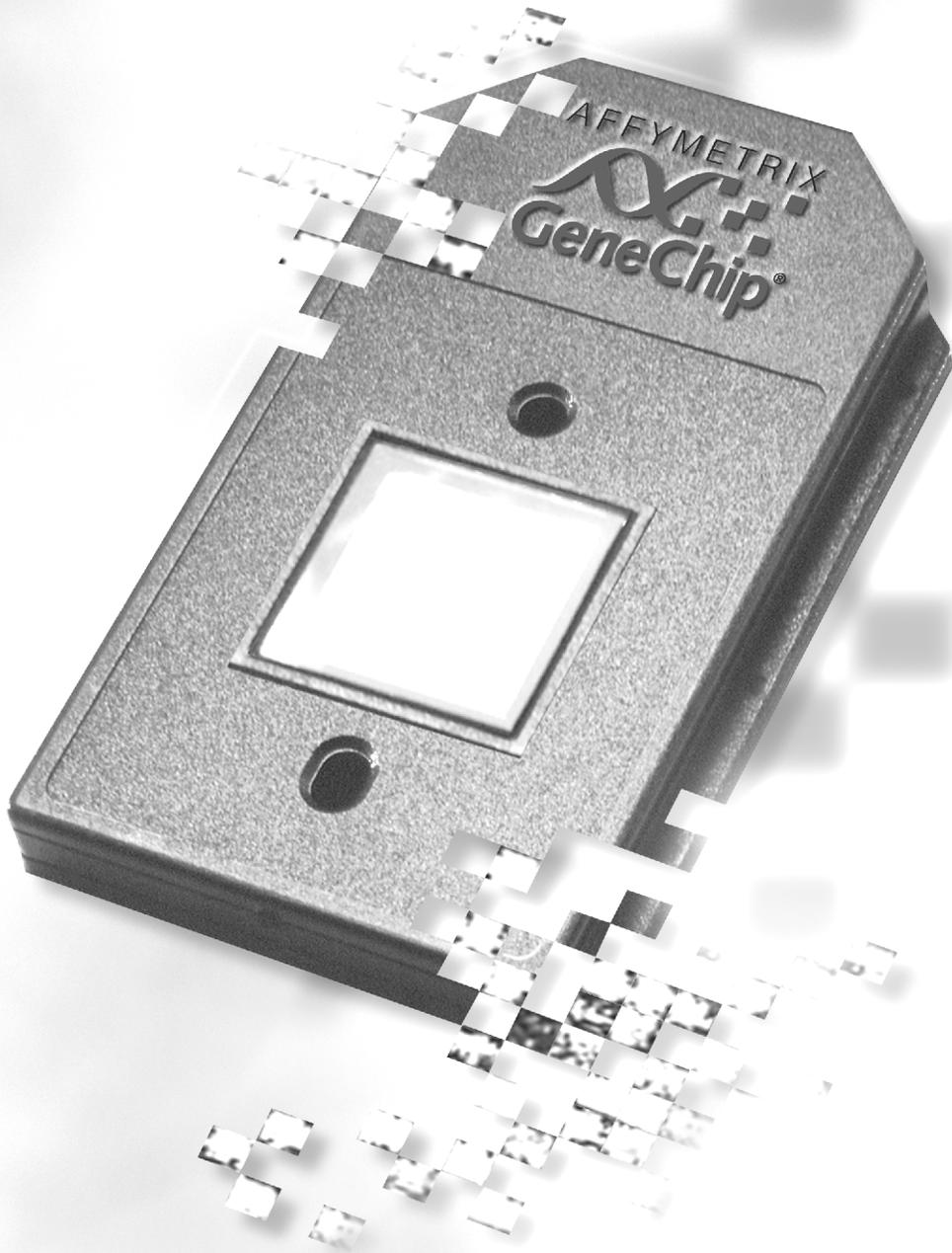
To analyze a *.cel:

- in the Probe Array Analysis tab, right-click the *.cel and click **Analyze Item** in the shortcut menu. Alternatively, select **View** → **Analyze Item** from the menu bar.

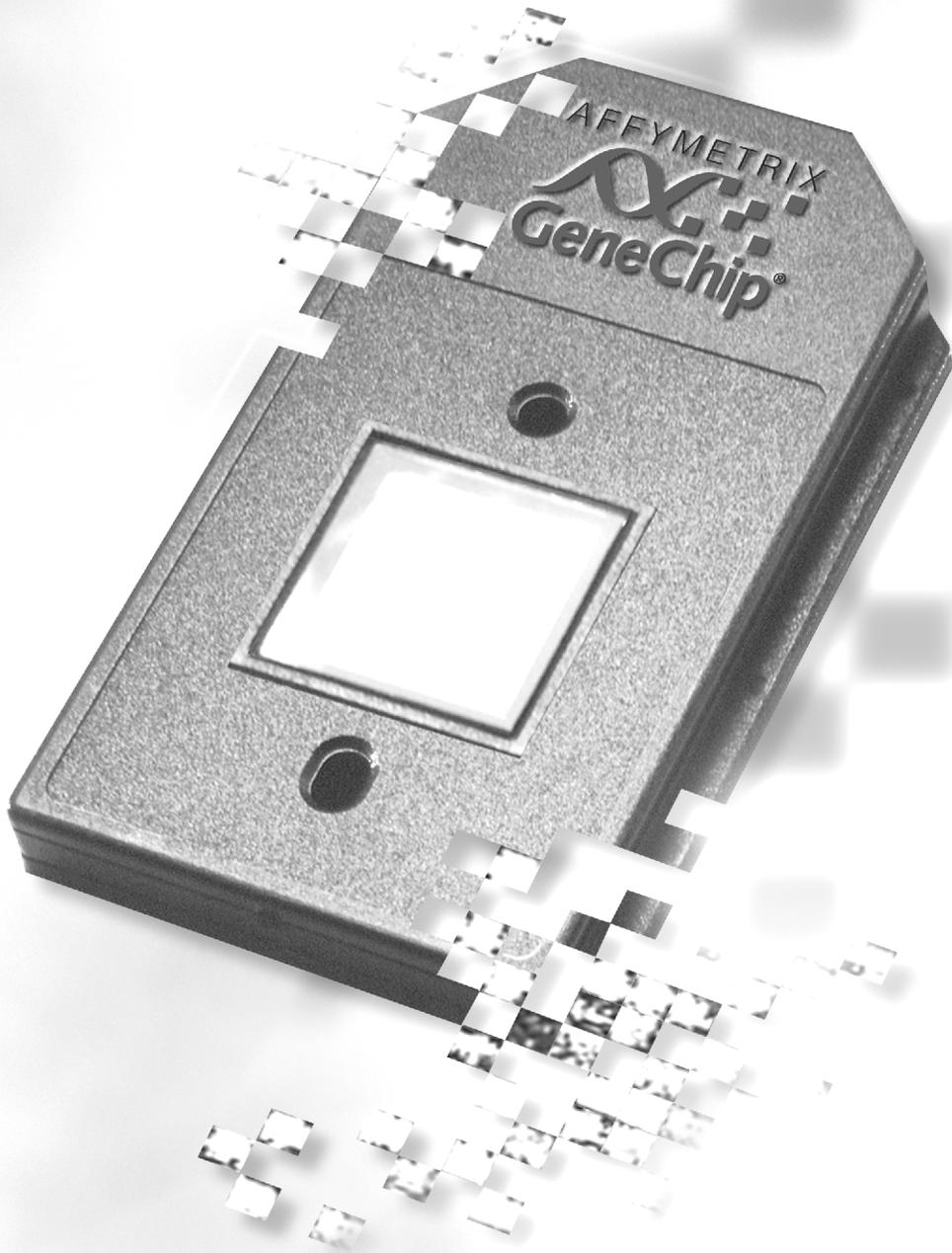
After the *.cel file has been analyzed and the analysis output file (*.chp) has been generated, the experiment is moved to the Publish tab.

Publish Tab

The Publish tab displays analysis output files (*.chp) that have not been published. The *.chp is removed from the Publish tab and the workflow monitor after it is published.



Chapter 7



Controlling the Instruments

This chapter describes how to use *Microarray Suite* to control the *GeneChip® Fluidics Station 400* and the *Agilent GeneArray® Scanner* (Figure 7.1).

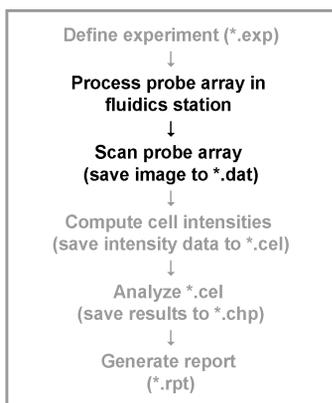


Figure 7.1
Assay & analysis flow chart

The Fluidics Station 400

The fluidics station is used to hybridize, wash, and stain probe arrays. One workstation can control up to eight fluidics stations. Each fluidics station contains four modules and each module can independently process a probe array using different fluidics protocols. Refer to the *GeneChip® Fluidics Station 400 User's Guide* for a description of the instrument, its components, and set up.

**Note**

Before you use the fluidics station, check the fluidics station configuration and prime the fluidics station with appropriate buffer.

Configuring the Fluidics Station

Before running a fluidics protocol, check to make sure the fluidics station(s) is properly configured.

1. Select **Tools** → **Defaults** from the menu bar.
⇒ The Defaults dialog box appears ([Figure 7.2](#)).

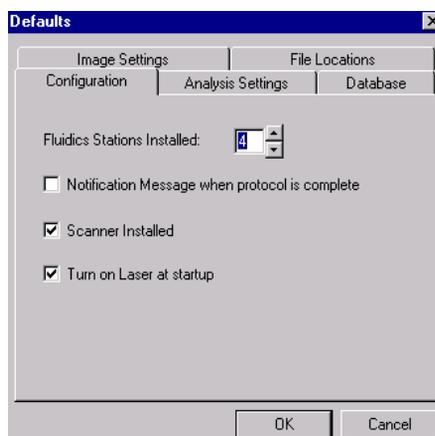


Figure 7.2
Defaults dialog box, Configuration tab

2. Click the Configuration tab.
3. Confirm the number of fluidics stations installed is correct or enter a new value.
4. If desired, check mark **Notification Message when protocol is complete** to display a notification message when a fluidics protocol is completed.
5. Click **OK** to close the Defaults dialog box.

Priming the Fluidics Station

Priming fills the fluidics station lines with appropriate buffer. You must prime the GeneChip® Fluidics Station 400:

- when the fluidics station is first turned on
 - before you run a fluidics protocol
 - when the wash solution is changed
 - before washing, after a shutdown has been performed
 - when a module LCD window instructs the user to prime
1. Click the **Fluidics** button  in the Instrument Control shortcut bar or the main toolbar. Alternatively, select **Run** → **Fluidics** from the menu bar.
 - ⇒ If more than one fluidics station is installed on the workstation, the Station Selection dialog box appears (Figure 7.3).
 - ⇒ If only one fluidics station is installed on the workstation, the Fluidics Station dialog box appears (Figure 7.4).
 2. If more than one fluidics station is installed on the workstation, select the number designation of the current fluidics station from the **Station Number** drop-down list.



Figure 7.3
Station Selection dialog box

3. Click **OK** to close the Station Selection dialog box.
 - ⇒ The **Fluidics Station** dialog box for the currently selected station appears (Figure 7.4).

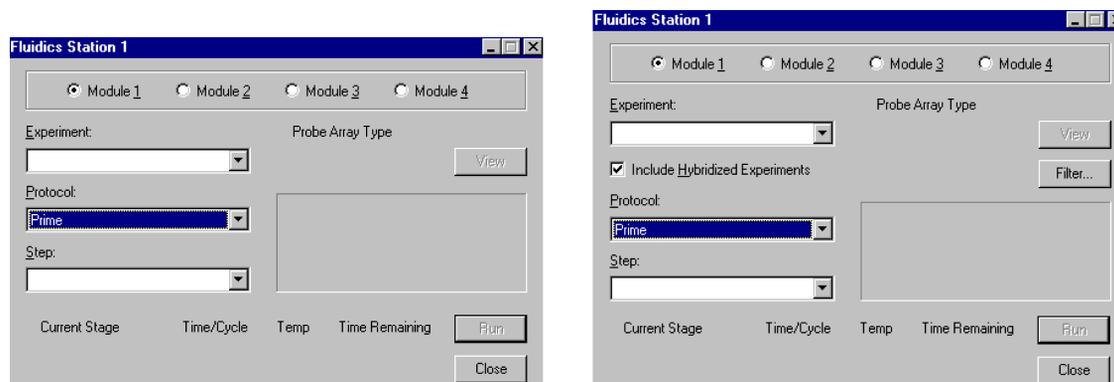


Figure 7.4 Fluidics Station dialog box, Prime protocol selected, disk files mode (left) and LIMS mode (right)

4. Select *Prime* from the **Protocol** drop-down list and *No Probe Array* from the **Experiment** drop-down list for each module to be used.
5. Fill the intake buffer reservoirs A and B with the appropriate priming buffer.
(Refer to the appropriate GeneChip[®] probe array package insert or the *Expression Analysis Technical Manual* for buffer preparation.)
6. Empty the waste bottle and fill the water reservoir with water.
7. Load an empty, standard 1.5 mL microcentrifuge tube in the sample holder of each module to be primed.
8. Click **Run** for each module to be primed and follow the prompts in the **Fluidics Station** dialog box (also shown in the module LCD window).

The Fluidics Station dialog box and the module LCD window display the status of the procedure. The fluidics station is ready to use when priming is completed and **Module primed, Ready** appears in the module LCD window.

Setting Up a Fluidics Protocol

1. In the Instrument Control shortcut bar, click **Fluidics** . Alternatively, select **Run** → **Fluidics** from the menu bar.
 - ⇒ If more than one fluidics station is installed on the workstation, the **Station Selection** dialog box appears (Figure 7.3).
 - ⇒ If only one fluidics station is installed on the workstation, the **Fluidics Station** dialog box appears (Figure 7.4).
2. If more than one fluidics station is installed on the workstation, select the number designation of the current fluidics station from the **Station Number** drop-down list.

Disk Files Mode

1. In the **Fluidics Station** dialog box (Figure 7.5):
 - a. Select the current fluidics station module.
 - b. Choose an experiment and protocol from the drop-down lists.

(Refer to the appropriate GeneChip® probe array package insert, the *Expression Analysis Technical Manual*, or the *HuSNP™ Mapping Assay User Manual* for protocol information.)

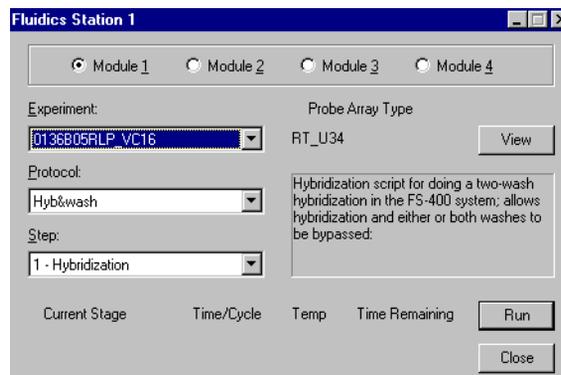


Figure 7.5
Fluidics Station dialog box, disk files mode

2. To display information about the selected experiment, click **View**.
⇒ This displays the experiment information window for the selected experiment.
3. Click **Run** to start the protocol on the selected module.
4. Repeat as necessary for other modules in the fluidics station(s) (**Figure 7.5**).

LIMS Mode

1. In the **Fluidics Station** dialog box (**Figure 7.6**):
 - a. Select the current fluidics station module.
 - b. Choose an experiment and protocol from the drop-down lists.
(Refer to the appropriate GeneChip® probe array package insert or the *Expression Analysis Technical Manual* for protocol information.)

✓ Note

If filters have been applied, the **Experiment** drop-down list may not include all experiments. Click **Filter** to open the **Filters** dialog box and view or specify filters.

Choose the **Include Hybridized Experiments** option to include hybridized experiments in the drop-down list.

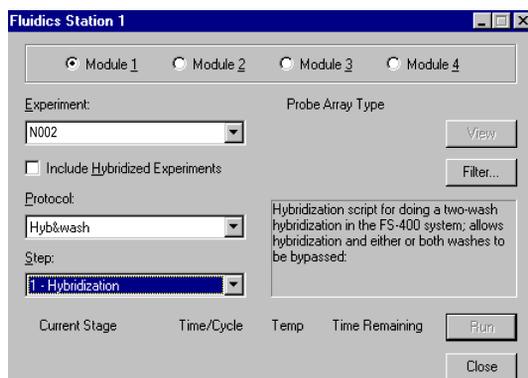


Figure 7.6
Fluidics Station dialog box, LIMS mode

2. Click **View** to display information about the selected experiment.
⇒ This displays the experiment information window for the selected experiment.
3. Click **Run** to start the protocol on the selected module.
4. Repeat as necessary for other modules in the fluidics station(s).

Running the Fluidics Protocol

1. Load the probe array and sample vial holder in each active module.
Sensors in the fluidics station detect when the probe array and sample vial holder have been loaded. The process will proceed automatically from this point. The Fluidics Station dialog box and each module LCD window display the status of the procedure.
2. After the protocol is finished, remove the probe array and inspect the probe array window for air bubbles.
If air bubbles are present, reinsert the probe array into the fluidics station to automatically drain and refill the probe array with the last wash buffer used. (Refer to the appropriate GeneChip® probe array package insert.) If no bubbles are present, the probe array is ready to be scanned.

Resuming a Fluidics Protocol

Microarray Suite software tracks the progress of a fluidics protocol run. If the protocol stops before completion, it can be resumed at the point where it was interrupted.

✓ Note

*The resume feature is only available for fluidics protocols that display multiple steps in the **Step** drop-down list of the Fluidics Station dialog box.*

1. When ready to resume the protocol, choose the protocol from the **Protocol** drop-down list in the Fluidics Station dialog box.
2. Click **Run** to continue the protocol from the point where it was interrupted.

Bypassing Steps in a Fluidics Protocol

Some multi-step fluidics protocols can be started at any step, so that part of a protocol can be bypassed.

1. In the Instrument Control shortcut bar, click **Fluidics** . Alternatively, select **Run** → **Fluidics** from the menu bar.
⇒ The Fluidics Station dialog box appears (Figure 7.7).

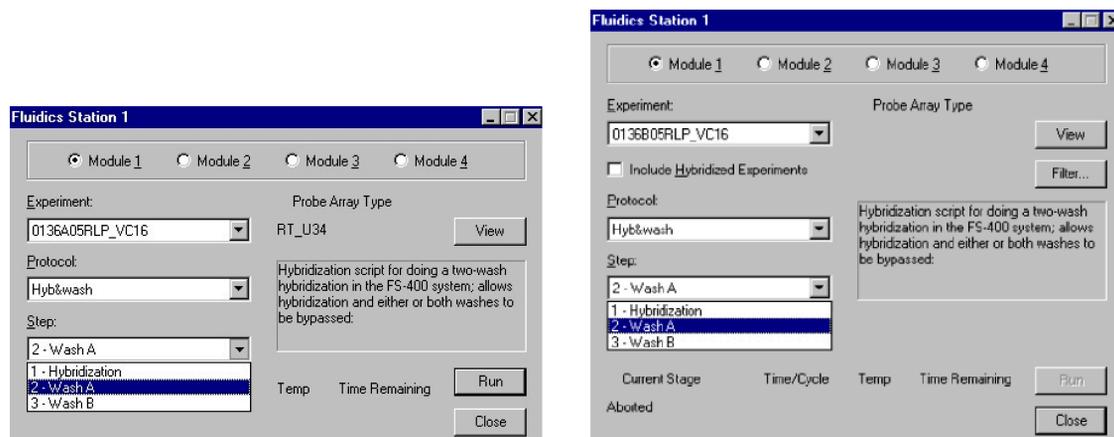


Figure 7.7 Fluidics Station dialog box, bypassing protocol step 1, disk files mode (left) and LIMS mode (right)

2. Select the desired experiment, module, and protocol from the drop-down lists in the Fluidics Station dialog box.

Note

*The bypass function is only available for fluidics protocols that display multiple steps in the **Step** drop-down list of the Fluidics Station dialog box.*

3. Select the desired beginning step from the **Step** drop-down list (Figure 7.7).
4. Click **Run** to start the fluidics protocol at the selected step.

Editing a Fluidics Protocol

You can edit some hybridization and wash (Hybwash) protocols.

✓ Note

A Hybwash protocol must be modified before it is run. Protocol changes made during a run do not affect the run in progress.

1. Select **Tools** → **Edit Protocol** from the menu bar.
⇒ The Fluidics Protocol dialog box appears ([Figure 7.8](#)).

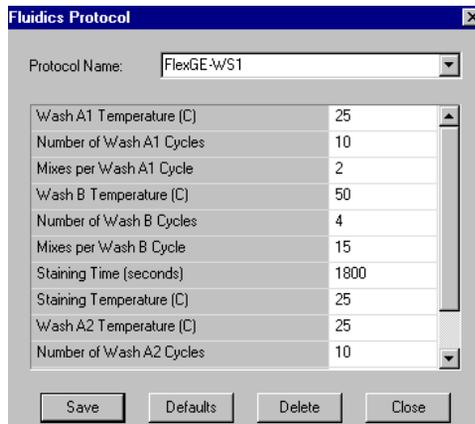


Figure 7.8
Fluidics Protocol dialog box

2. Choose the fluidics protocol you want to edit from the **Protocol Name** drop-down list.

✓ Note

Only the protocols in this list may be edited. All others are defined for specific applications and cannot be customized.

3. Highlight the parameter value you want to change and enter the new value. (Parameters values must be within the ranges in [Table 7.1](#).) Enter a Hybridization Time of zero if only a wash is desired. To omit Wash A or B, enter zero for the Number of Wash A or Wash B cycles.

Table 7.1

Valid ranges for hybridization or stain protocol parameters

Parameter	Valid Range
Hybridization or stain time	0 - 86,399 seconds
Hybridization temperature	15 - 50° C
Wash A temperature	15 - 50° C
Number of Wash A cycles	0 - 99
Mixes per Wash A cycle	1 - 99
Wash B temperature	15 - 50° C
Number of Wash B cycles	0 - 99
Mixes per Wash B cycle	1 - 99
Holding temperature	15 - 50° C

- 4.** To save the parameters under the same protocol name (overwrites the old protocol), click **Save**.
- 5.** To save the parameters under a new protocol name, enter a new name in the **Protocol Name** field, then click **Save**.
⇒ This adds the new protocol name to the drop-down list.
- 6.** Click **Defaults** to return the parameter settings to the default values.
- 7.** Click **Delete** to delete the currently selected protocol from the system.

The Scanner

Microarray Suite controls the Agilent GeneArray® Scanner and enables you to initiate a scan of a GeneChip® probe array, view the intensity data as they are collected during scanning, and analyze the data after the scan is completed. Refer to the *Agilent GeneArray Scanner User's Guide* for a description of the instrument, its components, and set up.

! CAUTION

Always turn on the scanner power switch first before starting the computer workstation and Microarray Suite. Allow the scanner laser 15 minutes of warm up time before scanning a probe array.

Scanning a Probe Array

1. Click the **Scanner** button  in the Instrument Control shortcut bar or the main toolbar. Alternatively, select **Run** → **Scanner** from the menu bar.
⇒ The Scanner dialog box appears ([Figure 7.9](#)).

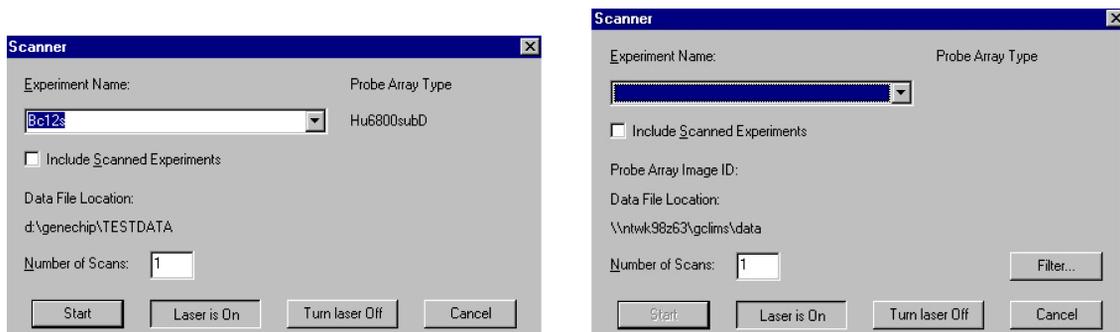


Figure 7.9
Scanner dialog box, disk files mode (left) and LIMS mode (right)

The **Experiment Name** drop-down list displays the experiments (*.exp) in the current directory that have not been scanned (no *.dat file containing scan data exists for the experiments).

 **Note**

*If filters have been applied in LIMS mode, the drop-down list may not include all experiments. Click **Filter** to open the **Filters** dialog box and view or specify filters.*

After a probe array has been scanned, Microarray Suite saves the scan image to an image data file (*.dat).

2. To include scanned experiments in the **Experiment Name** drop-down list, choose the **Include Scanned Experiments** option.
3. Select the experiment name of the probe array to be scanned from the **Experiment Name** drop-down list.

The **Probe Array Type** field automatically displays the probe array type that was entered during experiment setup. Microarray Suite sets the **Number of Scans** per probe array type. You may change the number of scans in the **Scanner** dialog box (Figure 7.9). Multiple scans are combined to create a single image data file (*.dat).

 **Note**

Increasing the number of scans increases the scan time as well as the amount of fluorophore bleaching and may result in lower fluorescence intensities.

4. If it is necessary to rescan a probe array, select the experiment name from the **Experiment Name** drop-down list.

 **IMPORTANT**

*The existing *.dat file will be overwritten when the probe array is rescanned. To retain the original scan data, use the **File** → **Save As** commands from the menu bar to save the original *.dat file under a different name.*

*In disk files mode, the software automatically creates an associated *.exp file with the new name. In LIMS mode, the software automatically adds experiment information for the renamed experiment to the process database.*

5. Click **Start** in the **Scanner** dialog box.
⇒ The GeneChip Scanner dialog box appears (Figure 7.10).



Figure 7.10
Scanner dialog box

6. Load the probe array in the scanner and close the scanner door.
7. To view the scanner options, click **Options** in the GeneChip Scanner dialog box (Figure 7.10).
⇒ The Scanner Options dialog box appears (Figure 7.11).

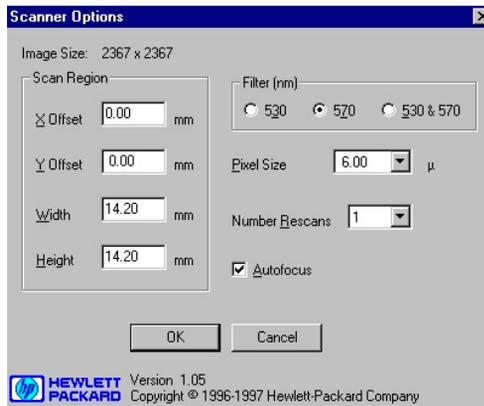


Figure 7.11
Scanner options dialog box

✓ Note

Do not modify the scanner settings. Microarray Suite automatically selects the appropriate settings based on the probe array type specified during experiment setup.

8. Click **OK** to close the Scanner Options dialog box.
9. Click **OK** in the GeneChip Scanner dialog box to start scanning the probe array.

If the view scan in progress feature is enabled (select **View** → **Scan in Progress** from the menu bar), the Image window automatically opens in the main display area when a scan starts. It displays the fluorescence intensity of the probe array line by line as the scan progresses.

To enable (or disable) this option, select **View** from the menu bar and place (or remove) a check mark next to **Scan in Progress**.

After the scan is completed, Microarray Suite:

- saves the image data to an image data file (*.dat) (displayed in the main display area)
- aligns a grid on the *.dat to identify the probe cells
- automatically computes probe cell intensities and saves the data to the cell intensity file (*.cel)

Stopping a Scan

1. Click the **STOP** toolbar button  or select **Run** → **Stop Scanner** from the menu bar.
2. At the prompt, click **Yes** to stop the scanner or **No** to resume scanning (Figure 7.12).

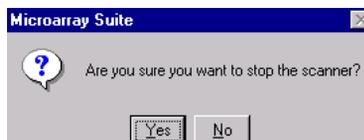


Figure 7.12
Stop scanner prompt

3. After the scan is stopped, the data from a partial scan may be saved. At the prompt (Figure 7.13), click **Yes** to save the data to a *.dat file.

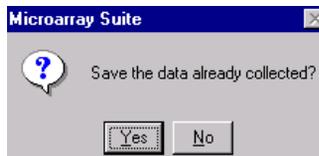


Figure 7.13
Save data prompt

 **Note**

Selecting the same experiment name and repeating a scan of the probe array overwrites the *.dat file of partial scan data that was saved. To retain the original scan data, use the **File** → **Save As** command from the menu bar to save the *.dat file under a different name prior to starting another scan of the probe array.

Shutting Down the Scanner

1. Click the **Scanner** toolbar button  or select **Run** → **Scanner** from the menu bar.
⇒ The Scanner dialog box appears (Figure 7.14).
2. Click **Turn laser Off** in the Scanner dialog box.

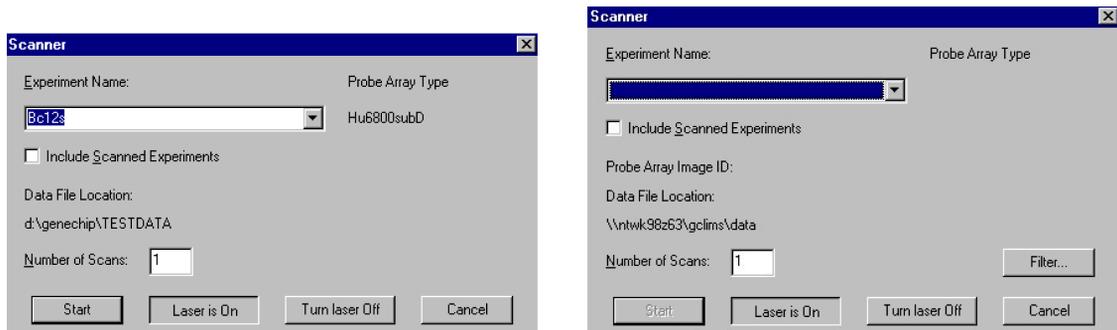
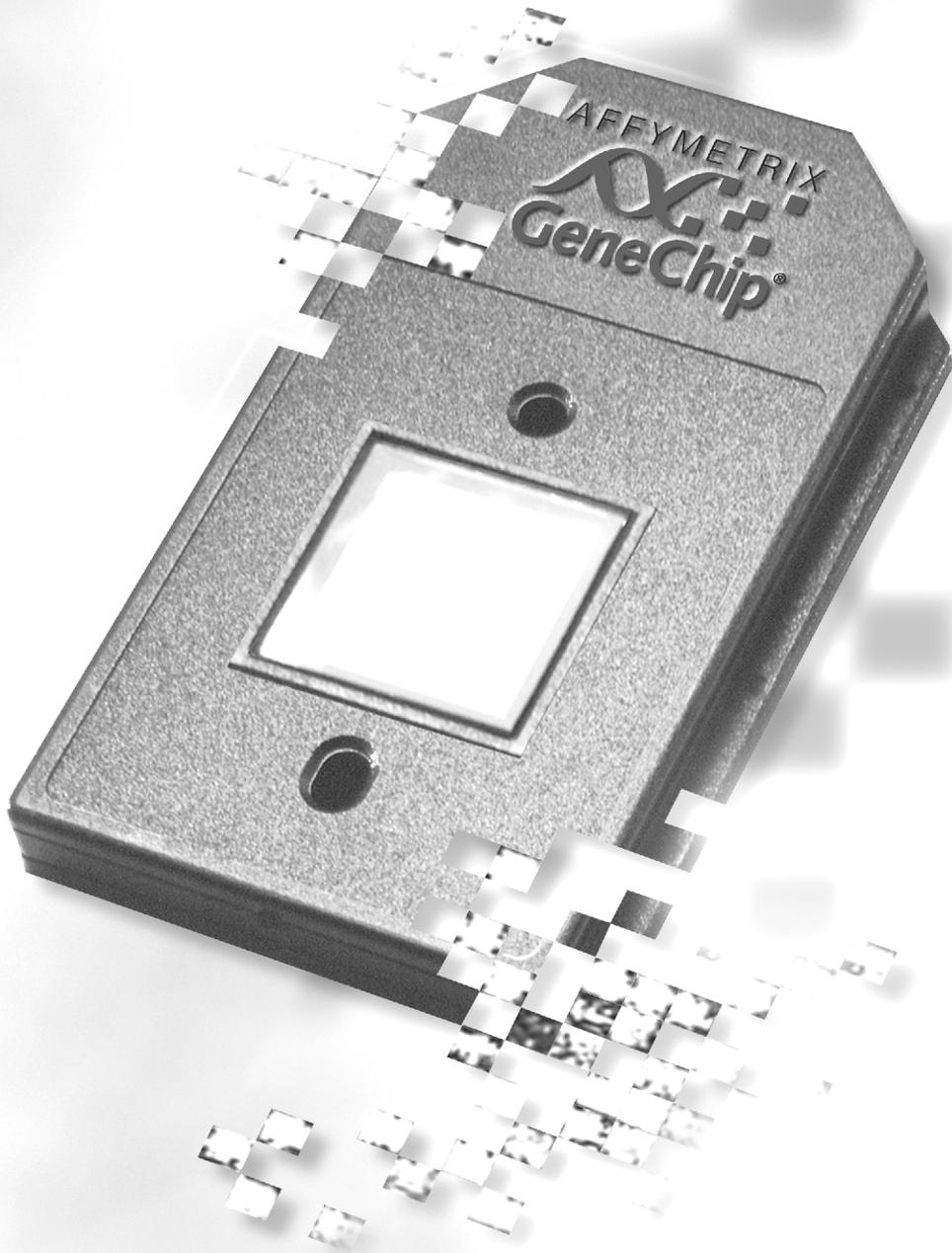


Figure 7.14
Scanner dialog box, disk files mode (left) and LIMS mode (right)

3. At the stop scanner prompt, click **Yes** to turn off the laser power immediately and start the laser cool down routine.

The laser cool down routine lasts for several minutes and is finished when the cooling fan in the laser power supply shuts down. After the cooling fan shuts down, you can safely turn off the scanner power switch.



Chapter 8





Working with Images

This chapter describes how to view an image data file (.dat) or cell intensity file (*.cel) in the Image window. In the Image window, you can:*

- adjust image settings for the *.dat or *.cel file
- view data for a user specified probe cell
- highlight a user specified tile(s) on the probe array
- create bookmarks in the *.dat or *.cel file
- view and mask probe cell outliers

The Image Window

The Image window ([Figure 8.1](#)) displays scan data saved to an image data file (*.dat) or the computed cell intensities saved to a cell analysis file (*.cel).

The Image window opens when:

- a probe array is scanned and the **Scan in Progress** feature is enabled (see the section Scanning a GeneChip® Probe Array in [Chapter 7](#)).
- a *.dat or *.cel is opened (double-click the file in the data tree or click the **Open** toolbar button . Alternatively, select **File** → **Open** from the menu bar to view the *.dat or *.cel files in the current directory).

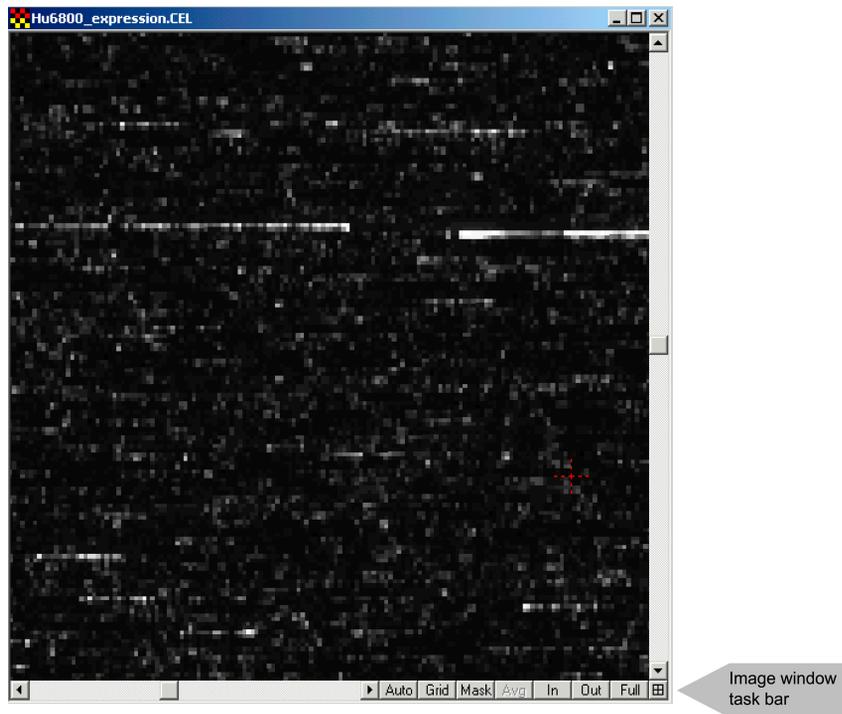


Figure 8.1
Image data file (*.dat) in the Image window

The Image window displays the *.dat or *.cel file name (same as the experiment name entered during experiment setup) in the upper left corner.

- Use the scroll bars at the bottom and right side of the Image window to navigate the scan image.

Commands in the Image Window Task Bar

Auto	Automatically scales the intensity of the current image based on the range of intensity levels in a selected region of the scan image. (For more information see, Adjusting the Image Settings, on page 121.)
Grid	Toggles the grid on and off the scan image. (Press the G key.) (For more information see, Viewing the Grid & Manually Adjusting Grid Alignment, on page 118.)
Mask	Masks user selected areas of the scan image. Masked probe cells are excluded from the analysis. (For more information see, Outliers, on page 141.)
Avg	Calculates the average intensity and standard deviation for a user selected region of the scan image. (For more information see, Calculating Average Intensity for User Specified Probe Cells, on page 121.)
In	Incrementally zooms in on the scan image. Press the I key. Alternatively, position the mouse arrow over the scan image, select an image area by dragging the mouse, and then click In to zoom directly to the selected area.
Out	Zooms out incrementally from the scan image (press the O key).
Full	Zooms out completely to the full scan image view (press the Shift+O keys).
	Enables the user to select and zoom directly to each corner of the scan image to check grid alignment. (For more information see, Viewing the Grid & Manually Adjusting Grid Alignment, on page 118.)

Opening Multiple Image Windows

You can open and view more than one Image window in the main display area.

1. In the data file tree, double-click each data (*.dat) or cell intensity (*.cel) file you wish to view.

This displays each file in the main display area.

2. Select **Window** → **Cascade** from the menu bar to display an overlapping cascade of the open Image windows. Select **Window** → **Tile** to display the open Image windows side by side (Figure 8.2).

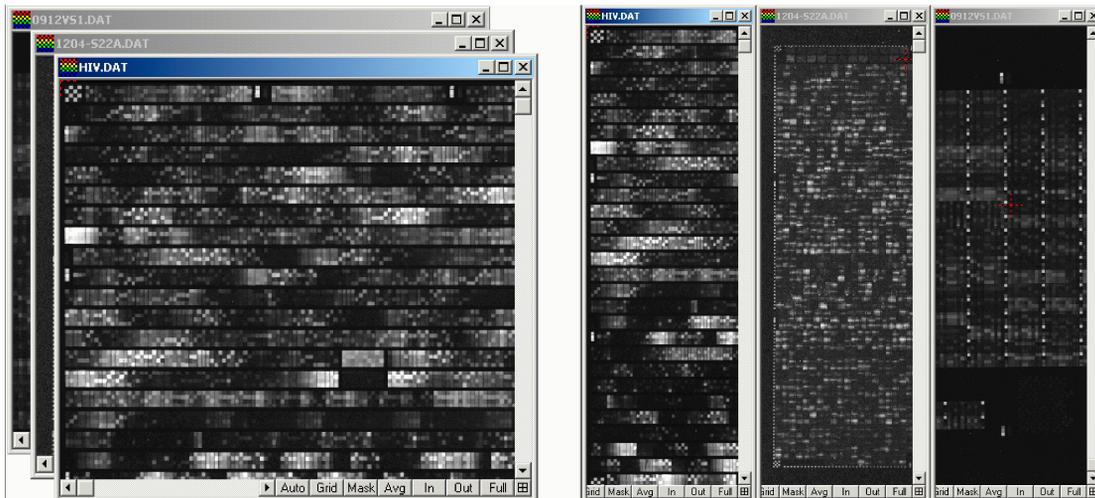


Figure 8.2
Multiple image windows, cascade view (left) and tile view (right)

3. If other file types are open (for example, *.exp or *.chp) and displayed in the main display area, click **Image Views**  in the Microarray Suite shortcut bar to display the open Image windows.

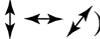
Viewing the Grid & Manually Adjusting Grid Alignment

The Alignment algorithm uses the checkerboard image of the control probes located at the corners of the probe array to superimpose a grid on the scan image. The algorithm aligns the grid so that it delineates the probe cells.

1. To toggle the grid on or off, do one of the following:
 - click **Grid** in the Image window task bar (Figure 8.1)
 - press the **G** key
 - select **View** → **Grid** from the menu bar
2. To confirm the grid is properly aligned, do one of the following to check each corner of the scan image close up:

- click the  button in the Image window taskbar and choose the desired corner
- press **F5** (upper left corner), **F6** (upper right corner), **F7** (lower left corner), or **F8** (lower right corner)
- select **View** → **Corner** from the menu bar and choose the desired corner

On a rare occasion the grid position may be misaligned and require manual adjustment ([Figure 8.3](#)).

1. To manually adjust the grid alignment, place the mouse arrow over the grid perimeter (the arrow becomes a double arrow, .

The diagonal orientation of the double arrow along the perimeter of each corner probe cell indicates horizontal and vertical adjustments can be made simultaneously using the click-and-drag method or by using the keyboard arrow keys.

2. Use the click-and-drag method or the keyboard arrow keys to adjust the horizontal or vertical position of the grid ([Figure 8.3](#)).

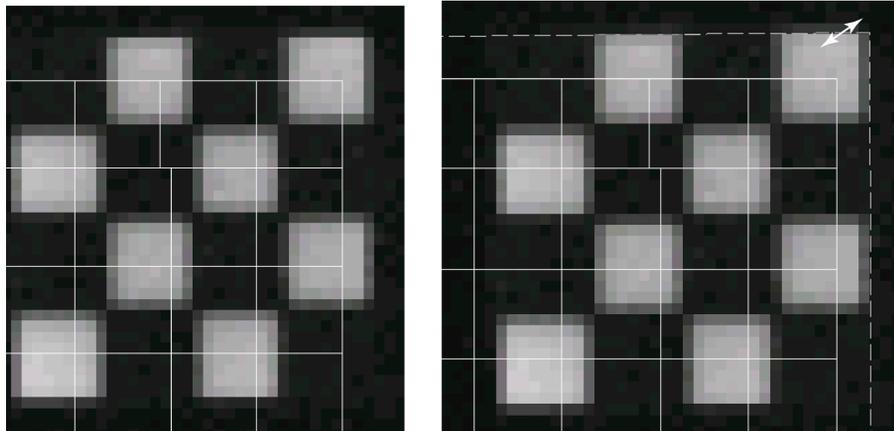


Figure 8.3
Misaligned grid (left) and adjusting the grid alignment (right)

Computing Cell Intensities

Microarray Suite automatically generates a cell intensity file (*.cel) from the image data file (*.dat) (Figure 8.4). The cell analysis algorithm analyzes the *.dat and computes a single intensity value for each probe cell on an array. The intensity data are saved to the *.cel.

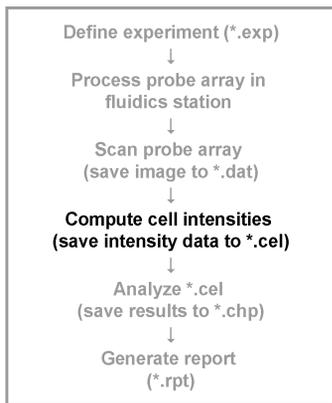


Figure 8.4
Assay & analysis flow chart

Reanalyzing an Image Data File

Microarray Suite automatically generates the cell intensity file (*.cel) from the image data file (*.dat). In some cases you may wish to generate a new cell intensity file (for example, if the grid has been realigned).

To reanalyze a *.dat:

1. Open the *.dat.
2. Select **Run** → **Analysis** from the menu bar.
⇒ The Cell Analysis algorithm recomputes the cell intensities and generates a new *.cel.



Note

*Microarray Suite will not reanalyze the *.dat if the *.cel exists and the grid position has not changed.*

Viewing Images (*.dat)

Calculating Average Intensity for User Specified Probe Cells

Microarray Suite can calculate the average intensity for a user-specified group of probe cells in an image data file (*.dat).

1. In the *.dat, use the click-and-drag method to select the probe cells of interest.
2. Click **AVG** in the Image window task bar (Figure 8.1).
⇒ The Microarray Suite status bar displays the average intensity and the standard deviation.
3. Position the mouse arrow inside the selected area of the *.dat.
⇒ A pop-up tool tip displays the average intensity and the standard deviation for the selected group of probe cells (Figure 8.5).

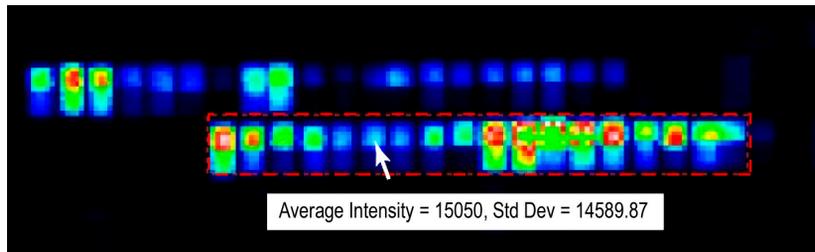


Figure 8.5

Average intensity and standard deviation for user specified probe cells in an image data file (*.dat)

Adjusting the Image Settings

You can adjust image settings (intensity range, color, coordinates, and highlights) for an image data (*.dat) or cell intensity (*.cel) file. The image settings may be adjusted for the current file only or globally for all subsequently opened *.dat or *.cel files.

Making Adjustments to the Current Image Only

- To adjust the settings for the current image only, do one of the following:
 - Click the **Image settings** toolbar button .
 - Press the **S** key.
 - Select **Edit** → **Image Settings** from the menu bar.
- ⇒ This displays the **Image Settings** dialog box for the active *.dat or *.cel (Figure 8.6).



Note

*All settings (except the color setting) in the Image Settings dialog box apply only to the current *.dat or *.cel file and do not affect other images. The settings (except for the color setting) are not saved when the image is closed.*

Click the **S** key.

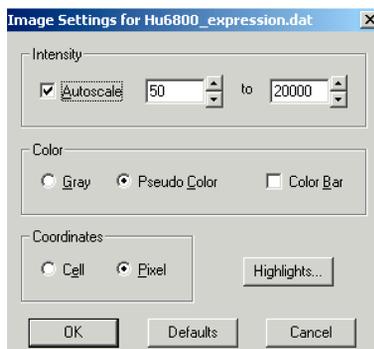
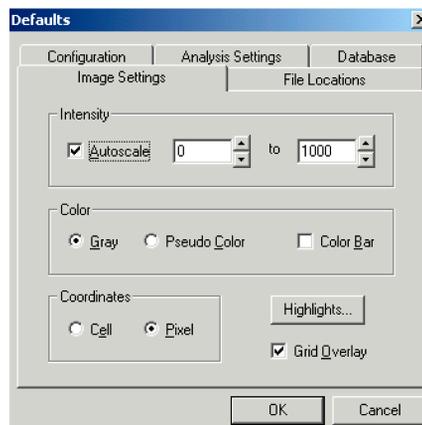


Image settings for the active image only

Select **Tools** → **Defaults** from the menu bar.



Default image settings applied globally to subsequently opened images

Figure 8.6

Image Settings dialog box for the active *.dat or *.cel (left); Defaults dialog box, Image Settings tab (right)

Making Global Adjustments to Subsequently Opened Images

- To adjust settings globally, select **Tool** → **Defaults** from the menu bar.
 - ⇒ This displays the Image Settings tab of the Defaults dialog box (Figure 8.6).

✓ Note

*The settings specified in the Image Settings tab of the Defaults dialog box do not affect an open *.dat or *.cel, but are applied to all subsequently opened *.dat or *.cel files. These default image settings are saved when the image is closed.*

Setting the Intensity Range

The Agilent GeneArray® Scanner software has a dynamic intensity range from 0 to 65,000 and provides 256 colors from black through white for image display. The intensity range of an image is divided into 256 bins and each bin is assigned a color or grayscale.

You may enter a new lower or upper limit for the intensity range associated with an image and apply this subset of the dynamic range to the image. Lowering the upper limit increases the image brightness and raising the lower limit decreases the brightness. This process is called *image scaling* and can be used to adjust the display of a *.dat or *.cel for optimum viewing. Alternatively, Microarray Suite can automatically scale the image using the minimum and maximum pixel intensities of the image.

1. To autoscale the current image, choose the **Autoscale** option in the Image Settings dialog box (Figure 8.6)
2. Click **OK** to automatically scale the image.

When the **Image Settings** dialog box is reopened, it displays the minimum and maximum pixel intensities used to scale the image.

3. To return the intensity settings (autoscale option and intensity range) to the defaults in the Defaults dialog box, click **Defaults** in the Image Settings dialog box.

Selecting Color Settings

- To display the *.dat or *.cel in:
 - **Gray scale** - choose the **Gray** color option (Figure 8.6).
 - **Pseudo color** - choose the **Pseudo Color** option.
 - ⇒ This applies rainbow colors to the intensity scale.
- To display a gray scale or pseudo color bar at the top of the Image window, choose the **Color Bar** option.

Selecting Display Coordinates

- Click the image to display image coordinates.
 - ⇒ A red cross hairs  marks the location.

The status bar (lower left in the main window) displays the intensity data and x,y coordinates of a probe cell or pixel indicated by the cross hairs.

If you choose the **Cell** option in the Image Settings dialog box (Figure 8.6), the status bar and a pop-up tool tip display the x,y coordinates of the cell, the cell intensity, and the pixel intensity (Figure 8.7).

If you choose the **Pixel** option, the pop-up tool tip and the status bar display the x,y coordinates of the pixel and the pixel intensity (Figure 8.7).



Only the Cell option is available for cell intensity files (.cel).*

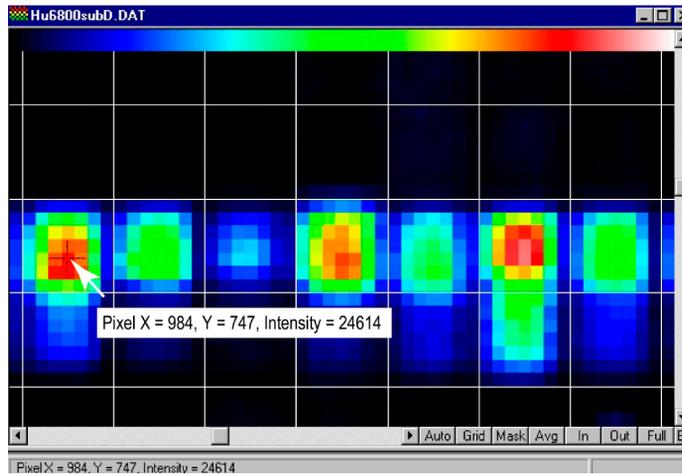
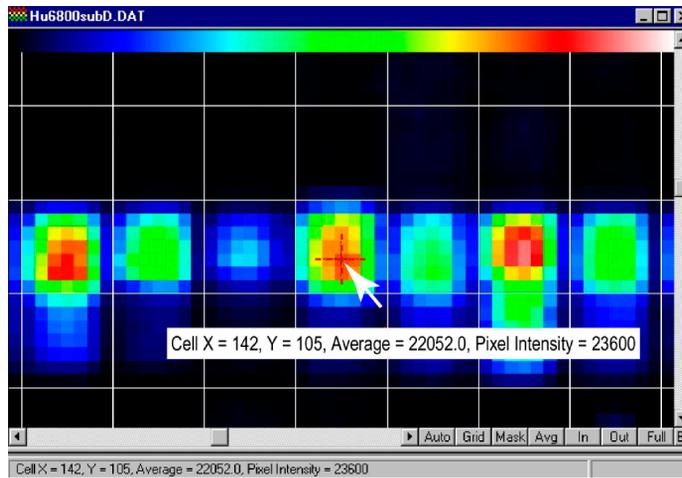


Figure 8.7

Image data file (*.dat), probe cell coordinates (top) and pixel coordinates (bottom)

Specifying Highlight Colors

You may specify new highlighting colors for the grid, masked cells, outlier cells, or highlighted probe array cells (features).

1. Click **Highlights** in the Image Settings dialog box (Figure 8.6).
⇒ The **Highlight Colors** dialog box appears (Figure 8.8).

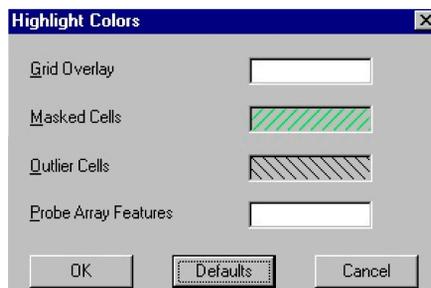


Figure 8.8
Highlight Colors dialog box

2. To change the color of a particular item (for example, the grid overlay):
 - a. Click the associated color box in the **Highlight Colors** dialog box.
⇒ This displays the Color palette (Figure 8.9).

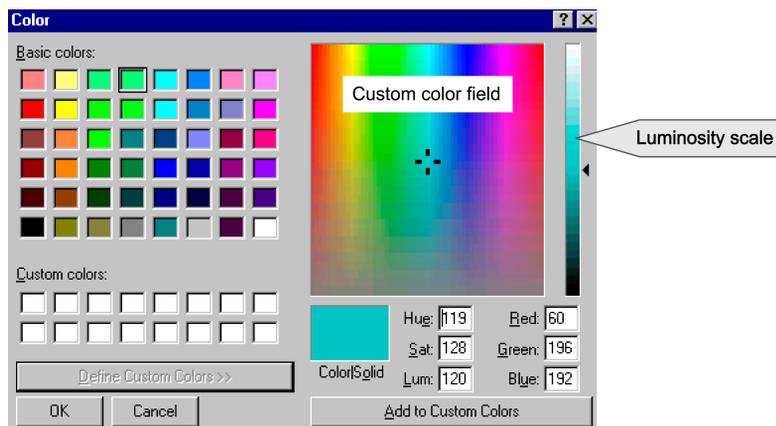


Figure 8.9
Color palette

- b. To select a predefined color, click one of the basic colors.
- c. To define a custom color, click **Define Custom Colors**, then use the click-and-drag method to move the cross hairs in the custom color field. Adjust the color brightness using the luminosity scale to the right. When finished, click **Add to Custom Colors** to apply the color.
⇒ This applies the new color to the selected item.
- d. Click **OK** to close the Color palette.

Image Display Options

Intensity Display

You can view the probe cell intensity data in the image data file (*.dat) as a measured, average, or difference image.

1. Open a *.dat file.
2. Select **View** → **Image** from the menu bar, then choose one of the three viewing options: **Measured**, **Difference**, or **Average**.

Measured Image View

The measured image view ([Figure 8.10](#)) displays the raw data from a scan (*.dat).



Figure 8.10
Measured image view

Cell Image View

The cell image view ([Figure 8.11](#)) displays the probe cell intensity values calculated by the cell analysis algorithm (*.cel).



Figure 8.11
Cell image view

Difference Image View

The difference image view ([Figure 8.12](#)) displays the absolute difference between the measured and the cell intensity values for each pixel. The difference image of a probe cell with hybridized target typically displays high intensity at the perimeter where there is a large difference between the measured and the cell intensity value, but displays low intensity at the center due to small differences between the measured and the cell intensity values.

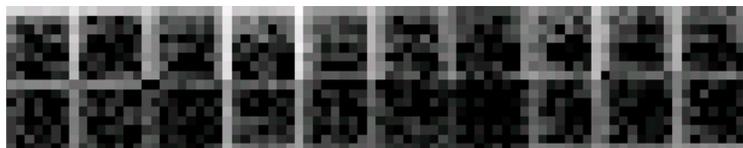


Figure 8.12
Difference image view

The difference image of a hybridized probe array is the reverse of the measured image ([Figure 8.10](#)) where the hybridized probe cells appear bright at the center and dark at the perimeter. The difference image helps assess the uniformity of the hybridization signal in the probe cell and whether the gridlines properly delineate the probe cells.

Viewing Probe Cell Data

You can click a probe cell (feature) in an image data file (*.dat) or cell intensity file (*.cel) to obtain more information about the feature and the associated gene.

Figure 8.13 through **Figure 8.17** display example probe cell data for different types of GeneChip® probe array assays.

To display probe cell data, do one of the following:

- Double-click the probe cell of interest.
- Click the probe cell of interest; right-click the image, and click **View Probe Cell Data** in the shortcut menu.
- Click a probe cell of interest and select **View** → **Probe Cell Data** from the menu bar.

Click other probe cells in the image to update the displayed probe cell data.

Probe Cell Data: Gene Expression Assay

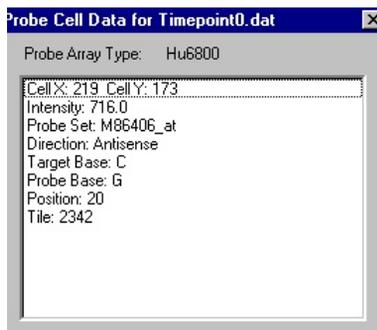


Figure 8.13

Probe cell data, GeneChip® Hu6800 probe array

Cell Coordinates	The probe cell x and y coordinates in cell units.
Intensity	The cell intensity calculated by the cell analysis algorithm.
Probe Set	The identifier of the probe set containing the probe cell.
Direction	The direction (sense or antisense) of the target (sample).
Target Base	The expected nucleotide at the position that corresponds to the substitution position of the probe.
Probe Base	The actual nucleotide located at the substitution position of the probe.
Position	The numbered position of the perfect match and mismatch probe pair within the probe set (for example, 1-20 for a probe set consisting of 20 probe pairs).
Tile	The identification number for the probe set.

Probe Cell Data: Genotyping Assay

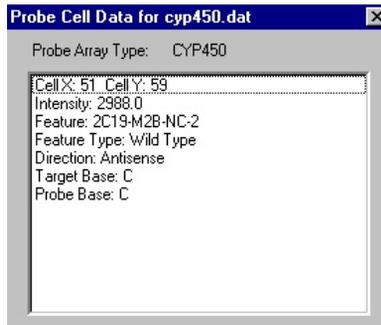


Figure 8.14

Probe cell data, GeneChip® CYP450 probe array

Cell Coordinates	The probe cell x and y coordinates in cell units.
Intensity	The cell intensity calculated by the cell analysis algorithm.
Feature	A polymorphic site.
Feature type	Mutant or wild-type (for CYP450) or allele A or allele B (for HuSNP).
Direction	The direction (sense or antisense) of the target (sample).
Target Base	The expected nucleotide at the position that corresponds to the substitution position of the probe.
Probe Base	The actual nucleotide located at the substitution position of the probe.

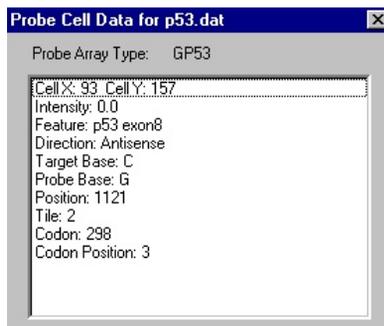
Probe Cell Data: Polymorphism & Mutation Detection Assay

Figure 8.15
Probe cell data, GeneChip® p53 probe array

Cell Coordinates	The probe cell x and y coordinates in cell units.
Intensity	The cell intensity calculated by the cell analysis algorithm.
Feature	A brief description of the region interrogated by the probe.
Direction	The direction (sense or antisense) of the target (sample).
Target Base	The expected nucleotide at the position that corresponds to the substitution position of the probe.
Probe Base	The actual nucleotide located at the substitution position of the probe.
Position	The nucleotide number in the sequence being examined.
Tile	The identification number for the probe set.
Codon	Corresponds to the amino acid number for the gene being examined.
Codon Position	The nucleotide position within the codon.

Probe Cell Data: Mapping Assay

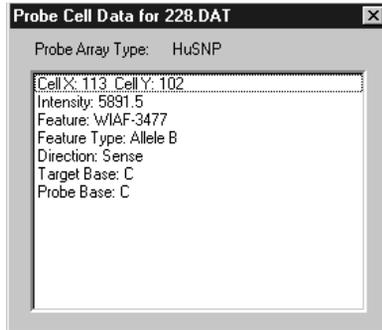


Figure 8.16

Probe cell data, GeneChip® HuSNP™ probe array

Cell Coordinates	The probe cell x and y coordinates in cell units.
Intensity	The cell intensity calculated by the cell analysis algorithm.
Feature	The name of the SNP.
Feature Type	Mutant or wild-type allele A or allele B.
Direction	The direction (sense or antisense) of the target (sample).
Target Base	The expected nucleotide at the position that corresponds to the substitution position of the probe.
Probe Base	The actual nucleotide located at the substitution position of the probe.

Probe Cell Data: Hybridization Assay

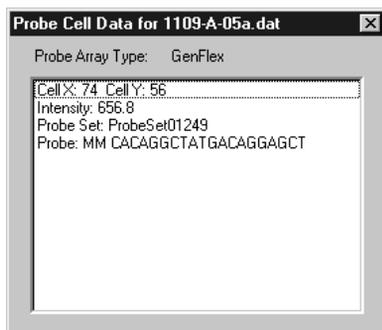


Figure 8.17

Probe cell data, GeneChip® GenFlex™ tag array

Cell Coordinates	The probe cell x and y coordinates in cell units.
Intensity	The cell intensity calculated by the cell analysis algorithm.
Probe Set	The identifier of the probe set containing the probe cell.
Probe	The probe type located in the probe cell (perfect match (PM), mismatch (MM), complement to the perfect match (CPM), or complement to the mismatch (CMM)).

Viewing Probe Set Information

Viewing Probe Array Tiling Information

You can view probe array tiling information for the current image data (*.dat) or cell intensity (*.cel) file in the Image window.

1. Right-click the *.dat or *.cel and click **View Probe Tiling** in the shortcut menu. Alternatively, select **View** → **Probe Tiling** from the menu bar.

⇒ The **View Probe Tiling** dialog box appears (Figure 8.18).



Figure 8.18
View Probe Tiling dialog box

The Features tab of the View Probe Tiling dialog box lists all probe sets tiled on the probe array of the current *.dat file.

2. In the Features tab, press and hold the **Ctrl** key while you click the features of interest (Figure 8.18). Click **Apply** when the selection is completed.

⇒ This highlights the probe sets in the scan image that correspond to the selected features so that hybridization data may be quickly located (Figure 8.19).

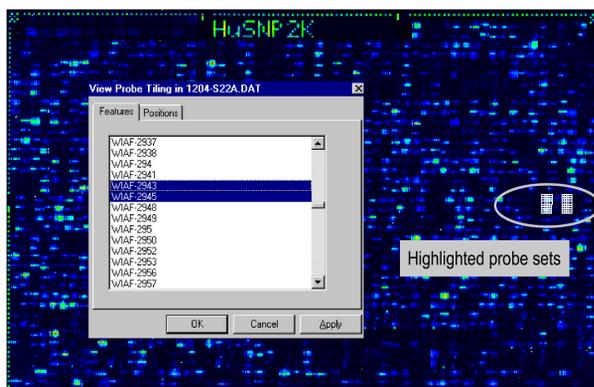


Figure 8.19
Highlighted probe cells in the scan image

TIP

Turn off the grid (press the **G** key) so that the highlighted probe cells can be easily seen. The highlighting color for the selected probe cells may be changed. (For more information see, [Specifying Highlight Colors](#), on page 126.)

3. Click **OK** to close the View Probe Tiling dialog box.

Viewing Probe Cell Data for a Highlighted Probe Set(s)

1. In the *.dat or *.cel, click a highlighted probe cell of interest.
⇒ This places the red cross hairs over the feature.
2. Right-click and select **View Probe Cell Data** from the shortcut menu. Alternatively, select **View** → **Probe Cell Data** from the menu bar.
⇒ This displays the probe cell data ([Figure 8.20](#)).

✓ Note

If you click another probe cell, Microarray Suite automatically updates the probe Cell Data dialog box.

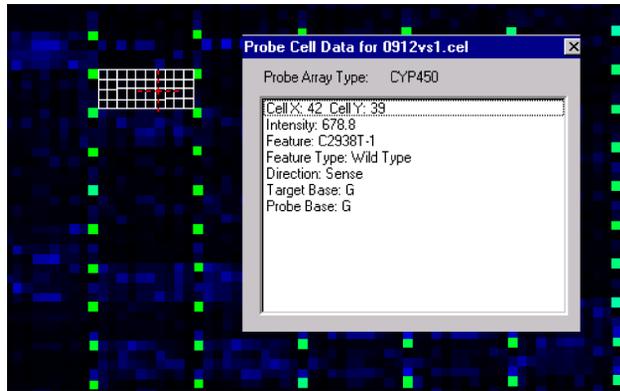


Figure 8.20
Probe cell data for a highlighted feature

3. To place the red cross hairs over the next feature of the highlighted probe set, click the image (to return focus to the image window) and do one of the following:
 - Press the **N** key.
 - Right-click the image and select **View Next Highlight** in the shortcut menu.
 - Select **View** → **Next Highlight** from the menu bar.

If it is open, the Probe Cell Data dialog box is automatically updated with the new probe cell data.

4. To return the cross hairs to a previous feature, do one of the following:
 - Press the **P** key.
 - Right-click the image and select **View Previous Highlight** in the shortcut menu.
 - Select **View** → **Previous Highlight** from the menu bar.
5. To remove all highlights, do one of the following:
 - Press the **E** key.
 - Right-click the image and select **Clear Highlights** from the shortcut menu.
 - Select **View** → **Clear Highlights** from the menu bar.

✓ **Note**

*Highlights are automatically cleared when the *.dat or *.cel file is closed.*

Highlighting Tiles on the Probe Array

You can specify highlighting options for:

- all tiles or a particular tile on the probe array (by tile number)
 - all positions or a particular position of the selected tile(s)
1. Right-click the *.dat or *.cel and click **View Probe Tiling** in the shortcut menu. Alternatively, select **View** → **Probe Tiling** from the menu bar.
 - ⇒ The View Probe Tiling dialog box appears (Figure 8.21).
 2. Click the Positions tab, select a **Position** and **Tile**, then click **Apply**.
 - ⇒ This highlights the selection in the image.

For example, in Figure 8.21 all positions of tile number 2 on the HIV PRT *Plus* probe array are highlighted. In Figure 8.22 only position number 45 of tile number 10 on the CYP450 probe array is highlighted.

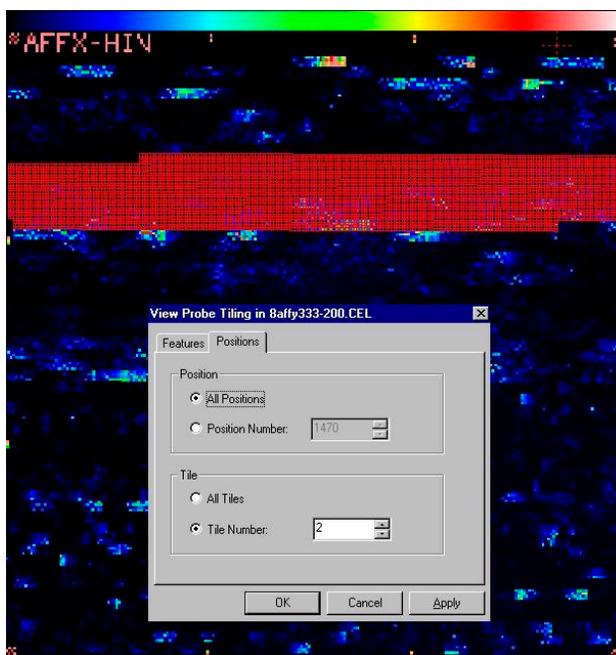


Figure 8.21
HIV PRT *Plus* cell intensity file, all positions of tile 2 highlighted

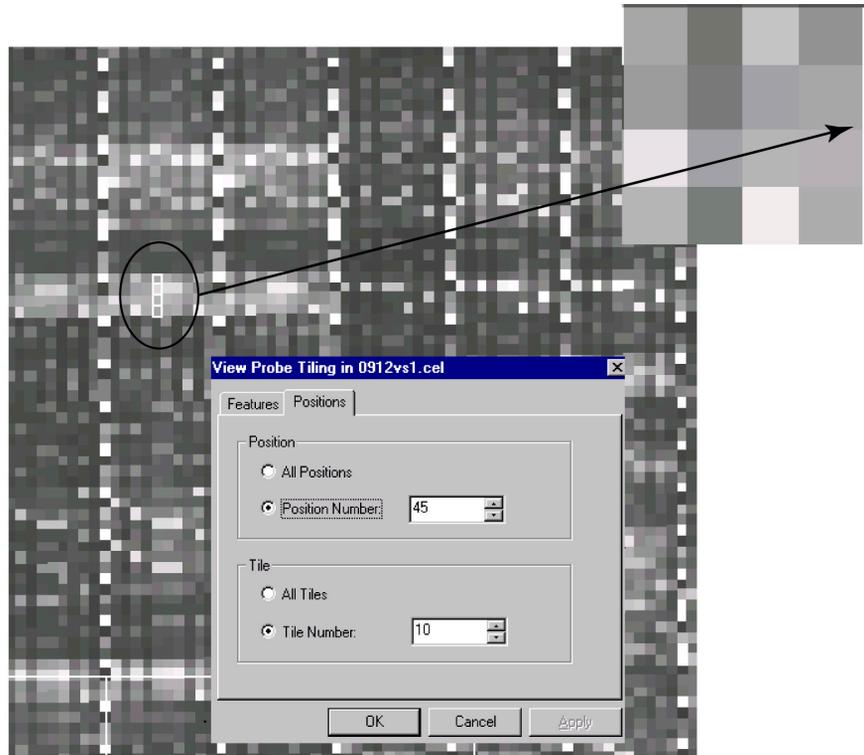


Figure 8.22
CYP450 cell intensity file, position number 45 of tile 10 highlighted

Bookmarks

A bookmark denotes a specific area of the image (*.dat or *.cel) at a particular magnification level. Bookmarks make it easy to quickly view the hybridization data at a specific location in different images.

When you select a bookmark, the Image window automatically displays the image location at the magnification specified by the bookmark.

Creating a Bookmark

1. Use the task bar zoom commands and scroll bars in the Image window to locate the image area of interest at the desired magnification level.
2. Right-click the image and select **View Bookmarks** → **Add Bookmark** from the shortcut menu. Alternatively, select **View** → **Bookmarks** → **Add Bookmark** from the menu bar.
⇒ The Add Bookmark dialog box appears (Figure 8.23).

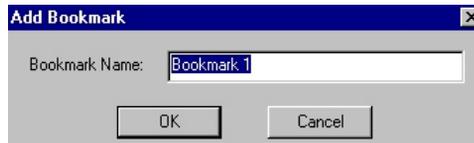


Figure 8.23
Add Bookmark dialog box

3. Enter a name in the **Bookmark Name** box or use the default name **Bookmark 1, 2,3...**
4. Click **OK** to create the bookmark.

Viewing Bookmarks

To view bookmarks, do either of the following:

- Right-click the image and select **View Bookmarks** → **Edit Bookmarks** from the shortcut menu. Alternatively, select **View** → **Bookmarks** → **Edit Bookmarks** from the menu bar.
⇒ The Edit Bookmarks dialog box displays a list of all the bookmarks (Figure 8.24).
- Select **View** → **Bookmarks** from the menu bar. (A maximum of 20 bookmark names are displayed here.)

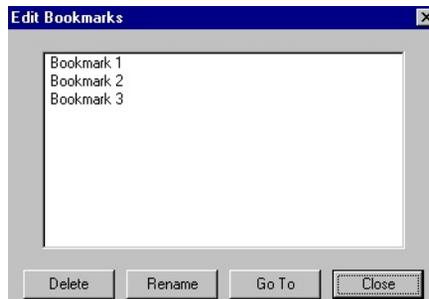


Figure 8.24
Edit Bookmarks dialog box

Commands in the Edit Bookmarks Dialog Box

Highlight the bookmark name in the Edit Bookmarks dialog box, then click the desired command.

- | | |
|---------------|--|
| Delete | Permanently removes the selected bookmark. |
| Rename | Allows the selected bookmark to be renamed. |
| Go To | Displays the selected bookmark in the image. |
| Close | Closes the Edit Bookmarks dialog box. |

Outliers

Outliers are probe cells that are obscured or non-uniform in intensity (for example, probe cells with bright or dark streaks). The Outlier Detection algorithm automatically determines whether a probe cell is an outlier.

You can exclude or *mask* outliers from a probe array in an expression analysis. The expression analysis algorithm ignores the probe cell data from masked outliers.

Viewing Outliers

1. Select **View** → **Cells** → **Outliers** from the menu bar.
⇒ This highlights outlier probe cells with diagonal lines .

✓ Note

The color that highlights outliers may be changed. (For more information see, [Specifying Highlight Colors](#), on page 126.)

Masking Outliers

1. Select **Edit** → **Mask All Outliers** from the menu bar.
⇒ This automatically masks all outliers identified by the Outlier Detection algorithm.
2. To mask user-specified probe cells:
 - a. Use the click-and-drag method to select a probe cell or area of the image to be masked.
 - b. Select **Edit** → **Mask Cells** from the menu bar.

💡 TIP

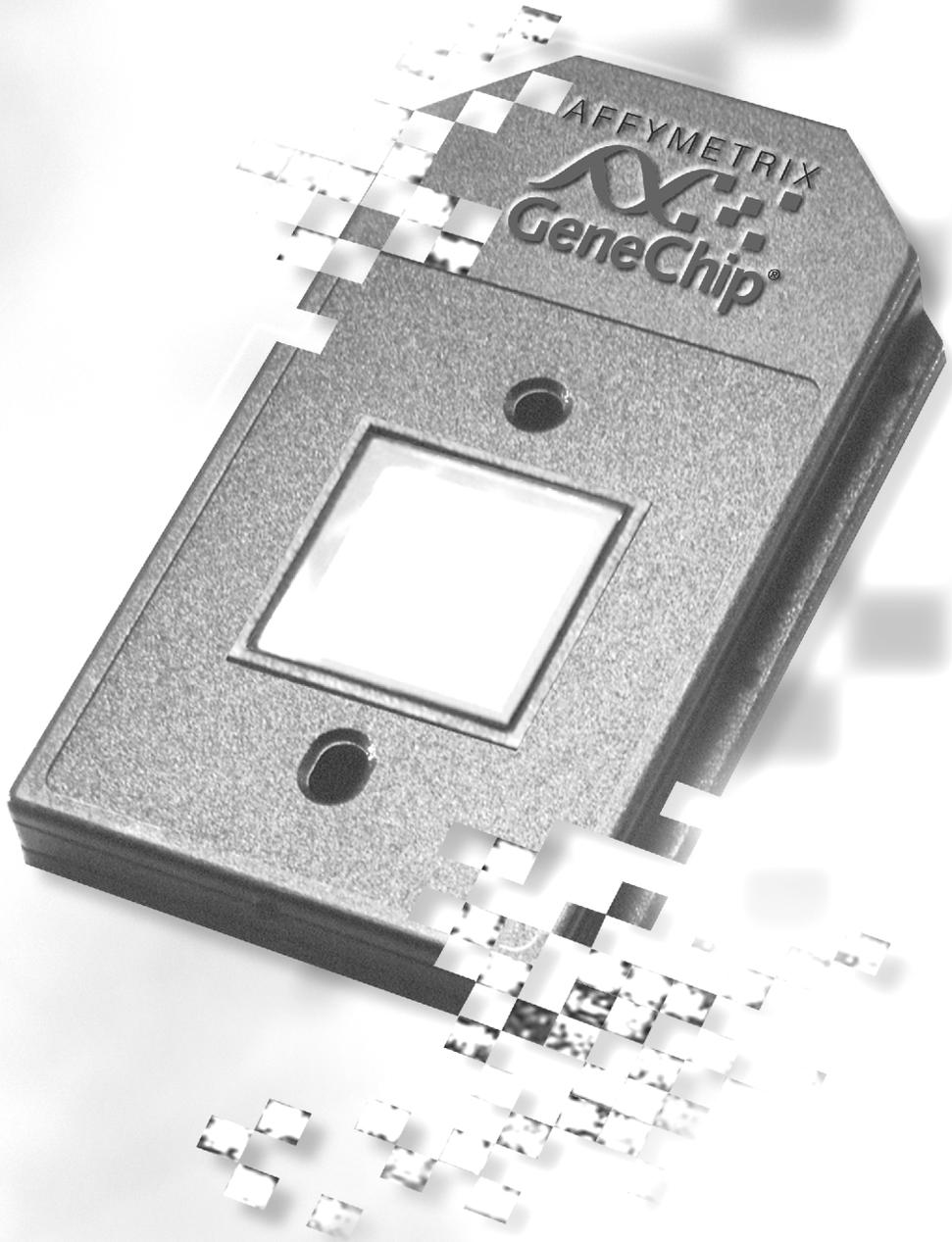
*Turn off the grid (press the **G** key) to easily view the masked probe cells highlighted with diagonal lines.*

Unmasking Outliers

1. Select **Edit** → **Unmask All Cells** from the menu bar.
⇒ This removes all masking.
2. To unmask only a selected probe cell or area:
 - a. Use the click-and-drag method to select a probe cell or area of interest.
 - b. Select **Edit** → **Unmask Cells** from the menu bar.

✓ Note

*If you manually adjust the grid or change the probe array type in the experiment, the mask or outlier data will be lost when a new *.cel file is generated (select **Run** → **Analysis** from the menu bar).*



Chapter 9



Gene Expression Analysis

The Expression Analysis algorithm analyzes cell intensities (.cel) from GeneChip® expression arrays (Figure 9.1).*

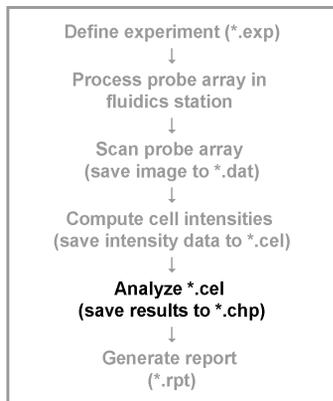


Figure 9.1
Assay & analysis flow chart

This chapter describes how to:

- run an absolute or comparison expression analysis
- view the expression analysis output (*.chp) in graphical or tabular format in the Expression Analysis window (EAW)
- generate a report

Overview of Absolute & Comparison Expression Analysis

Absolute Expression Analysis

An absolute expression analysis examines the cell intensity file (*.cel) from one experiment (GeneChip® probe array). For each transcript represented on the probe array, the expression algorithm computes the:

- detection call (present, absent, or marginal (unable to call the transcript present or absent), or no call)
- detection p-value
- signal (background-subtracted and adjusted for noise)
- stat pairs
- stat pairs used

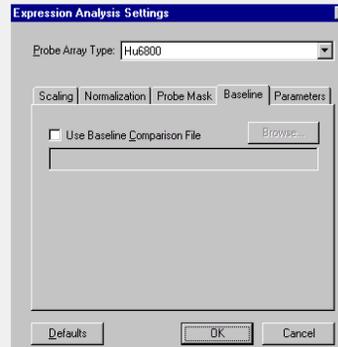
Microarray Suite saves the data to an analysis output or *chip* file (*.chp).

To run an absolute analysis ([Figure 9.2](#)):

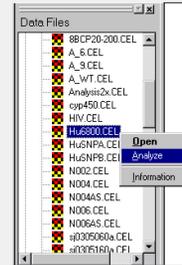
- confirm the defaults or specify new values in the Expression Analysis Settings dialog box
- in the data tree, select the cell intensity file (*.cel) for the absolute analysis
- analyze the *.cel and view the (*.chp) in the Expression Analysis window (EAW)

Absolute Expression Analysis

- 1 Confirm or specify new expression analysis settings. Do not choose **Use Baseline Comparison File** option.



- 2 In the data tree, right-click the *.cel file, then select **Analyze**.



- 3 Expression Analysis window (EAW) displays absolute analysis results.

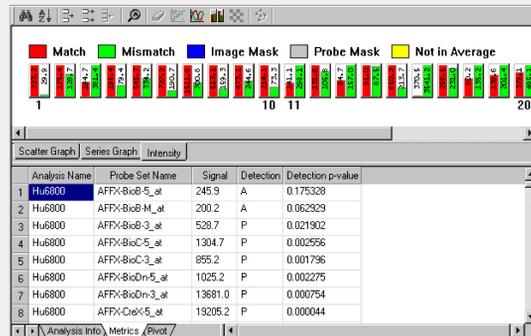


Figure 9.2
Absolute expression analysis

Comparison Expression Analysis of an Experiment & Baseline

A comparison analysis compares the cell intensities of the same probe set on two different probe arrays (an experiment and baseline of the same probe array type) to determine the relative change in the expression level of a transcript.

For each transcript represented on the probe array, the expression algorithm computes the:

- change call (increase, marginal increase, no change, marginal decrease, decrease) or no call
- change detection p-value
- signal log ratio (background-subtracted and adjusted for noise)
- signal log ratio low
- signal log ratio high
- stat common pairs

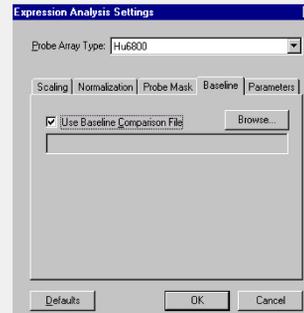
Microarray Suite saves the data to an analysis output or *chip* file (*.chp).

To run a comparison analysis ([Figure 9.3](#)):

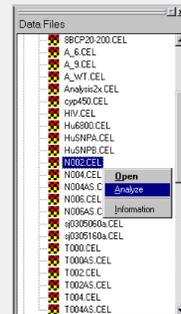
- select a baseline analysis output file (*.chp)
- confirm the default expression analysis settings or specify new values
- select the experiment cell intensity file (*.cel) from the data tree
- analyze the experiment *.cel and view the comparison expression results (*.chp) in the EAW

Comparison Expression Analysis: Compare an Experiment & Baseline

- 1 Confirm or specify new expression analysis settings. Choose the **Use Baseline Comparison File** option. Click **Browse** to specify the baseline *.chp.



- 2 In the data tree, right-click the experiment *.cel, then select **Analyze**.



- 3 Expression Analysis window (EAW) displays absolute & comparison analysis results for the experiment.

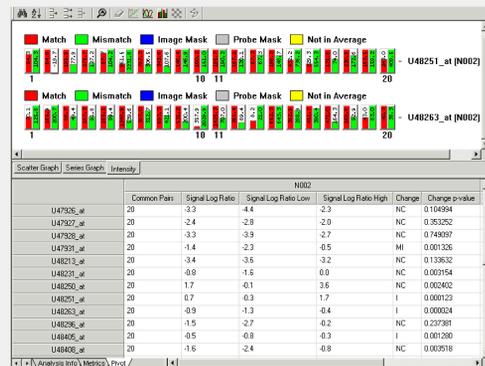


Figure 9.3
Comparison expression analysis of an experiment and baseline

Expression Analysis Settings

The Statistical Expression algorithm relies on the expression analysis settings to derive biologically meaningful results from the hybridization intensity data. The Expression Analysis Settings dialog box ([Figure 9.4](#)) displays the:

- the type of expression probe array for the analysis
- a scaling or normalization factor (see [Appendix D](#))
- probe mask definition (see [Appendix D](#))
- the baseline file selected for a comparison analysis
- the user-modifiable parameters of the expression analysis algorithm (see [Appendix D](#))

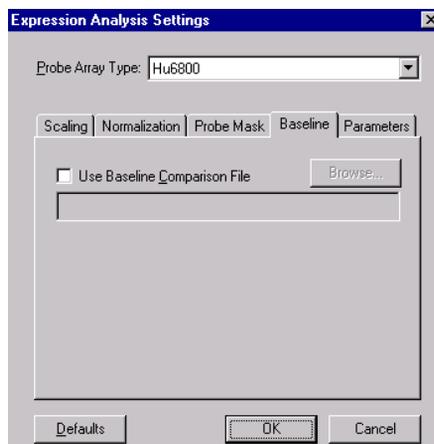


Figure 9.4
Expression Analysis Settings dialog box, Baseline tab

✓ Note

Before running an absolute or comparison analysis, confirm the defaults or specify new values for the expression analysis settings.

1. In the Analysis Settings shortcut bar, click **Expression** .
⇒ The Expression Analysis Settings dialog box appears ([Figure 9.4](#)).

2. Click a tab to view each type of setting. (See Appendix D for more information about the settings and how to change them.)

In LIMS mode, the expression analysis settings are saved on a per user basis. Changes made by one user (identified by the logon name) do not affect the settings of other users.



TIP

To compare the analysis output files (*.chp) of several experiments, scale the experiments to the same target intensity using the **All Probe Sets** or **Selected Probe Sets** scaling option (see Appendix D).

Absolute Expression Analysis

An absolute expression analysis analyzes the cell intensity data (*.cel file) of a GeneChip® expression probe array. For each transcript represented on the probe array, the algorithm computes a detection call, detection p-value, and signal.



Note

In LIMS mode, an expression analysis run from the data tree, menu bar, toolbar, or workflow monitor uses the expression analysis settings in the Expression Analysis Settings dialog box. The user set specified during experiment setup is not applied.

1. Confirm or specify new settings in the Expression Analysis Settings dialog box (Figure 9.4).
2. To analyze a *.cel, do one of the following:
 - in the data tree, right-click the *.cel (or *.dat), then click **Analyze** in the shortcut menu
 - select **File** → **Analyze** from the menu bar and choose the *.cel you want to analyze from the Analyze dialog box that appears.
 - if the *.cel or *.dat is open, click the **Analyze** toolbar button . Alternatively, select **Run** → **Analysis** from the menu bar.

⇒ The Save Results As dialog box appears (if chosen as a default option, see Chapter 16).

It displays the analysis output file (*.chp) default name (same as the experiment name specified during experiment setup) (Figure 9.5).

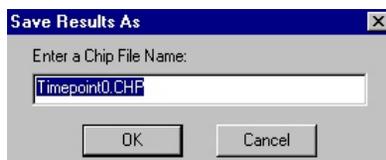


Figure 9.5
Save Results As dialog box

3. If an analysis was previously run and you wish to save the current results without overwriting the previous results, enter a new name in the Save Results As dialog box.
4. Click **OK**.
⇒ The Expression Analysis Settings dialog box appears next (if chosen as a default option, see [Chapter 16](#)) ([Figure 9.6](#)).

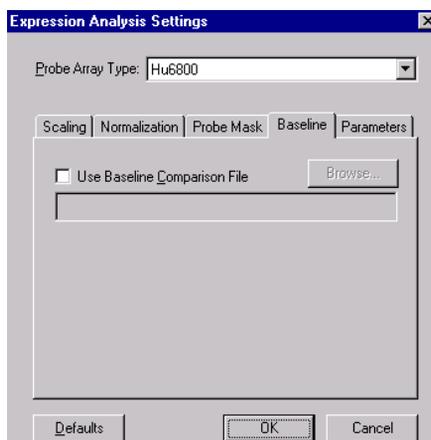


Figure 9.6
Expression Analysis Settings, baseline tab

5. In the Expression Analysis Settings dialog box, click the Baseline tab and make sure the **Use Baseline Comparison File** option is not chosen ([Figure 9.6](#)).
6. Confirm or change the other expression analysis settings as desired (see [Appendix D](#)).

7. Click **OK**.

⇒ This closes the Expression Analysis Settings dialog box and runs the absolute analysis.

When the analysis is finished, the Expression Analysis window (EAW) displays the analysis output file (*.chp) (Figure 9.7).

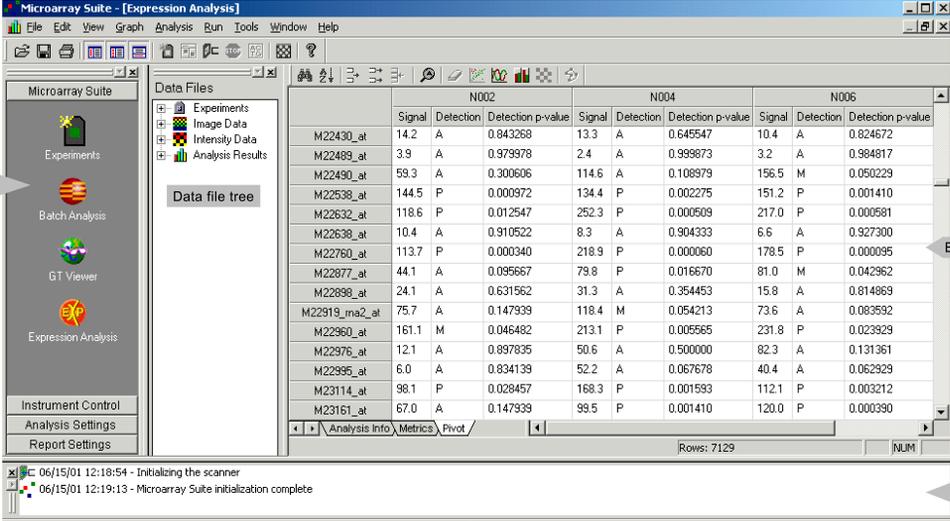
The status log (Figure 9.7) displays:

- the name of the analysis output file (*.chp)
- the location of the *.chp
- a message indicating when the analysis is completed

If the status log is not displayed, click the **Status Log** toolbar button  or select **View** → **Status Bar** from the menu bar.

 **Note**

If the EAW is already open, the results are added to the open window and it may be necessary to use the scrollbars at the bottom and right side of the EAW to see the newly added results.



The screenshot shows the Microarray Suite - [Expression Analysis] window. On the left is a 'Shortcut bar' with icons for Experiments, Batch Analysis, GT Viewer, and Expression Analysis. Below it are 'Instrument Control', 'Analysis Settings', and 'Report Settings'. The main area is a 'Data Files' pane with a 'Data file tree' and a pivot table. The pivot table has columns for Signal, Detection, and Detection p-value, grouped by probe sets N002, N004, and N006. The status log at the bottom shows two entries: '06/15/01 12:18:54 - Initializing the scanner' and '06/15/01 12:19:13 - Microarray Suite initialization complete'.

	N002			N004			N006		
	Signal	Detection	Detection p-value	Signal	Detection	Detection p-value	Signal	Detection	Detection p-value
M22430_at	14.2	A	0.843268	13.3	A	0.645547	10.4	A	0.824672
M22489_at	3.9	A	0.979378	2.4	A	0.999673	3.2	A	0.984617
M22490_at	59.3	A	0.300606	114.6	A	0.108979	156.5	M	0.050229
M22538_at	144.5	P	0.000972	134.4	P	0.002275	151.2	P	0.001410
M22632_at	118.6	P	0.012547	252.3	P	0.000509	217.0	P	0.000581
M22639_at	10.4	A	0.910522	8.3	A	0.904333	6.6	A	0.927300
M22760_at	113.7	P	0.000340	218.9	P	0.000060	178.5	P	0.000095
M22877_at	44.1	A	0.095657	79.8	P	0.016670	81.0	M	0.042962
M22898_at	24.1	A	0.631562	31.3	A	0.354453	15.8	A	0.814669
M22919_ma2_at	75.7	A	0.147939	118.4	M	0.054213	73.6	A	0.083592
M22960_at	161.1	M	0.046482	213.1	P	0.005665	231.8	P	0.023929
M22976_at	12.1	A	0.897835	50.6	A	0.500000	82.3	A	0.131361
M22995_at	6.0	A	0.834139	52.2	A	0.067678	40.4	A	0.062929
M23114_at	98.1	P	0.028457	168.3	P	0.001593	112.1	P	0.003212
M23161_at	67.0	A	0.147939	99.5	P	0.001410	120.0	P	0.000390

Figure 9.7
Expression Analysis Window (EAW), pivot tab, absolute analysis output (*.chp)

Comparison Expression Analysis

A comparison expression analysis compares the cell intensity data (*.cel file) of an experiment to a baseline GeneChip® expression probe array (of the same probe array type). The comparison analysis identifies the relative change in the expression level of each transcript represented on the probe array.

For each transcript represented on the array, the algorithm computes a change call, change p-value, signal log ratio, signal log ratio low, and signal log ratio high.



Note

An absolute analysis of the baseline must be run prior to running a comparison analysis of the experiment and baseline.

Selecting the Baseline

1. In the Analysis Settings shortcut bar, click **Expression** .
⇒ The Expression Analysis Settings dialog box appears ([Figure 9.8](#)).

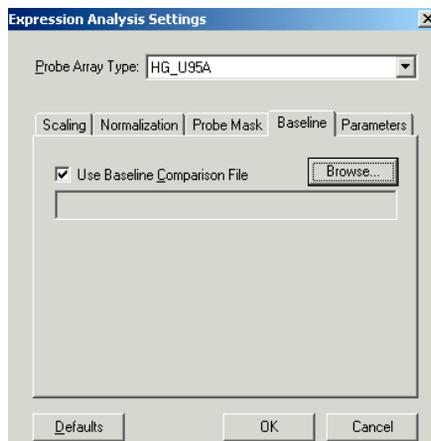


Figure 9.8
Expression Analysis Settings dialog box

2. Select the probe array type for the baseline from the **Probe Array Type** drop-down list (Figure 9.8).
3. Click the Baseline tab.
4. Choose the **Use Baseline Comparison File** option and click **Browse**.
⇒ This displays the Baseline Comparison File dialog box (disk files mode) or the Open dialog box (LIMS mode) (Figure 9.9).

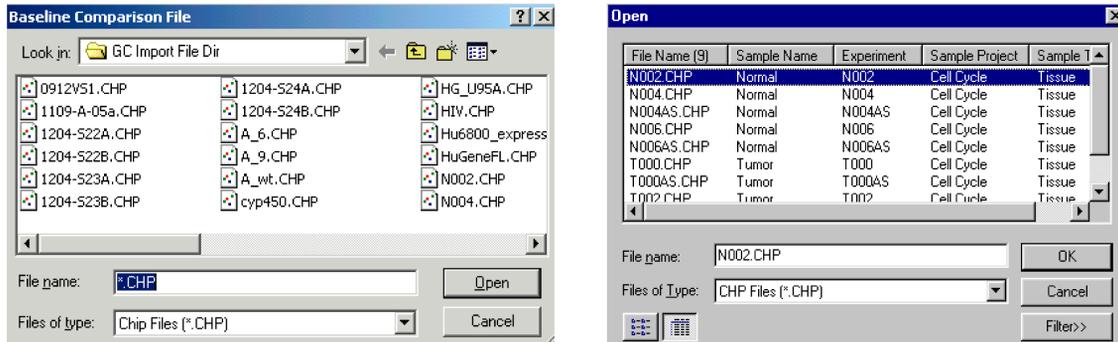


Figure 9.9
Baseline Comparison File dialog box, file mode (left) and Open dialog box, LIMS mode (right)

5. Select the baseline file and click OK.

Running the Comparison Analysis

1. To run the comparison analysis of an experiment and baseline, do one of the following:
 - in the data tree, right-click the experiment *.cel (or *.dat), then click **Analyze** in the shortcut menu
 - select **File** → **Analysis** from the menu bar and choose the experiment *.cel from the Analyze dialog box that appears
 - if the *.cel or *.dat is open, click the **Analyze** toolbar button . Alternatively, select **Run** → **Analysis** from the menu bar.
 ⇒ The Save Results As dialog box appears (if chosen as a default option, see Chapter 16) and displays the analysis output file (*.chp) default name (same as the experiment name specified during experiment setup) (Figure 9.10).



Figure 9.10
Save Results As dialog box

2. Enter a new name in the **Save Results As** dialog box if an analysis was previously run and you wish to save the current results without overwriting the previous results. Click **OK**.
 - ⇒ The **Expression Analysis Settings** dialog box appears next (if chosen as a default option, see [Chapter 16](#)) (**Figure 9.11**).

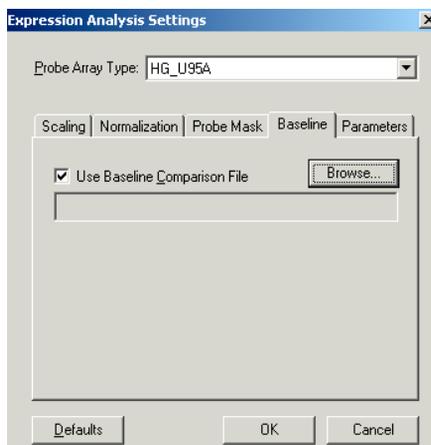


Figure 9.11
Expression Analysis Settings dialog box, baseline tab

3. If the Expression Analysis Settings dialog box appears, review and confirm the settings for the probe array type and baseline *.chp:
 - a. In the Baseline tab, select the probe array type for the baseline from the **Probe Array Type** drop-down list (**Figure 9.11**).

- b. Choose the **Use Baseline Comparison File** option, then click **Browse**
- ⇒ This displays the Baseline Comparison File dialog box (disk files mode) or the Open dialog box (LIMS mode) (**Figure 9.12**).

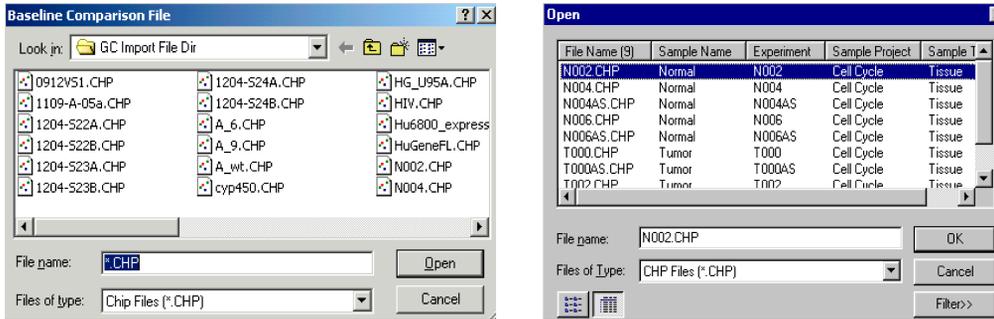


Figure 9.12
Baseline Comparison File dialog box, file mode (left) and Open dialog box, LIMS mode (right)

- c. Double-click the baseline *.chp. Alternatively, click the baseline *.chp, then click **Open** or **OK**.
- ⇒ The Expression Analysis Settings dialog box displays the selected baseline file (**Figure 9.13**).

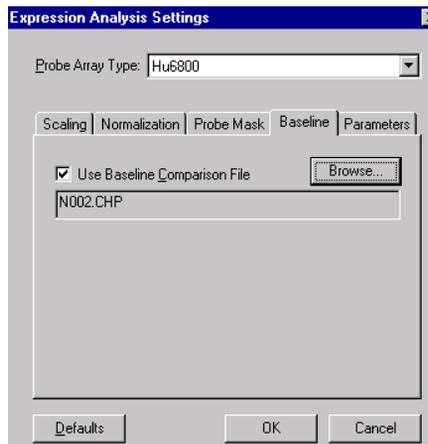


Figure 9.13
Expression Analysis Settings, baseline tab displays the baseline file for the comparison analysis

4. In the Expression Analysis Settings dialog box (**Figure 9.13**), click the normalization tab and specify a normalization option for the experimental data.

✓ Note

You need only specify a normalization option for the experiment, not the baseline. You may use the scaling option instead of the normalization option (see Appendix D). If you use the scaling option, scale both the experiment and the baseline.

5. Confirm or specify other probe array analysis parameters as desired (see Appendix D).
6. Click **OK**.
⇒ This closes the Expression Analysis Settings dialog box and runs the absolute analysis.

When the analysis is finished, the Expression Analysis window (EAW) displays the analysis output file (*.chp) (**Figure 9.14**).

The status log (**Figure 9.14**) displays:

- the name of the analysis output file (*.chp)
- the location of the *.chp
- a message indicating when the analysis is completed

If the status log is not displayed, click the **Status Log** toolbar button  in or select **View** → **Status Bar** from the menu bar.

✓ Note

If the EAW is already open, the results are added to the open window and it may be necessary to use the scrollbars at the bottom and right side of the EAW to see the newly added results.

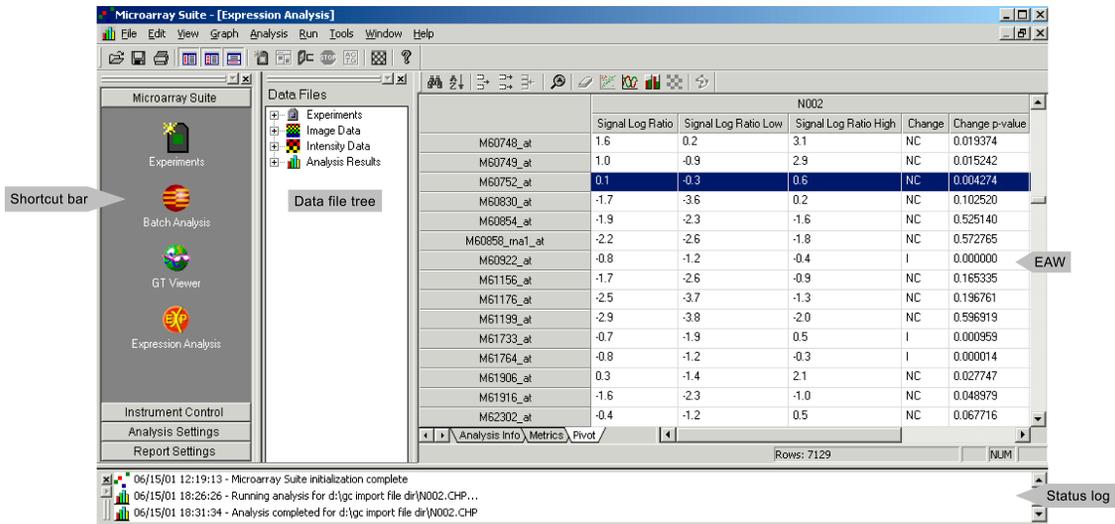


Figure 9.14
Expression Analysis Window (EAW), comparison analysis output (*.chp), pivot tab

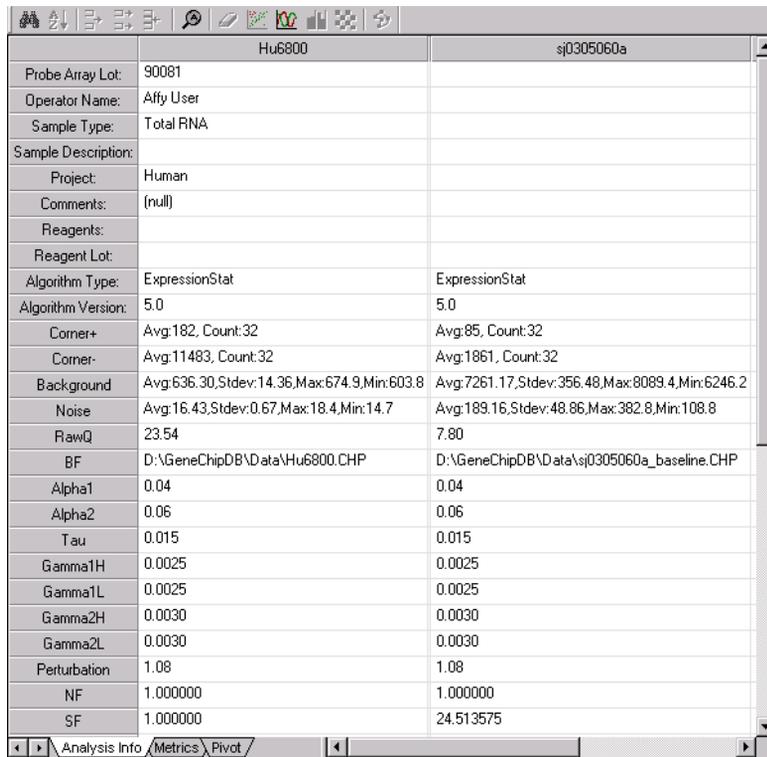
Expression Analysis Tabular Data

The Expression Analysis window (EAW) has three tables (Analysis Info, Metrics, and Pivot) that display the tabular and graphic analysis output for gene expression assays. The EAW opens after an expression analysis is completed or when an expression analysis output file (*.chp) is opened.

You can open more than one *.chp file in the EAW. Use the EAW scroll bars to view all of the results. A second toolbar (see Appendix J) and additional menu bar commands are available in the EAW.

Analysis Information Table

The analysis information table (Figure 9.15) displays experiment and sample information entered during experiment setup, values for user-modifiable expression analysis algorithm settings, and other information. Each column in the Analysis Information table represents one analysis (probe array).



	Hu6800	sj0305060a
Probe Array Lot:	90081	
Operator Name:	Aify User	
Sample Type:	Total RNA	
Sample Description:		
Project:	Human	
Comments:	(null)	
Reagents:		
Reagent Lot:		
Algorithm Type:	ExpressionStat	ExpressionStat
Algorithm Version:	5.0	5.0
Corner+	Avg:182, Count:32	Avg:85, Count:32
Corner-	Avg:11483, Count:32	Avg:1861, Count:32
Background	Avg:636.30,Stdev:14.36,Max:674.9,Min:603.8	Avg:7261.17,Stdev:356.48,Max:8089.4,Min:6246.2
Noise	Avg:16.43,Stdev:0.67,Max:18.4,Min:14.7	Avg:189.16,Stdev:48.86,Max:382.8,Min:108.8
RawQ	23.54	7.80
BF	D:\GeneChipDB\Data\Hu6800.CHP	D:\GeneChipDB\Data\sj0305060a_baseline.CHP
Alpha1	0.04	0.04
Alpha2	0.06	0.06
Tau	0.015	0.015
Gamma1H	0.0025	0.0025
Gamma1L	0.0025	0.0025
Gamma2H	0.0030	0.0030
Gamma2L	0.0030	0.0030
Perturbation	1.08	1.08
NF	1.000000	1.000000
SF	1.000000	24.513575

Figure 9.15
Expression Analysis window, Analysis Info tab

Exporting Analysis Information

You can export the data in the analysis information table to a tab delimited text file (*.txt) or Microsoft® Excel file (*.xls).

1. If it is not already open, open the EAW (in the data file tree, double-click each analysis output file (*.chp) of interest).
2. Click the Analysis Information tab.
3. To export a user-selected portion of the data table, highlight the cells of interest.
4. Click the **Save** toolbar button . Alternatively, select **File** → **Save As** from the menu bar.
⇒ The Save As dialog box appears (Figure 9.16).

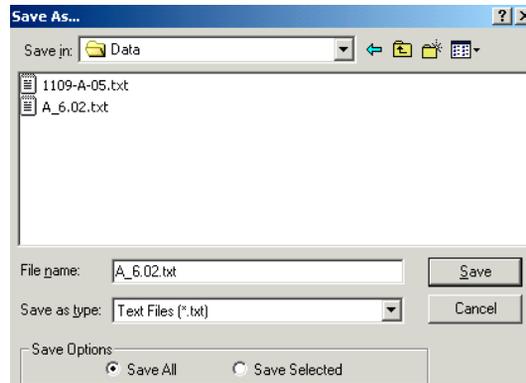


Figure 9.16
Save As dialog box

5. Choose a **Save Option**:
 - **Save All** exports all of the analysis information table data to the text file.
 - **Save Selected** exports only the user-selected portion of the analysis information table to the text file.
6. Enter a name for the text file and confirm or select another directory for the text file.

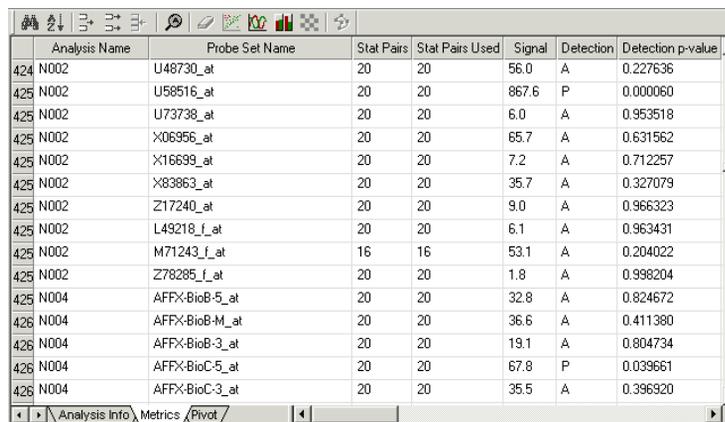
7. Click **Save**.
⇒ This exports the experiment information.

Metrics Table

The metrics table ([Figure 9.17](#)) displays data for each probe set in the selected analysis output file(s) (*.chp) ([Table 9.1](#)). See Appendix E for definitions of the Statistical expression metrics.

Table 9.1
Statistical expression algorithm metrics

Absolute Analysis	Comparison Analysis
Stat Pairs	Stat Common pairs
Stat Pairs Used	Signal log ratio
Signal	Signal log ratio low
Detection	Signal log ratio high
Detection p-value	Change
	Change p-value



Analysis Name	Probe Set Name	Stat Pairs	Stat Pairs Used	Signal	Detection	Detection p-value
424 N002	U48730_at	20	20	56.0	A	0.227636
425 N002	U58516_at	20	20	867.6	P	0.000060
425 N002	U73738_at	20	20	6.0	A	0.953518
425 N002	X06956_at	20	20	65.7	A	0.631562
425 N002	X16699_at	20	20	7.2	A	0.712257
425 N002	X83863_at	20	20	35.7	A	0.327079
425 N002	Z17240_at	20	20	9.0	A	0.966323
425 N002	L49218_f_at	20	20	6.1	A	0.963431
425 N002	M71243_f_at	16	16	53.1	A	0.204022
425 N002	Z78285_f_at	20	20	1.8	A	0.998204
425 N004	AFFX-BioB-5_at	20	20	32.8	A	0.824672
426 N004	AFFX-BioB-M_at	20	20	36.6	A	0.411380
426 N004	AFFX-BioB-3_at	20	20	19.1	A	0.804734
426 N004	AFFX-BioC-5_at	20	20	67.8	P	0.039661
426 N004	AFFX-BioC-3_at	20	20	35.5	A	0.396920

Figure 9.17
Expression Analysis window, metrics tab, absolute analysis output for experiments N002 and N004

Exporting Metrics Data

You can export metrics data and the associated analysis information as a text file (*.txt) or Microsoft® Excel file (*.xls).

Setting Metrics Export Preferences

1. Click the **Analysis Options** toolbar button .
⇒ The Analysis Options dialog box appears (Figure 9.18).
2. Click the metrics tab to view export preferences.

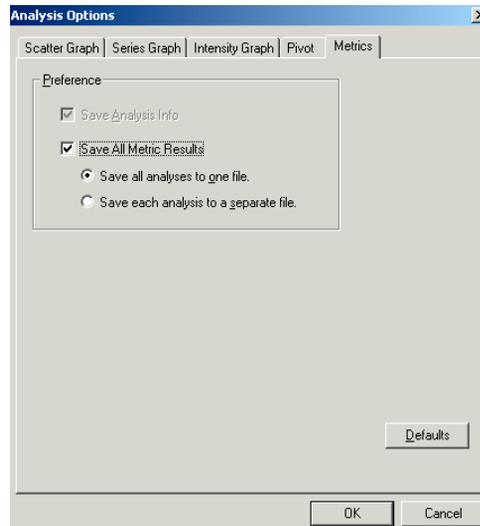


Figure 9.18
Analysis Options, Metrics tab

3. Choose the **Save Analysis Info** option to include the analysis information with the metrics data for export.
4. Choose the **Save All Metric Results** option to view the save preferences:
 - The **Save all analyses to one file** option exports the analysis information and metrics data for all analyses in the EAW to one text file (*.txt).

- The **Save each analysis to a separate file** option exports the analysis information and metrics data for each analysis in the EAW to a separate text file.

✓ Note

The save preferences set in the metrics tab do not affect the save function in the Pivot or Analysis Information tab in the EAW.

5. Click **OK** to close the Analysis Options dialog box.

Exporting Metrics Data

1. If it is not already open, open the EAW (in the data file tree, double-click each analysis output file (*.chp) of interest), then click the Metrics tab.
2. To export:
 - All of the metrics table, click the **Save** toolbar button . Alternatively, select **File** → **Save As** from the menu bar.
 - A selected portion of the metrics table, highlight the rows of interest, then click the **Save** toolbar button  or select **File** → **Save As** from the menu bar.

✓ Note

*Do not choose the **Save All Metric Results** option (Figure 9.18) if you wish to save selected results.*

⇒ The Save As dialog box appears (Figure 9.19).



Figure 9.19
Save As dialog box

3. Enter a name for the text file and confirm or select another directory for the text file.
4. Click **Save** to export the metrics data according to the preferences set in the Analysis Options (Metrics tab) (Figure 9.18).

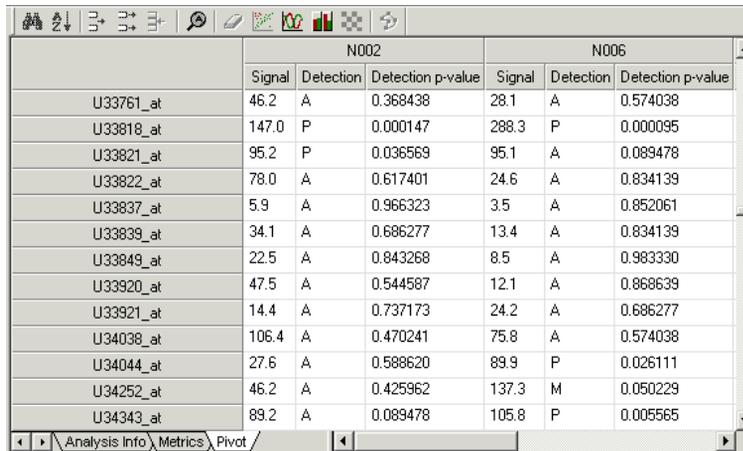
Pivot Table

To view the pivot table:

- open an expression analysis output file (*.chp) (double-click the file in the data tree)
- click the Pivot tab in the EAW

Each row in the pivot table (Figure 9.20) begins with a probe set name and each set of columns displays the results from one analysis output file (*.chp). The pivot table displays analysis output and descriptions for each transcript represented on the probe array.

If several *.chp from the same probe array type are open in the EAW, each row of the pivot table displays the data for one probe set from the different experiments. The pivot table is organized so that probe set data from different experiments may be easily compared.



	N002			N006		
	Signal	Detection	Detection p-value	Signal	Detection	Detection p-value
U33761_at	46.2	A	0.368438	28.1	A	0.574038
U33818_at	147.0	P	0.000147	288.3	P	0.000095
U33821_at	95.2	P	0.036569	95.1	A	0.089478
U33822_at	78.0	A	0.617401	24.6	A	0.834139
U33837_at	5.9	A	0.966323	3.5	A	0.852061
U33839_at	34.1	A	0.686277	13.4	A	0.834139
U33849_at	22.5	A	0.843268	8.5	A	0.983330
U33920_at	47.5	A	0.544587	12.1	A	0.868639
U33921_at	14.4	A	0.737173	24.2	A	0.686277
U34038_at	106.4	A	0.470241	75.8	A	0.574038
U34044_at	27.6	A	0.588620	89.9	P	0.026111
U34252_at	46.2	A	0.425962	137.3	M	0.050229
U34343_at	89.2	A	0.089478	105.8	P	0.005565

Figure 9.20

Expression Analysis window (EAW), pivot table, absolute analysis output (*.chp) for experiments N002 and N006

Pivot Table Display Options

Absolute or Comparison analysis output

You can select the types of absolute or comparison analysis output (columns) displayed in the pivot table.

1. Click the **Analysis Options** toolbar button . Alternatively, select **Analysis** → **Options** from the menu bar.
⇒ The Analysis Options dialog box appears (Figure 9.21).
2. Click the Pivot tab.

The pivot table includes column headers for check marked items in the Analysis Options dialog box. You can also view these items by selecting **View** → **Pivot Data** from the menu bar.

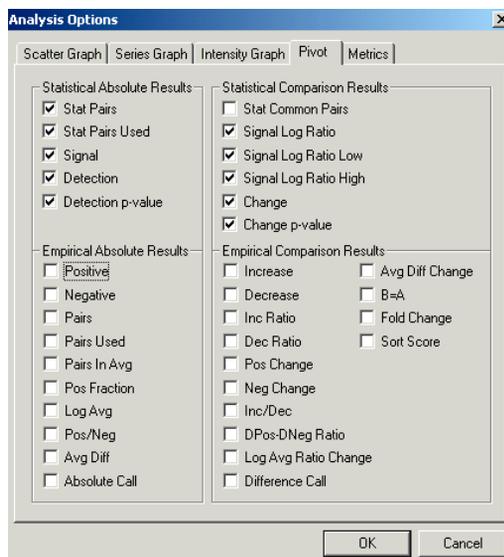


Figure 9.21
Analysis Options dialog box, Pivot tab

Note

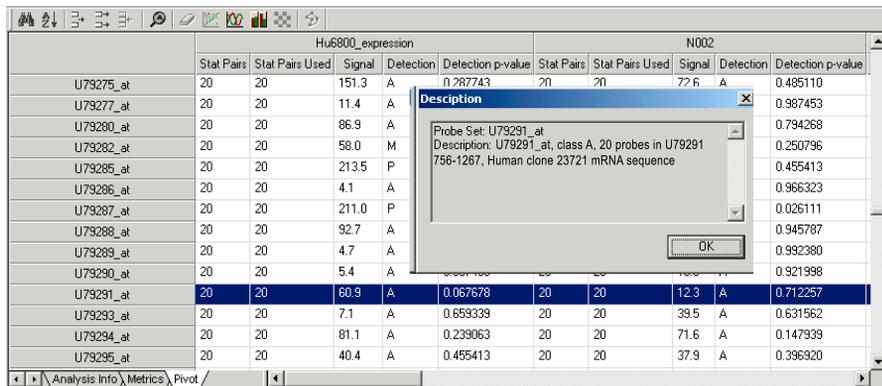
The pivot table can also display analysis output (*.chp) for experiments analyzed using the Empirical expression analysis algorithm (in Microarray Suite versions lower than 5.0).

3. Click an unchecked item to add it to the pivot table or click a checked item to remove it from the table.
4. To reorder an experiment result column in the pivot table, use the drag-and-drop method to move the column header in the table.

Probe Set Descriptions

To view probe set descriptions, do either of the following:

- scroll to the Descriptions (far right column) in the pivot table
- double-click the row of interest in the pivot table
 - ⇒ This displays the Description dialog box (Figure 9.22).



	Hu6800_expression				N002					
	Stat Pairs	Stat Pairs Used	Signal	Detection	Detection p-value	Stat Pairs	Stat Pairs Used	Signal	Detection	Detection p-value
U79275_at	20	20	151.3	A	0.287743	20	20	77.6	A	0.485110
U79277_at	20	20	11.4	A						0.987453
U79280_at	20	20	86.9	A						0.794268
U79282_at	20	20	58.0	M						0.250796
U79285_at	20	20	213.5	P						0.455413
U79286_at	20	20	4.1	A						0.966323
U79287_at	20	20	211.0	P						0.026111
U79288_at	20	20	92.7	A						0.945787
U79289_at	20	20	4.7	A						0.932380
U79290_at	20	20	5.4	A						0.921998
U79291_at	20	20	60.9	A	0.067678	20	20	12.3	A	0.712257
U79293_at	20	20	7.1	A	0.659339	20	20	39.5	A	0.631562
U79294_at	20	20	81.1	A	0.239063	20	20	71.6	A	0.147939
U79295_at	20	20	40.4	A	0.455413	20	20	37.9	A	0.396920

Figure 9.22
EAW, probe set Description dialog box

Obtaining Further Information

You can search for further information about a probe set at an Internet site.

1. In the metrics or pivot table:
 - a. Right-click the probe set (row) of interest.
 - b. Select **External Database** → **NetAffx** (or **Entrez**) from the shortcut menu.
 - ⇒ This starts the default Internet browser (for example, Microsoft® Internet Explorer or Netscape® Communicator) and automatically opens the selected Internet site.

Exporting the Pivot Table

You can export the pivot table as a text file (*.txt) or Excel file (*.xls).

1. If it is not already open, open the EAW (in the data file tree, double-click each analysis output file (*.chp) of interest).
2. Click the Pivot tab.
3. To export only a user-selected portion of the data table, highlight the cells of interest.
4. Click the **Save** toolbar button  or select **File** → **Save As** from the menu bar.
⇒ The **Save As** dialog box appears ([Figure 9.23](#)).

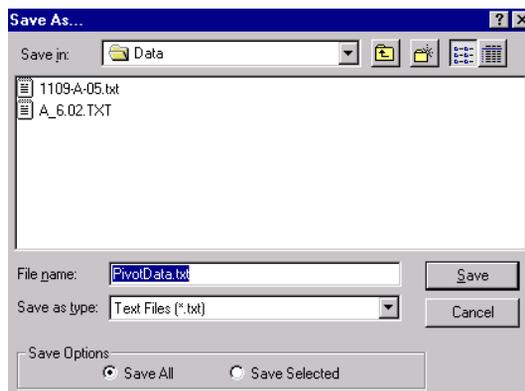


Figure 9.23
Save As dialog box

5. Choose a save option:
 - **Save All** places all of the pivot table in the text file.
 - **Save Selected** places only the selected pivot data in the text file.
6. Enter a name for the text file and confirm or select another directory for the text file.
7. Click **Save** to export the pivot table or selected pivot data.

Finding Probe Set Information

You can perform text searches in the EAW.

- To start a text search, do one of the following:
 - click the **Find** toolbar button 
 - select **Edit** → **Find** from the menu bar
 - right-click the metrics or pivot table and select **Find** in the shortcut menu.

⇒ The Find dialog box appears (Figure 9.24).

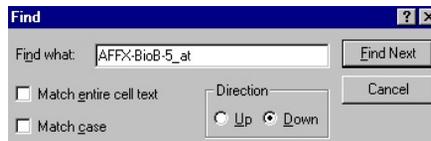


Figure 9.24
Find dialog box

- Enter the text string for the search (up to 256 alphanumeric characters).
- Choose the **Match** or **Direction** search option, then click **Find Next**. Click **Find Next** again to continue the search.

✓ Note

The **Find** command finds all strings that match the text string for the search. For example, using the **Find** command to search for the text string *biob* would find *AFFX-BioB-5* as well as other occurrences of *BioB*.

Sorting Probe Set Information

You can sort the metrics or pivot table by any category of data displayed in the table.

Single Column Sort

- In the EAW, click the Metrics or Pivot tab.
- Right-click the column header of interest and click **Sort Ascending** or **Sort Descending** in the shortcut menu.

Multiple Column Sort

1. In the EAW, click the Metrics or Pivot tab.
2. Click the **Sort** toolbar button . Alternatively, select **Edit** → **Sort** from the menu bar.
⇒ The Sort dialog box appears ([Figure 9.25](#)).

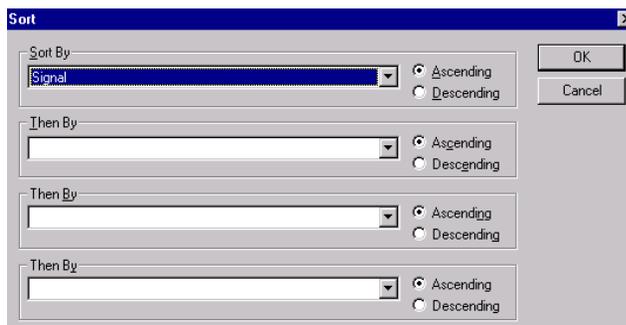


Figure 9.25
Sort dialog box

3. Select the desired sort category from the first (top) drop-down list and choose an ascending or descending sort order.
Up to three additional sort categories may be specified.
4. Click **OK** when finished specifying the sort.
The sort category and order are displayed at the bottom of the EAW.

Reordering or Removing Probe Set Information

Reordering Columns in the Metrics or Pivot Table

1. In the Metrics or Pivot tab, click the column header of interest
2. Use the drag-and-drop method to move the column to a new position in the table.

Reordering Rows in the Metrics or Pivot Table

Use the **Edit** commands in the menu bar to remove or reorder tabular data in the metrics or pivot table:

Edit → Move Selected to Top Moves a selected row(s) to the top of the metric or pivot table.

Edit → Remove Analyses When the pivot table contains only one analysis, removes the analysis from the table (equivalent to closing the *.chp file). When the pivot table includes more than one analysis, this displays an Open dialog box that enables you to select the analyses you wish to remove from the table.

Edit → Select All Highlights all rows in the metric or pivot table.

Hiding Probe Set Information

You can hide tabular data in the metrics or pivot table. Use the View commands in the menu bar to hide (unhide) selected rows or all of the tabular data.

Alternatively, right-click the metrics or pivot table to display a shortcut menu of hide commands:

View → Hide Selected Hides the currently selected row(s) in the metrics or pivot table. The status bar at the bottom of the EAW displays the remaining number of rows in the table.

View → Hide Unselected Hides all unselected rows in the metrics or pivot table. The status bar at the bottom of the EAW displays the remaining number of rows in the data table.

View → Unhide All Displays all previously hidden rows at the bottom of the data table.

Expression Analysis

Measured Images

A measured image displays the probe set hybridization intensity data (from the *.dat).

Displaying Measured Images

1. Click the desired probe set name(s) in the metrics or pivot table.

To select adjacent probe set names, press and hold the **Shift** key and click the first and last probe set name in the selection. To select non-adjacent probe set names, press and hold the **Ctrl** key while you click the desired names.

2. Click the **Image** toolbar button . Alternatively, select **Graphs** → **Measured Images** from the menu bar.

⇒ The graph pane (Intensity tab) displays the measured image and the intensity signal range for the selected probe set ([Figure 9.26](#)).

When an additional image is created, the graph pane automatically scrolls to the new image. Use the scroll bar to view all of the measured images created during a session. (See [Appendix K](#) for information about resizing the graph pane.)

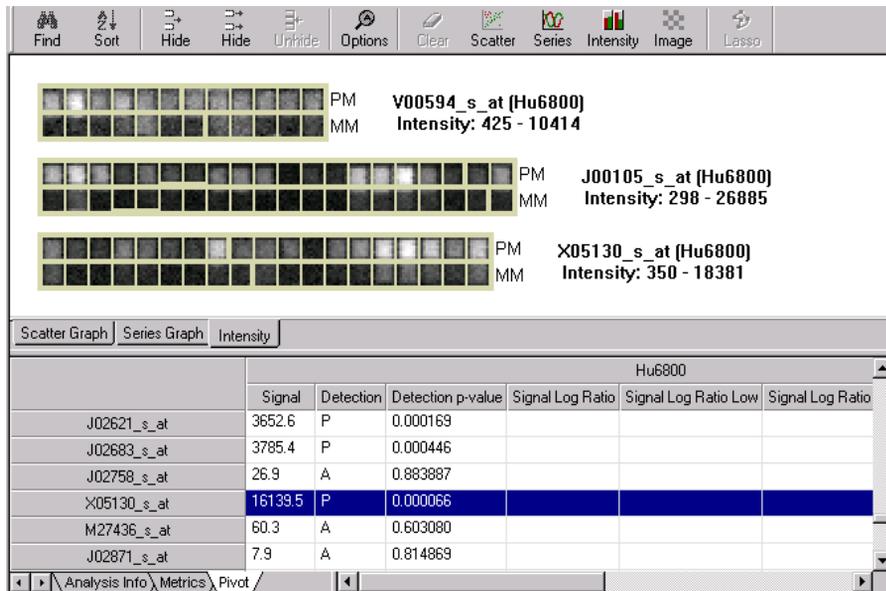


Figure 9.26
EAW, measured images in the graph pane (top)

Clearing Measured Images

You can remove all or selected images from the graph pane.

1. To remove all measured images from the graph pane, right-click the graph pane and select **Clear All Graphs** from the shortcut menu. Alternatively, select **Graph** → **Clear All Graphs** from the menu bar.
2. To remove selected images from the graph pane:
 - a. Double-click the measured image.
 - b. Right-click the image and select **Clear Selected Graphs** from the shortcut menu or select **Graph** → **Clear Selected Graphs** from the menu bar.

Expression Analysis Graphs

Microarray Suite plots three types of expression analysis graphs ([Table 9.2](#)). The graphs are displayed in the graph pane that appears at the top of the EAW ([Figure 9.27](#)).

Table 9.2
Expression analysis graphs

Expression Graph Type	Plots...
Scatter	a user-specified metric for all probe sets from two experiments (of the same probe array type) using a traditional scatter plot
Series	a user-specified metric for each probe set on the probe array in a bar or line graph format
Intensity	the intensities of the probe pairs in a probe set using a bar graph format

The graph pane has three tabs: Scatter Graph, Series Graph, and Intensity. Click the graph pane tabs to toggle the display between the different graph types. (See [Appendix K](#) for information about resizing the graph pane.)

Note

The intensity graph  and images  functions are unavailable until a probe set(s) has been highlighted in the metrics or pivot table.

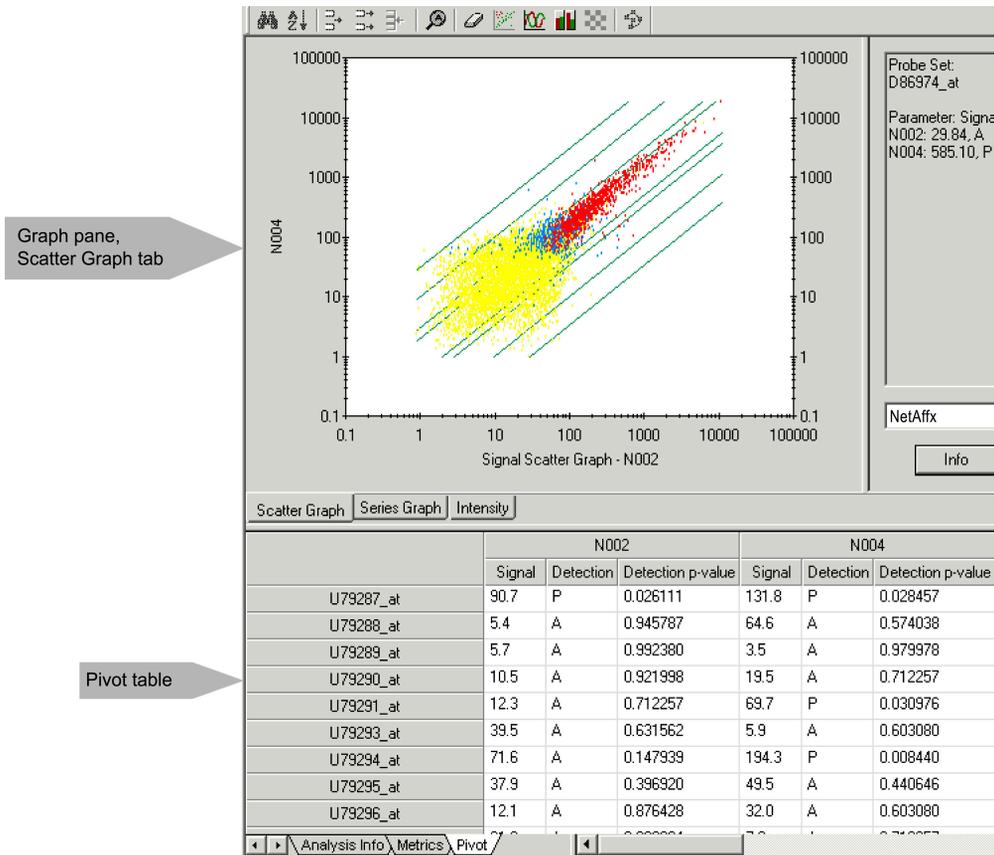


Figure 9.27
Expression Analysis window (EAW), Scatter graph tab of the graph pane (top) and Pivot tab (bottom)

Scatter Graph

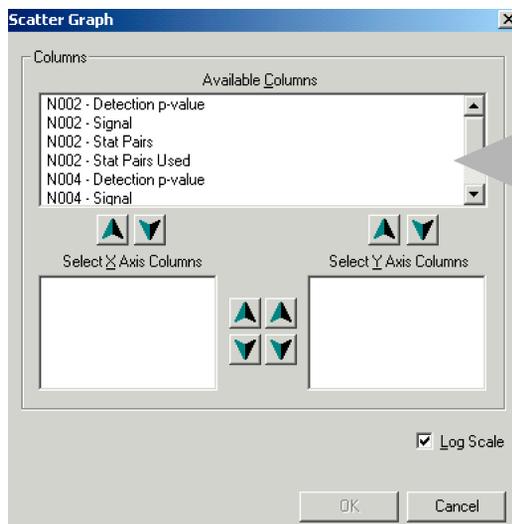
The scatter graph plots user-selected expression data for all probe sets from one or more pairs of analyses using a traditional scatter plot (Figure 9.27). Each point in the scatter graph represents a probe set common to both experiments and is defined by the intersection of the analysis result value on the x and y axes for the common probe set. The scatter graph displays fold change lines (four pairs) that help identify probe sets in the comparison that change expression levels.

Plotting a Scatter Graph

1. Click the Scatter graph toolbar button . Alternatively, select **Graph** → **Scatter Correlation Graph** from the menu bar.
⇒ The Scatter Graph dialog box displays the pivot table columns available for the scatter graph (Figure 9.28).

✓ Note

Select pairs of analyses (*.chp) for the scatter plot that are derived from the same probe array type. The order of the analyses is critical. For example, the analyses from a multiple probe array set must be ordered identically along the x and y-axes since the scatter graph compares the first set of analysis output on the x-axis with the first set of analysis output on the y-axis (and so forth). If the analysis output are not in identical order, many probe sets will not be compared and plotted (only the probe sets in common such as the controls).



This box shows the pivot table columns available for the scatter graph

Figure 9.28
Scatter graph dialog box

2. To select a column for the x-axis, do either of the following:
 - Use the drag-and-drop method to move the column name from the **Available Columns** box to the **Select X Axis Columns** box (Figure 9.28).
 - Highlight the column name in the **Available Columns** box, then click the down arrow  (on the left side) to place the column in the **Select X Axis Columns** box.
3. To select a column for the y-axis, do either of the following:
 - Use the drag-and-drop method to move the column name from the **Available Columns** box to the **Select Y Axis Columns** box (Figure 9.28).
 - Highlight the column name in the **Available Columns** box and click the down arrow  (on the right side) to place the column in the **Select Y Axis Columns** box.
4. To change the order of columns, do either of the following:
 - Use the drag-and-drop method to move the column to the desired location in the list.
 - Highlight the column name, then use the up  and down  arrows located to the inside of the **Select Axis Columns** boxes.
5. To select linear scaling, remove the check mark from the **Log Scale** check box.

Logarithmic scaling is the default.
6. Click **OK** to generate the scatter graph.
 - ⇒ The graph pane (Scatter Graph tab) displays the scatter graph (Figure 9.27).

Working with the Scatter Graph

Right-click the scatter graph to display a shortcut menu.

Points	Displays the absolute call combinations and the color assigned to each. The scatter graph displays the check marked items. Click an item to remove or add a check mark. The scatter graph is immediately updated.
Copy Graph	Copies the Scatter graph to the system clipboard.
Close Graph Pane	Closes the graph pane. Graphs created during the session are not saved.
Full Out Zoom	Restores the graph to the original magnification view.
Redraw Graph	Redraws the scatter graph so that it plots only the points that correspond to rows displayed in the pivot table. Points that correspond to hidden rows are not plotted.
Select Highlighted Points	Selects rows in the Pivot and Metric tables that correspond to points selected (<i>roped</i>) in the Scatter graph.
Options	Displays the Analysis Options dialog box.

Magnifying the Graph

1. Press and hold the **Shift** key while using the click-and-drag method to draw a rectangle over the graph area of interest ([Figure 9.29](#)).
2. Release the mouse key.
⇒ This magnifies the area selected by the rectangle ([Figure 9.30](#)).
3. To zoom out and restore the graph, right-click the graph and select **Full Out Zoom** from the shortcut menu.

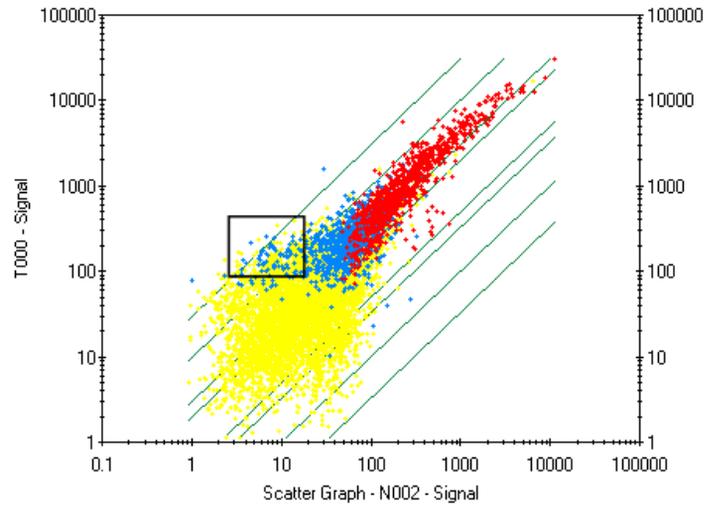


Figure 9.29
Scatter graph, rectangle selects an area to magnify (GeneChip® data mode)

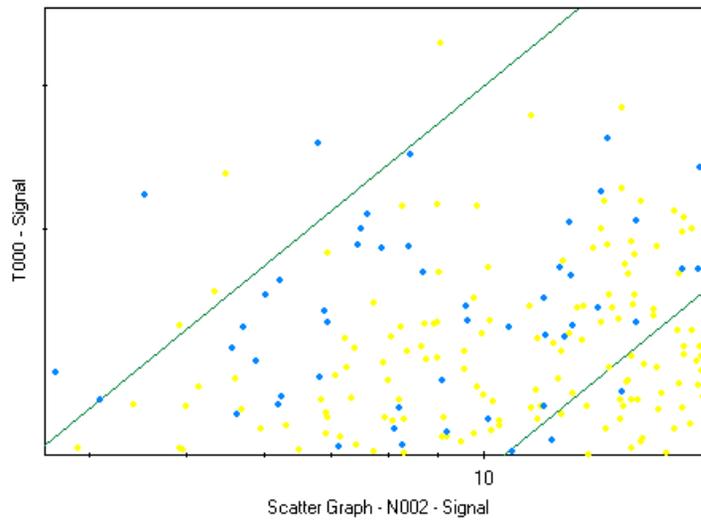


Figure 9.30
Magnified area in the scatter graph

Displaying Probe Set and Gene Information

1. Click a point in the scatter graph.
⇒ This displays the probe set name, parameter value, detection call, and a brief description of the probe set (Figure 9.31). The corresponding row in the pivot table is highlighted.

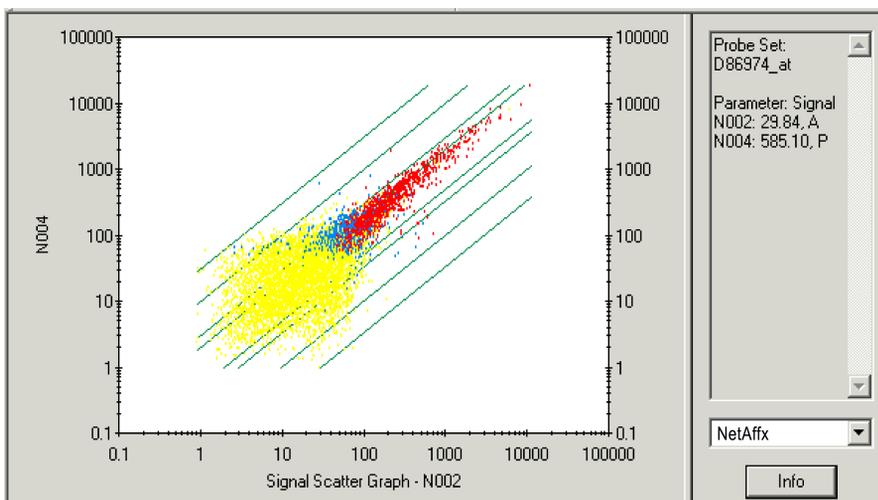


Figure 9.31
Scatter graph, click a point to display corresponding probe set and gene information

2. To obtain further gene information, select a database from the drop-down list and click **Info**.
⇒ The Internet browser window opens and displays information for the selected gene (Figure 9.32).

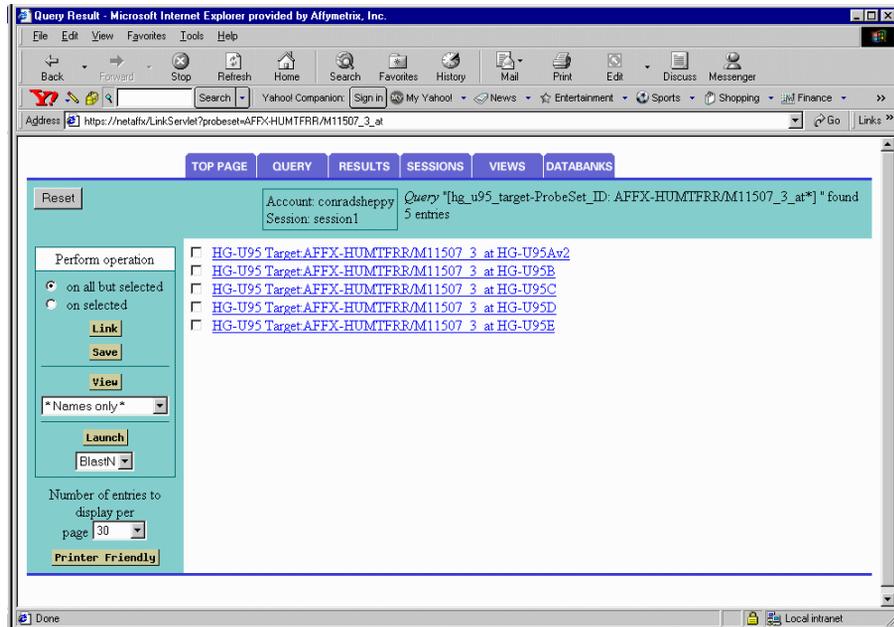


Figure 9.32
Internet browser window

Selecting Points in the Scatter Graph

1. Click the **Lasso** toolbar button . Alternatively, select **Graph** → **Lasso Points** from the menu bar.
 - ⇒ The mouse pointer changes to a cross hairs (+) when it is positioned over the scatter graph
2. To rope points of interest, position the cross hairs near the group of points, then do either of the following:
 - Use the click-and-drag method to draw a circle around the points, then release the mouse button near the starting point.
 - Click the mouse, move it to draw a line segment, then click the mouse again to start drawing a new line segment. Repeat until you return the cross hairs to the starting point and the line segments enclose the points of interest.
 - ⇒ This completes the lasso operation (**Figure 9.33**).

The scatter graph displays the selected points in orange. The metrics and pivot tables highlight the corresponding probe sets.

**TIP**

Press the **Esc** key to cancel the lasso operation.

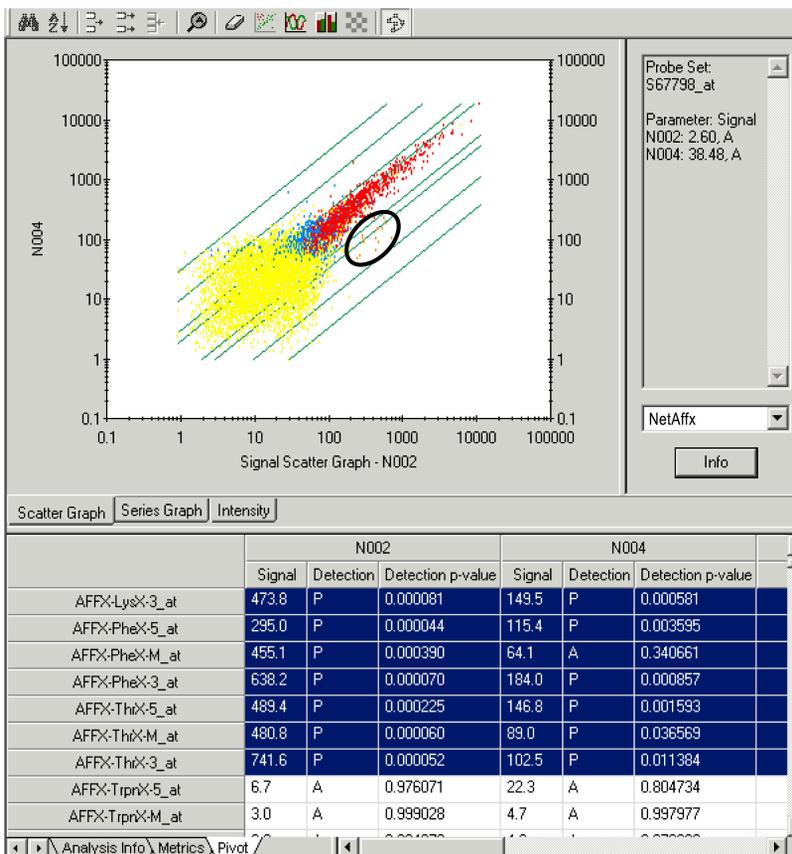


Figure 9.33

Roping points in the scatter graph highlights the corresponding probe set rows in the metrics and pivot tables

**TIP**

After you select points in the scatter graph, hide the unselected rows in the pivot table, then use the **Redraw Graph** command to display only the selected points in the scatter graph.

Hiding Points in the Scatter Graph

You can hide points in the scatter graph according to detection or change call (if the scatter graph points are color coded using the **Color by Detection/Absolute Call** or **Color by Change/Difference Call** option, see Point Options on page 184).

1. Right-click the scatter graph and select **Points** from the shortcut menu.
 - ⇒ This displays a key for the scatter graph point colors (Figure 9.34).

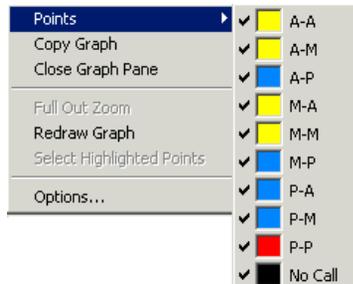


Figure 9.34

Scatter graph shortcut menu, point color combinations for detection calls

2. Click a point color combination to remove (or add) a check mark.
 - ⇒ This hides (unhides) the selected points in the scatter graph.

Scatter Graph Options

You can modify some display features of the scatter graph. The Analysis Options dialog box displays the user-modifiable features (Figure 9.35).

1. To open the Analysis Options dialog box, do one of the following:
 - Click the **Options** toolbar button .
 - Right-click the scatter graph and select **Options** from the shortcut menu.
 - Select **Analysis** → **Options** from the menu bar.
 - ⇒ The Analysis Options dialog box appears (Figure 9.35).
2. Click the Scatter Graph tab.

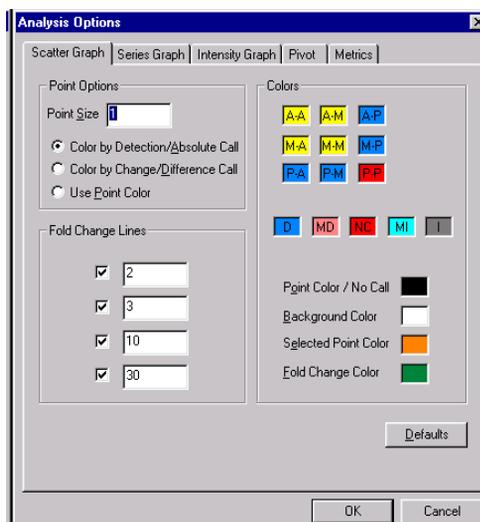


Figure 9.35
Analysis Options, Scatter Graph tab

Point Options

Point Size Determines the size of the graph points. Enter a larger point size number for easier viewing, but use a smaller point size for high resolution graphs.

Color by Detection/Absolute Call Displays the graph points using the colors assigned to the absolute call combination of the x and y-axis analyses. There are nine possible call combinations:

	Absent in Y	Marginal in Y	Present in Y
Absent in X	A-A	M-A	A-P
Marginal in X	M-A	M-M	M-P
Present in X	P-A	P-M	P-P

**TIP**

Color coding the combinations can help quickly determine whether a highly expressed gene in one experiment was present or absent in another experiment.

**Note**

If the call is No Call in either analysis, the point is displayed using the No Call point color (the default is black).

Color by Change/ Difference Call	Displays the data points using the colors assigned to the comparison call of the y-axis analysis. There are five possible calls: Decrease (D), Marginal Decrease (MD), No Change (NC), Marginal Increase (MI), and Increase (I).
Use Point Color	Displays points using the color assigned to the Point Color.

Fold Change Lines

The default fold change lines are defined in pairs: $y = 2x$ and $y = 1/2x$, $y = 3x$ and $y = 1/3x$, $y = 10x$ and $y = 1/10x$, $y = 30x$ and $y = 1/30x$.

The fold change lines (four pairs) are drawn using the values entered in the edit boxes. Only integer values may be entered. Removing a check mark turns off the display of that pair of fold change lines.

Colors

You can change the colors assigned to the:

- absolute call categories
 - difference call categories
 - scatter graph points
 - graph background
 - selected (roped) scatter graph points
 - fold change lines
1. In the Analysis Options dialog box ([Figure 9.35](#)), click the color square associated with the item you wish to change.
 - ⇒ This displays the color palette ([Figure 9.36](#)).

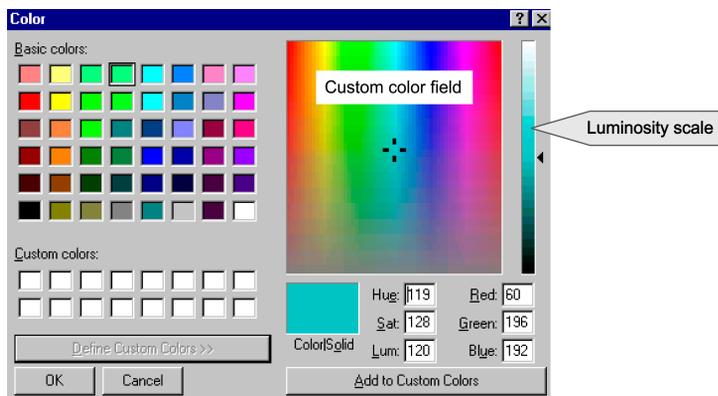


Figure 9.36
Color palette

2. To select a predefined color, click one of the basic colors.
3. To define a custom color, click **Define Custom Colors**, then use the click-and-drag method to move the cross hairs in the custom color field. Adjust the color brightness using the luminosity scale to the right. When finished, click **Add to Custom Colors** to apply the color.
4. Click **OK** to close the color palette.

The color choices are saved on a per user basis.

Defaults

Click **Defaults** to return all analysis option settings to the factory set defaults.

Series Graph

The series graph plots a line or bar graph of user-selected expression data for all analyses or for all probe sets of analyses open in the EAW (Figure 9.37 and Figure 9.38). The line graph format is the default.



Hidden probe sets are not included in the series graph.

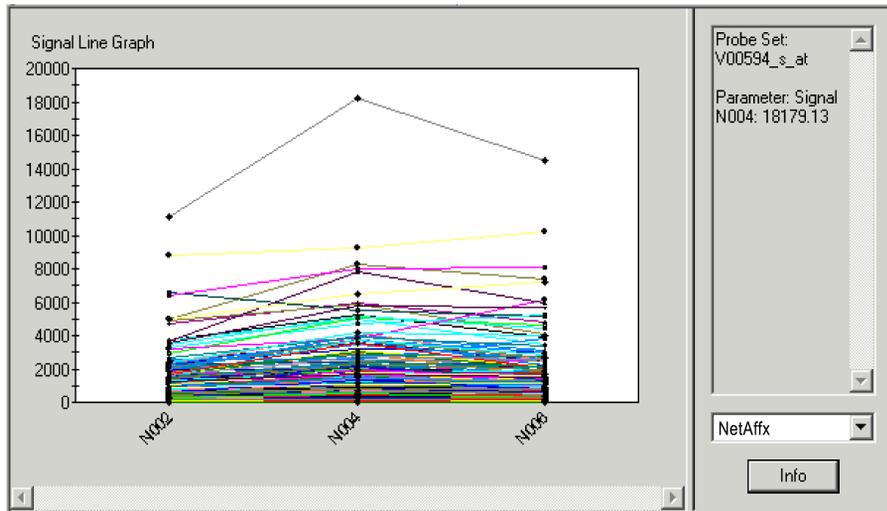


Figure 9.37
Series line graph, signal

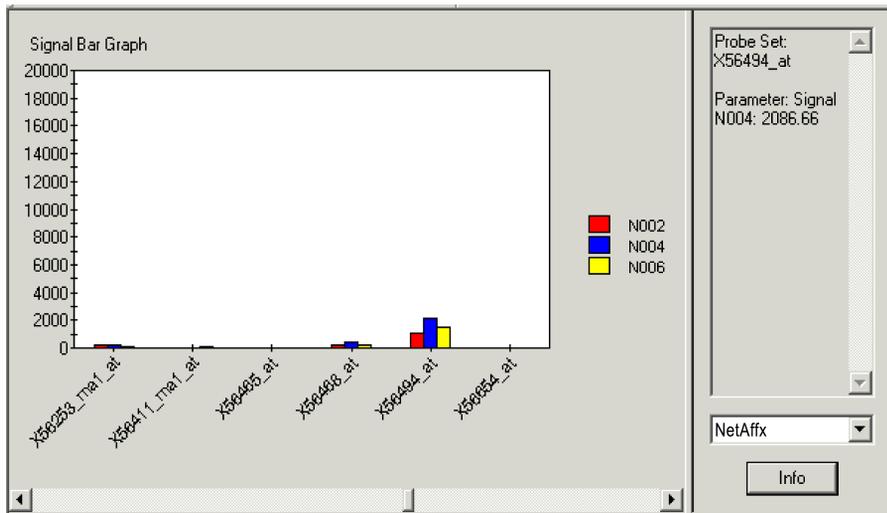


Figure 9.38
Series bar graph, signal

Plotting a Series Graph

1. Click the **Series Graph** toolbar button . Alternatively, select **Graph** → **Series Graph** from the menu bar.
⇒ The Series Graph dialog box appears ([Figure 9.39](#)).

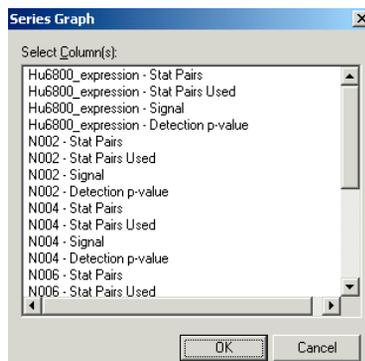


Figure 9.39
Series Graph dialog box

2. Select the pivot table columns for the series graph.
⇒ This generates the series line graph for the probe sets displayed in the data tables ([Figure 9.37](#)).

The line graph format is the default. For information on how to change the bar graph format, see [Series Graph Options](#) below.

Working with the Series Graph

1. Place the mouse arrow over a probe set name, data point, or bar to view a pop-up tool tip that displays the analysis or probe set name and parameter value.
2. Right-click the series graph to display a shortcut menu.

- | | |
|-------------------------|--|
| Copy Graph | Copies the graph to the system clipboard. |
| Close Graph Pane | Closes the graph pane (graphs created during the session are not saved). |
| Options | Displays the Analysis Options dialog box. |

Series Graph Options

Some display features of the series graph may be modified. The Analysis Options dialog box displays the user-modifiable features (Figure 9.40).

- To open the Analysis Options dialog box, do one of the following:
 - Click the **Options** toolbar button .
 - Right-click the scatter graph and select **Options** from the shortcut menu.
 - Select **Analysis** → **Options** from the menu bar.

⇒ The Analysis Options dialog box appears (Figure 9.40)
- Click the Series Graph tab and choose the **Bar Graph** or **Line Graph** option.

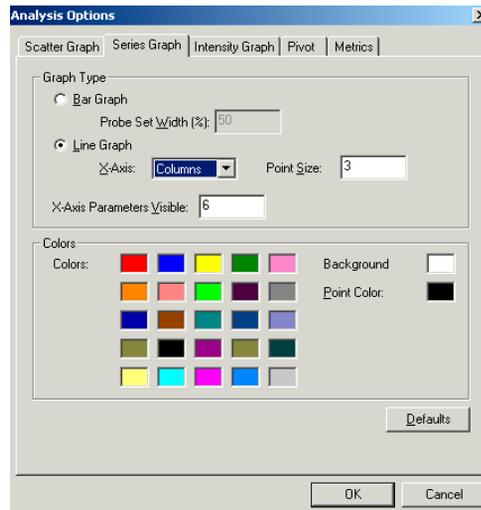


Figure 9.40
Analysis Options, Series Graph tab

Series Bar Graph Options

Probe Set Width (%)	Determines the graph bar width.
X-Axis Parameters Visible	The number of probe sets displayed on the x-axis within the viewable portion of the graph.

Series Line Graph Options

X-Axis	Select Probe Sets or Columns (from the X-Axis drop-down list) for display on the x-axis.
Point Size	Determines the size of the graph points.
X-Axis Parameters Visible	Specifies the number of probe sets or columns displayed on the x-axis within the viewable portion of the graph.

Colors

The color of the points or background in the series line graph may be changed.

1. In the Analysis Options dialog box (Figure 9.40), click the **Point Color** or **Background** color square.
⇒ This displays the color palette (Figure 9.41).

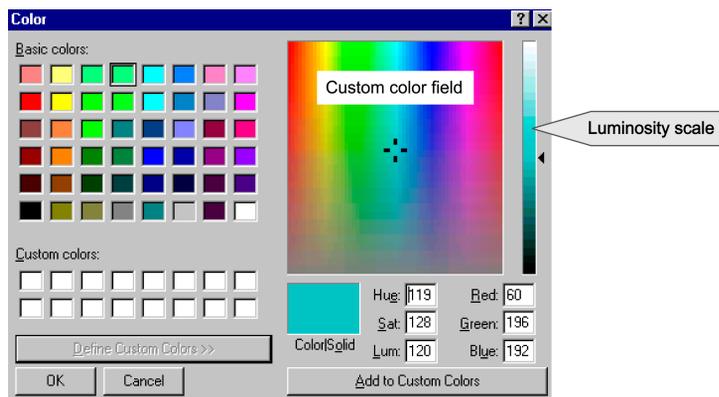


Figure 9.41
Color palette

2. To select a predefined color, click one of the basic colors.
3. To define a custom color, click **Define Custom Colors**, then use the click-and-drag method to move the cross hairs in the custom color field. Adjust the color brightness using the luminosity scale to the right. When finished, click **Add to Custom Colors** to apply the color

4. Click **OK** to close the color palette.

The color choices are saved on a per user basis.

Intensity Bar Graph

The intensity bar graph displays the intensities of the probe pairs of a probe set in a bar graph format (Figure 9.42). The intensity bar graph is a useful way to:

- evaluate the relative performance of the probe pairs in a probe set
- determine whether an image or probe mask has been applied to a probe pair

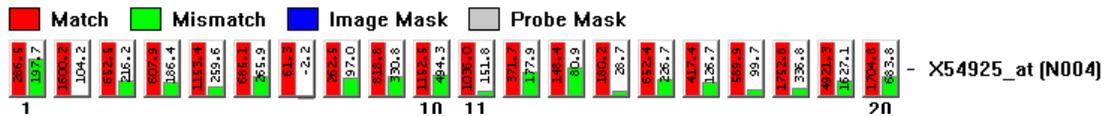


Figure 9.42
Intensity Bar graph of probe set X54925_at (from experiment N004)

Plotting the Intensity Bar Graph

1. Highlight a row or cell in the metrics or pivot table.
2. To plot the intensity bar graph(s) for the highlighted probe set(s), do either of the following:
 - Click the **Intensity Bar Graph** toolbar button .
 - Select **Graph** → **Intensity Bar Graph** from the menu bar.
3. If necessary, use the scroll bar to view all of the intensity bar graphs created during a session.

Intensity Bar Graph Options

You can modify some display features of the intensity bar graph. The Analysis Options dialog box displays the user-modifiable features (Figure 9.43).

1. To open the Analysis Options dialog box, do one of the following:
 - click the **Options** toolbar button 
 - right-click the intensity bar graph, then click **Options** in the shortcut menu
 - select **Analysis** → **Options** from the menu bar⇒ The Analysis Options dialog box appears (Figure 9.43)
2. Click the Intensity Graph tab.

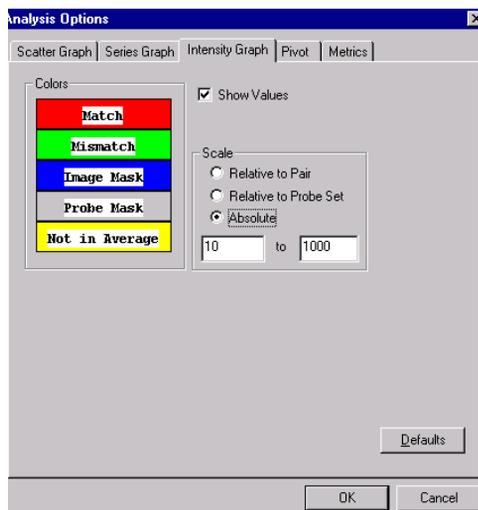


Figure 9.43
Analysis Options, Intensity Graph tab

Show Values

Choose the **Show Values** option to display numerical intensity values on the intensity bar graph.

Scale

Relative to Pair	Scales each probe pair set independently. The scale for each PM and MM pair is 0→max (PM, MM).
Relative to Probe Set	Scales each probe pair relative to all probe pairs of a particular probe set. The scale is determined by all pairs of the probe set and ranges from 0→max (all PM/MM probe pairs in the set).
Absolute	Scales all probe pairs according to a user specified intensity range (for example, 10 to 1000 in Figure 9.43).

Colors

The color code specifies the colors for:

Match	Perfect match probe intensities
Mismatch	Mismatch probe intensities
Image Mask	Probe pairs included in an image mask
Probe Mask	Probe pairs included in a probe mask
Not in Average	Probe pairs excluded from the Average Difference calculation (Empirical expression algorithm only)

To change the colors specified by the color code:

1. Click the color square of the item you wish to change (for example, **Match** in [Figure 9.43](#)).
⇒ This displays the color palette ([Figure 9.44](#)).

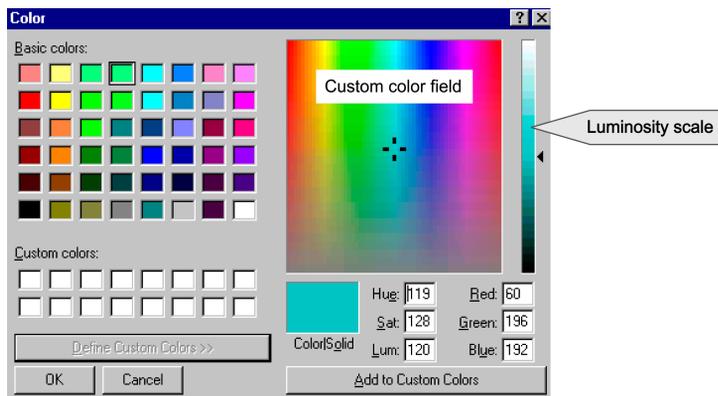


Figure 9.44
Color palette

2. To select a predefined color, click one of the basic colors.
3. To define a custom color, click **Define Custom Colors**, then use the click-and-drag method to move the cross hairs in the custom color field. Adjust the color brightness using the luminosity scale to the right. When finished, click **Add to Custom Colors** to apply the color
4. Click **OK** to close the color palette.

The color choices are saved on a per user basis.

Defaults

Click **Defaults** to restore the factory set defaults for all intensity bar graph options.

Clearing Intensity Bar Graphs

1. To remove all intensity bar graphs from the graph pane, right-click the graph pane and select **Clear All Graphs** from the shortcut menu. Alternatively, select **Graph** → **Clear All Graphs** from the menu bar.
2. To remove user-selected intensity bar graphs from the graph pane:
 - a. double-click the graph(s) you want to remove
 - b. right-click the selected graph and select **Clear Selected Graphs** from the shortcut menu. Alternatively, choose **Graph** → **Clear Selected Graphs** from the menu bar.

Expression Report

The expression report summarizes information about the expression analysis settings, algorithm settings, and probe set hybridization intensity data.

Generating an Expression Report

1. Confirm or change the user-modifiable report settings (see the next section *Expression Report Settings*).
2. To generate the report, do either of the following:
 - In the data file tree, right-click the analysis output file (*.chp) and select **Report** from the shortcut menu.
 - Select **File** → **Report** from the menu bar, then double-click an analysis output file (*.chp) in the Report dialog box (Figure 9.45) that appears.

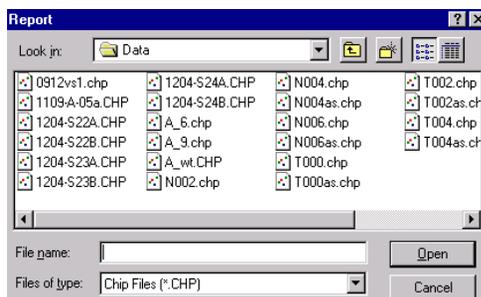


Figure 9.45
Report dialog box

⇒ This displays the expression report in the main display area.
Figure 9.46 and **Figure 9.47** show an example expression report.

- 3.** To best view the expression report:
 - a.** Use Arial font (10 point) (select **View** → **Set Font** from the menu bar).
 - b.** Set the tab stops set at 10 (select **View** → **Set Tab Stops** from the menu bar).

✓ Note

*A previously generated report will be overwritten. To save a previous report, rename it and save it under the new name. (See the section *Saving an Expression Report.*)*

Report Type: Expression Report
Date: 11:00AM 06/13/2001

Filename: N006AS.CHP
Probe Array Type: Hu6800
Algorithm: Statistical
Probe Pair Thr: 8
Controls: Antisense

Alpha1: 0.04
Alpha2: 0.06
Tau: 0.015
Noise (RawQ): 34.840
Scale Factor (SF): 1.000
Norm Factor (NF): 0.910

Baseline filename: N004AS.CHP
Gamma1H: 0.0025
Gamma1L: 0.0025
Gamma2H: 0.0030
Gamma2L: 0.0030
Perturbation: 1.22
Baseline Noise (RawQ): 36.920
Baseline Scale Factor (SF): 1.000000

Background:				
	Avg: 826.54	Std: 18.25	Min: 788.10	Max: 869.20
Noise:				
	Avg: 26.71	Std: 1.06	Min: 25.00	Max: 30.40
Corner+				
	Avg: 24217	Count: 32		

Figure 9.46

Expression report, comparison analysis, Statistical algorithm (page 1)

The following data represents probe sets that exceed the probe pair threshold and are not called "No Call".

Total Probe Sets:	7056	
Number Present:	1878	26.6%
Number Absent:	4988	70.7%
Number Marginal:	190	2.7%
Average Signal (P):	1202.8	
Average Signal (A):	113.7	
Average Signal (M):	294.0	
Average Signal (All):	408.4	
Number Increase:	94	1.3%
Number Decrease:	64	0.9%
Number MIncrease:	6	0.1%
Number MDecrease:	3	0.0%
Number No Change:	6889	97.6%
Number (A/M->P, M/I)	23	0.3%
Number (P->A/M, MD/D)	12	0.2%
Number (P->P, M/I)	63	0.9%
Number (P->P, MD/D)	41	0.6%
Number (A/M->A/M, M/I)	14	0.2%
Number (A/M->A/M, MD/D)	14	0.2%
Number (P->P, NC)	1468	20.8%
Number (A/M->A/M, NC)	4941	70.0%
Number Increase:		
(0<=SLR<1):	73	
(SLR>=1):	27	
(SLR>=2):	15	
(SLR>=3):	2	
(SLR>=4):	0	
Number Decrease:		
(0>=SLR>-1):	38	
(SLR<=-1):	29	
(SLR<=-2):	9	
(SLR<=-3):	5	
(SLR<=-4):	0	

Housekeeping Controls:

Probe Set	Sig(5)	Det(5)	Sig(M)	Det(M)	Sig(3)	Det(3)	Sig(all)	Sig(3/5)
HUMISGF3A/M97935	20.4	A	154.6	A	195.3	P	123.42	9.56
HUMRGE/M10098	10.6	A	30.7	A	34.9	A	25.42	3.29
HUMGAPDH/M33197	15088.9	P	14182.4	P	13387.5	P	14219.58	0.89
HSAC07/A00351	16005.2	P	26001.9	P	14967.8	P	18991.62	0.94
HUMTFRR/M11507	120.7	A	55.8	A	231.0	P	135.85	1.91
M27830	23.9	A	229.0	A	99.4	A	117.43	4.16

Spike Controls:

Probe Set	Sig(5)	Det(5)	Sig(M)	Det(M)	Sig(3)	Det(3)	Sig(all)	Sig(3/5)
BIOB	53.0	A	23.4	A	82.0	A	52.78	1.55
BIOC	481.1	P			293.9	P	387.48	0.61
BIODN	970.7	P			3778.4	P	2374.53	3.89
CREX	5405.5	P			7817.0	P	6611.24	1.45
DAPX	2146.1	P	2199.2	P	2794.3	P	2379.85	1.30
LYSX	2217.5	P	4278.5	P	3546.6	P	3347.55	1.60
PHEX	2495.5	P	3087.7	P	3865.5	P	3149.54	1.55
THRX	3749.3	P	2322.5	P	3804.9	P	3292.25	1.01
TRPNX	18.9	A	8.9	A	14.5	A	14.12	0.77

Other Controls:

Probe Set	Sig	Det
AFFX-BIOB-5_AT	53.0	A
AFFX-BIOB-M_AT	23.4	A
AFFX-BIOB-3_AT	82.0	A
AFFX-BIOB-5_ST	44.6	A
AFFX-BIOB-M_ST	39.8	A
AFFX-BIOB-3_ST	17.8	A

Figure 9.47

Expression report, comparison analysis, Statistical algorithm (page 2)

Expression Report Components

Statistical Expression Algorithm

Probe Pair Threshold	The minimum number of probe pairs a probe set must have in order for the probe set data to be included in the calculation of the report statistics
Alpha1	<p>The significance level for the detection p-value in an absolute analysis. Alpha1 is a user-modifiable parameter that is set in the Parameters tab of the Expression Analysis Settings (see Appendix D).</p> <p>If the probe set detection p-value $<$ alpha1, the call is present.</p>
Alpha2	<p>The second significance level for the detection p-value in an absolute analysis. Alpha2 is a user-modifiable parameter set in the Parameters tab of the Expression Analysis Settings (see Appendix D).</p> <p>If the probe set detection p-value \geq alpha2, the call is absent. If $\text{alpha1} \leq \text{detection p-value} < \text{alpha2}$, the call is marginal.</p>
Tau	Tau is a user-modifiable parameter that is set in the Parameters tab of the Expression Analysis Settings (see Appendix D). Ideally, tau should be set to a value that is a little larger than the median of the discrimination scores of the probe sets whose targets are absent to avoid false detected calls.
Noise (Raw Q)	The degree of pixel-to-pixel variation among the probe cells used to calculate the background (see the Appendix C).
Scale Factor (SF)	The scale factor specified in the Scaling tab of the Expression Analysis Settings dialog box or computed by the algorithm. (See Appendix D).

TGT Value	The user-specified target signal for scaling of the experiment probe array. The target signal is set in the Scaling tab of the Expression Analysis Settings dialog box (see Appendix D).
Norm Factor (NF)	The normalization factor specified in the Normalization tab of the Expression Analysis Settings dialog box or computed by the algorithm (see Appendix D).
Gamma1H	The small significance level for change calls at high intensities. Gamma1H is a user-modifiable parameter that is set in the Parameters tab of the Expression Analysis Settings (see Appendix D).
Gamma2H	The large significance level for change calls at high intensities. Gamma2H is a user-modifiable parameter that is set in the Parameters tab of the Expression Analysis Settings (see Appendix D).
Gamma1L	The small significance level for change calls at low intensities. Gamma2L is a user-modifiable parameter that is set in the Parameters tab of the Expression Analysis Settings (see Appendix D).
Gamma2L	The large significance level for change calls at low intensities. Gamma2L is a user-modifiable parameter that is set in the Parameters tab of the Expression Analysis Settings (see Appendix D).
Perturbation	A user-modifiable expression algorithm parameter that is set in the parameters tab of the Expression Analysis Settings (see Appendix D). Perturbation influences the p-value computed for a probe set in a comparison analysis.
Baseline Noise (Raw Q)	The degree of pixel-to-pixel variation among the probe cells used to calculate the background in the baseline probe array. (See Appendix C.)
Baseline Scale Factor (SF)	The scale factor specified for the baseline probe array in the Scaling tab of the Expression Analysis Settings dialog box or computed by the algorithm. (See Appendix D.)

Background	Minimum, maximum, average, and standard deviation of the background intensity calculated for the probe array.
Noise	The minimum, maximum, average, and standard deviation of the noise calculated for the probe array.
Corner+	The average cell intensity for the sense probe cells used in the grid alignment process.
Corner-	The average cell intensity of the antisense probe cells used in the grid alignment process.
Central+	The average cell intensity for the nine probe cells that compose the cross at the center of a sense probe array.
Central-	The average cell intensity for the nine probe cells that compose the cross at the center of an antisense probe array.
Total Probe Sets	The number of probe sets on the array that exceed the probe pair threshold and are not called <i>No Call</i> .
Average Signal	The average signal for all probe sets that exceed the probe pair threshold and are not called <i>No Call</i> .
Controls	The expression report includes the signal and call data for the probe sets that correspond to the housekeeping or spike control transcripts. Separate signal and call data are reported for the probe pairs specific to the 5', middle (M'), and 3' regions of the control transcripts.
Sig(all)	The average signal for all control probe sets.
Sig(3'/5')	For a probe set, $\text{Sig}(3')/\text{Sig}(5')$.

Expression Report Settings

The expression report has user-modifiable settings.

1. To view the user-modifiable settings, click **Report Settings** in the shortcut bar, then click **Expression Report** .
⇒ The Expression Report dialog box appears ([Figure 9.48](#)).

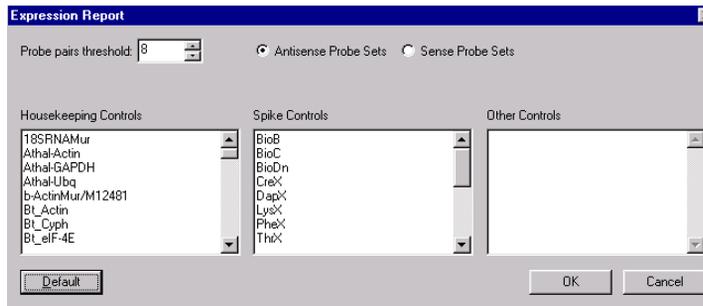


Figure 9.48
Expression Report dialog box

Probe Pairs Threshold	The minimum number of probe pairs a probe set must have in order for the probe set data to be included in the calculation of the report statistics.
Antisense Probe Sets	Choose this option for probe arrays that have antisense control probes. (Refer to the probe array product insert.)
Sense Probe Sets	Choose this option for probe arrays that have sense control probes. (Refer to the probe array package insert.)

Housekeeping Controls

Housekeeping controls are housekeeping gene transcripts that are usually constitutively expressed in the sample. The transcripts serve as endogenous controls and are useful for monitoring the quality of the target.

The report includes the signal (Sig) for probe sets that are designed to be specific to the 5', middle, or 3' portion of the transcript. These data are informative about the reverse transcription and IVT steps in sample preparation.

Differences greater than three fold between the Sig 3' and Sig 5' (Sig 3'/5') for a housekeeping control indicate the target may be degraded and should be prepared again.

- To delete a housekeeping control from the expression report:
 1. Double-click the probe set name.
 2. Right-click probe set and select **Delete** from the shortcut menu.

Spike Controls

Spike controls are transcripts that are added to or *spiked* in the sample. These are exogenous transcripts that are useful for monitoring assay procedures such as hybridization, washing, and staining.

- To delete a spike control from the expression report:
 1. Double-click the probe set name.
 2. Right-click probe set and select **Delete** from the shortcut menu.

Other Controls

You can choose any probe set on the array as a control. Enter the probe set name in the **Other Controls** box. The expression report includes the signal and detection call for user-specified controls.

Expression Report View Options

Find Feature

The Find feature searches the report.

1. Click the **Find** toolbar button  or select **Edit** → **Find** from the menu bar.
⇒ The Find dialog box appears (Figure 9.49).

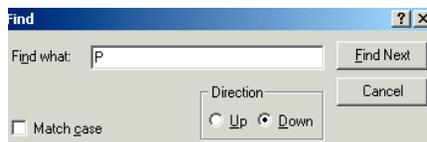


Figure 9.49
Find dialog box

2. Enter the desired text for the search (up to 256 alphanumeric characters including spaces) in the **Find what** box, then click **Find Next** to view the first search result.
3. Click **Find Next** or the **Find Next** toolbar button  to view each successive search result.

Editing an Expression Report

Edit commands for the expression report are available in:

- the Edit menu bar
- report window toolbar
- a shortcut menu (right-click the report)

Replace Feature

The replace feature performs a text search and replaces the search item with specified text.

1. Select **Edit** → **Replace** from the menu bar.
⇒ The Replace dialog box appears (Figure 9.50).



Figure 9.50
Replace dialog box

2. Enter the text for the search in the **Find what** box and enter the replacement text in the **Replace with** box.
3. Click **Find Next** to begin the search, then click **Replace** to replace the search item with the specified text. Repeat to find and replace successive instances of the search item. Alternatively, click **Replace All** to find and replace all occurrences of the search item in the report.

Saving an Expression Report

When a report for an analysis output file (*.chp) is generated, it overwrites a report previously generated for the *.chp. To prevent a previous report from being overwritten, save it under a new name.

1. To open the report, click the **Open** toolbar button .
⇒ This displays the Open dialog box (**Figure 9.51**).



Figure 9.51
Open dialog box

2. Double-click the report name.
⇒ This displays the report.
3. Select **File** → **Save As** from the menu bar.
⇒ The Save As dialog box appears (**Figure 9.52**).

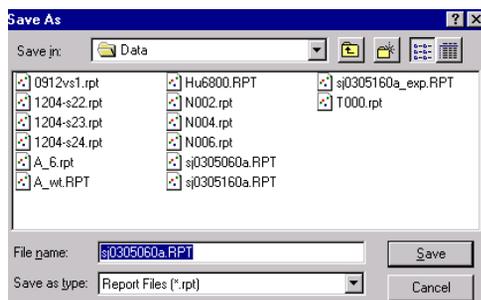
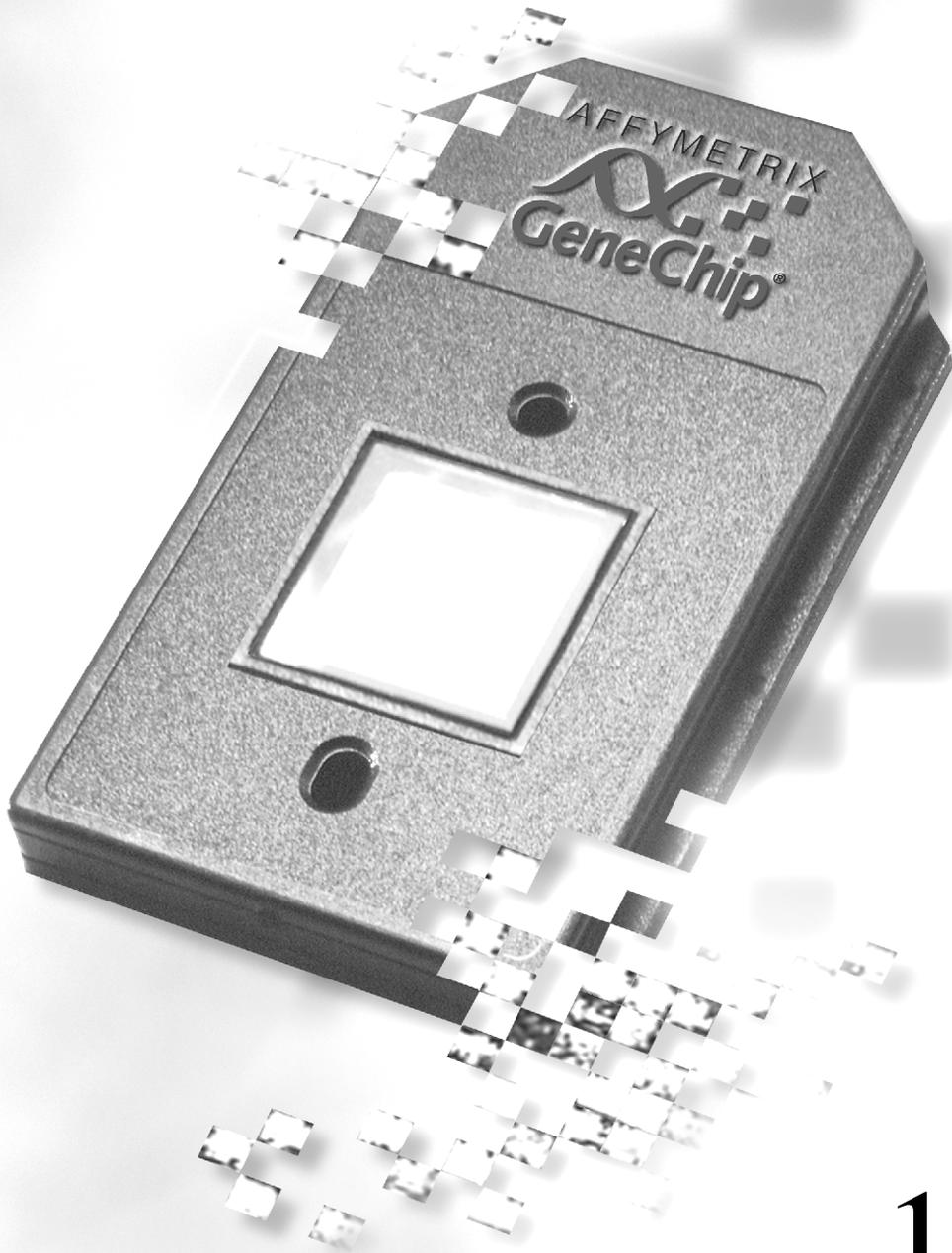


Figure 9.52
Save As dialog box

4. Click the report to be saved, then enter a new name for the report in the **File name** box.
5. Click **Save**.



10

Chapter 10





Mutation & Polymorphism Analysis

GeneChip® probe arrays for mutation and polymorphism analysis interrogate the target to identify point mutations, deletions, or polymorphisms relative to known bases.

This chapter describes how to:

- run a polymorphism & mutation analysis
- view the analysis output (*.chp) in graphical or tabular format in the Sequence Analysis window (SAW)
- generate a report

Overview

For a mutation and polymorphism analysis, standard and alternative tiling strategies specify probe sets that examine each base position of interest. The reference bases for the standard tile span the entire set of bases of interest. Alternative tiles anticipate specific mutations and polymorphisms and specify probes that interrogate selected regions of the sequence where known mutations or polymorphisms occur.

Microarray Suite aligns the bases of the probes that hybridize to complementary bases in the target, calls a consensus base call at each position of the target, and identifies changes from the reference bases.

After Microarray Suite analyzes the hybridization intensity data, the Sequence Analysis Window (SAW) displays the analysis output or *chip* file (*.chp). If the SAW is already open, the results are added to the open window and it may be necessary to scroll down to see the newly added results.

A second toolbar (see Appendix J) and additional menu bar commands are available in the SAW. In the SAW, you can:

- compare reference and consensus base calls and automatically highlight mismatches, amino acid changes, or silent mutations
- view graphs of the data
- search for nucleotides and amino acids or move between user-specified nucleotide or amino acid positions of interest
- import reference bases from an ASCII text file

Mutation & Polymorphism Analysis Algorithms

Microarray Suite selects the probe array calling algorithm for an experiment based on the probe array type specified during experiment setup ([Table 10.1](#)).

Table 10.1
Probe array analysis algorithms for mutation and polymorphism analysis

Mutation & Polymorphism Analysis		
Probe Array Analysis Algorithm	GeneChip® Probe Array	Regions Analyzed
Mixture Detection	p53	Exons 2-11 of the human p53 gene.
Consensus Rules	HIV PRT <i>Plus</i>	The 18 bp upstream from the protease gene, the entire protease gene, and codons 1-400 of the RT gene of HIV-1.

Mixture Detection Algorithm

The Mixture Detection algorithm analyzes hybridization intensity data from the GeneChip® p53 probe array to identify a single mutant base in a background of wildtype base at each position in exons 2-11 of the human p53 tumor suppressor gene and performs splice junction analyses on the flanking intron sequences.

The algorithm incorporates both standard and alternative tiling information to specify probe sets that interrogate the target for single base deletions,

specific insertions, and multi-base deletions. A given position may be analyzed by one or more (up to 14) probe sets. The algorithm does not detect multiple mutations at a single base position or quantitate the amount of mutant base relative to wildtype base.

The algorithm relies on the Mixture Detection analysis settings to derive biologically meaningful results from the hybridization intensity data. The Mixture Detection analysis settings are user-modifiable (see Appendix G). The default analysis settings were empirically optimized through extensive testing at Affymetrix.

Viewing Mixture Detection Algorithm Settings

1. Click **Analysis Settings** in the shortcut bar, then click **Mixture Detection Algorithm** .
⇒ The Mixture Detection Analysis Settings dialog box appears ([Figure 10.1](#)).

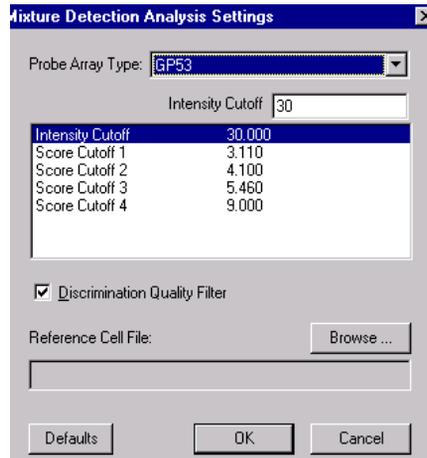


Figure 10.1
Mixture Detection Analysis Settings

2. Confirm the defaults or change the settings as described in Appendix G.

Consensus Rules Algorithm

The Consensus Rules algorithm analyzes hybridization intensity data from the GeneChip® HIV PRT *Plus* probe array to determine base changes in the target relative to known bases.

The HIV PRT *Plus* assay examines the 18 base pairs upstream from the protease gene, the entire protease gene, and codons 1-400 of the reverse transcriptase (RT) gene of HIV-1. The algorithm incorporates both standard and alternate tiling information to make a base call (IUPAC codes) at each target position interrogated.

The Consensus Rules algorithm has no user-modifiable analysis settings.

Running a Mutation & Polymorphism Analysis

Before analyzing a GeneChip® p53 probe array, you must specify a wildtype reference cell intensity file (*.cel). (See Appendix G or the GeneChip P53 probe array package insert for more information about the reference cell intensity file.)

If not analyzing a GeneChip p53 probe array, proceed to [step 5](#).

1. Click **Analysis Settings** in the shortcut bar, then click **Mixture**

Detection .

⇒ The Mixture Detection Analysis Settings dialog box appears ([Figure 10.2](#)).

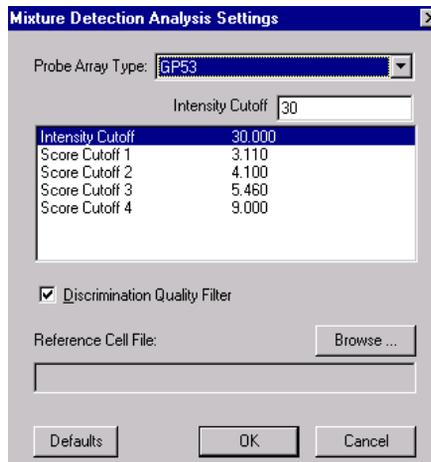


Figure 10.2
Mixture Detection Analysis Settings

2. Click **Browse** to open a directory of cell intensity (*.cel) files.
3. Double-click the p53 wildtype reference cell intensity file (*.cel).
⇒ The Mixture Detection Analysis Settings dialog box displays the reference cell intensity file (**Figure 10.3**).

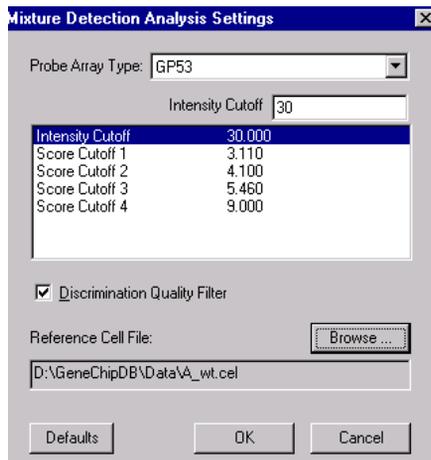


Figure 10.3
Mixture Detection Analysis Settings, p53 wildtype reference cell intensity file selected

4. Click **OK** to close the Mixture Detection Analysis Settings dialog box.
5. Do one of the following to analyze the cell file (*.cel) of interest:
 - In the data file tree, right-click the *.cel and select **Analyze** from the shortcut menu.
 - Select **File** → **Analysis** from the menu bar, then choose the *.cel of interest from the Analyze dialog box that appears.
 - If the *.cel image is displayed, click the **Analyze** toolbar button  or select **Run** → **Analysis** from the menu bar.

The status log displays:

- the name of the analysis output file (*.chp)
- the location of the *.chp
- a message indicating when the analysis is completed

If the status log is not displayed, click the **Status Log** toolbar button  in or select **View** → **Status Bar** from the menu bar.

When the analysis is finished, the Sequence Analysis window (SAW) displays the analysis output file (*.chp) (Figure 10.4). If the SAW is already open, the results are added to the open window and it may be necessary to use the scroll bars at the bottom and right side of the SAW to see the newly added results.

Sequence Analysis Window (SAW)

The Sequence Analysis Window (SAW) displays the tabular and graphic analysis output for mutation and polymorphism assays. The SAW automatically opens when you view an analysis output file (*.chp) for this type of probe array.

The SAW is divided into five areas or *windowpanes* (Figure 10.4). The windowpanes may be resized (see Appendix K). Use the scroll bars at the right side or bottom of the SAW to view the entire contents of the window.

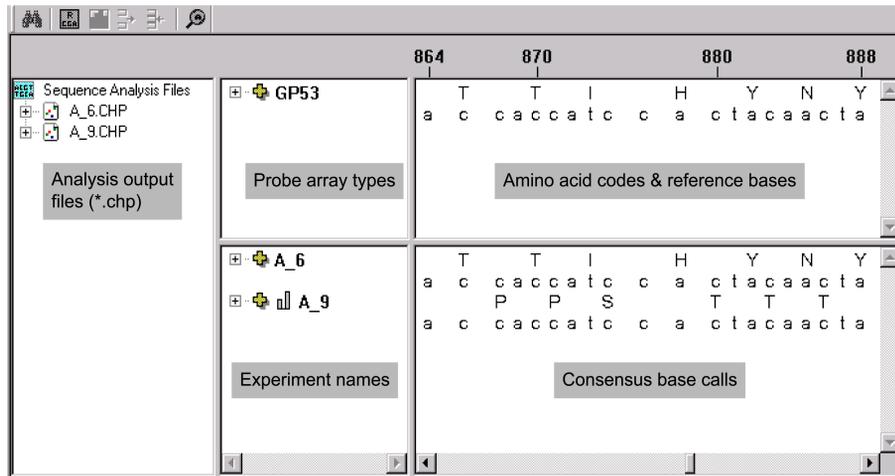


Figure 10.4
Sequence Analysis window (SAW)

Sequence Analysis Window View Options

Viewing Expanded Information

1. To expand the information in the SAW, click the plus symbol (+) to the left of a *.chp file name, probe array type, or experiment name.

The expanded file information (Figure 10.5) displays the analysis algorithm. You can also expand the algorithm to view any user-modifiable parameters.

2. To condense the information, click the minus symbol (-) to the left of the item.

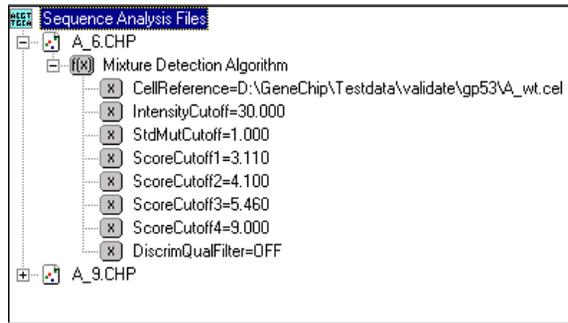


Figure 10.5

Left windowpane in the SAW, expanded file and algorithm information for GeneChip® p53 probe array

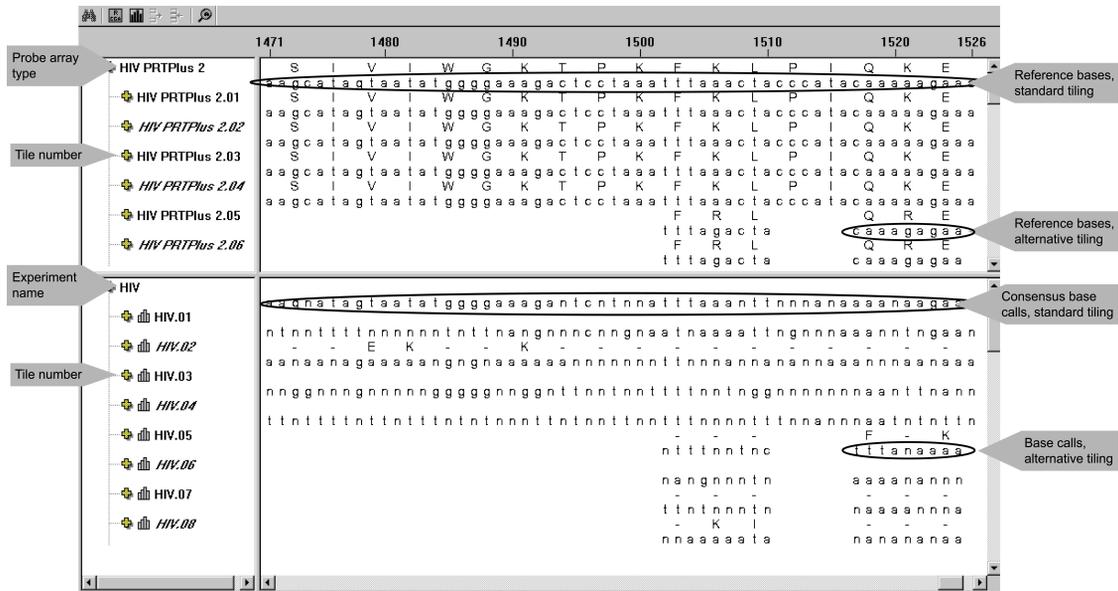


Figure 10.6

SAW, expanded display of probe array tiling and reference and consensus information for GeneChip® HIV PRT Plus 2 probe array

The expanded probe array type displays the tile numbers of the probe sets on the array and the corresponding reference bases (Figure 10.6).

GeneChip® probe arrays for mutation and polymorphism analysis contain both sense and antisense probes. A sense probe hybridizes to the sense (coding) strand of the target and an antisense probe hybridizes to the antisense (noncoding) strand of the target. The tile numbers of antisense probes are displayed in italics. The SAW displays all bases as sense strands (5'→3', left to right).

The expanded experiment name displays the tile numbers of the probe sets on the array and the corresponding called bases (Figure 10.6).

Viewing Nucleotides, Amino Acids, or Both

The SAW displays nucleotides or nucleotide mixtures (IUPAC codes, see Appendix L), amino acids, or both nucleotides and amino acids.

To toggle the view between these options, do either of the following:

- Click the **Toggle** toolbar button  until the desired format is displayed.
- Select **View** → **Amino Acids**, or **View** → **Nucleotides**, or **View** → **Nucleotides and Amino Acids** from the menu bar. Continue to select from the menu until the desired format is displayed.

The status bar at the bottom of the SAW displays information about a highlighted amino acid or nucleotide. It displays the codon number and gene name for a highlighted amino acid. When a base in a consensus sequence is highlighted, it displays the nucleotide number, tile number, codon number, and gene name. The hybridization intensity and background data are also displayed for called bases (Figure 10.7).

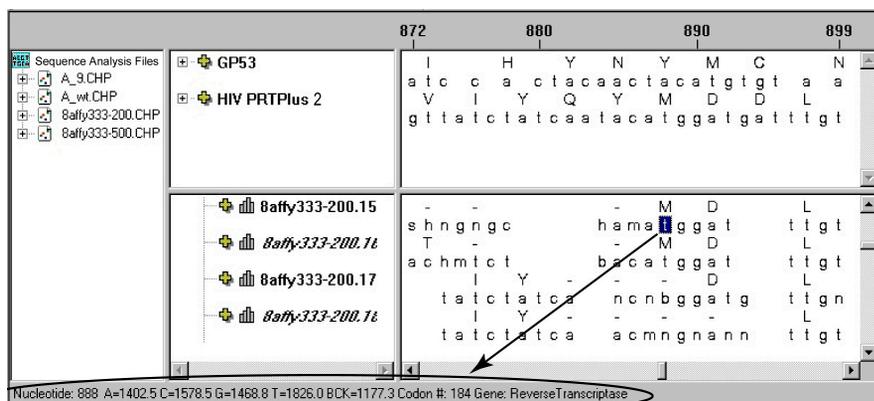


Figure 10.7
Information for a called base

Other possible characters in the SAW include:

- n The IUPAC ambiguity code (indicates any base A, C, G, or T) that denotes a position which could not be called.
- a/c, a/g, a/t, c/a, c/g, c/t, g/a, g/c, g/t, t/a, t/c, t/g A mixture of mutant (first letter) and wildtype (second letter) at a position in the Affymetrix p53 assay.
- A mixture of a deletion and wildtype base at a position in the p53 assay.

Viewing the Complement of a Selected Tile Number

Do one of the following to view the complementary bases:

- Double-click the plus sign  at the left of the tile number.
- Highlight a tile number, then select **View** → **Complement** from the menu bar.
- Highlight a tile number, right-click the windowpane and select **Complement** from the shortcut menu.

Comparing Nucleotide Data

You can automatically highlight nucleotides in the SAW according to the comparison categories listed in [Table 10.2](#).

Table 10.2
SAW, nucleotide comparison options

Nucleotide Comparison Option	Highlights...
Any Mismatch	all mismatches between bases
Non-Ambiguous Mismatch	mismatches resulting from unique base calls. Does not highlight mismatches resulting from an ambiguity code indicating several base possibilities for the position
All Ambiguities	all ambiguities
All Ambiguous Matches	ambiguous matches where one of the possibilities for the ambiguous position matches the reference sequence
All Ambiguous Mismatches	ambiguous mismatches where none of the possibilities for the ambiguous position matches the reference sequence
Unable to call	positions where the base was not called
Deletions	deleted base positions

TIP

Comparison options and corresponding highlight colors can be chosen to help compare reference and called bases. Highlight a general category in one color and a subcategory in another color (for example, highlight Mismatches red and Non-ambiguous Mismatches green). All tiles are compared to the top most selected tile.

Selecting a Nucleotide Comparison Option and Highlight Color

1. To select the probe arrays or tile numbers for the comparison:
 - a. Expand the reference and experiment probe array list (click the + sign next to the name).
 - b. Press and hold the **Ctrl** key while you click the probe arrays or tile numbers for the comparison ([Figure 10.8](#)).

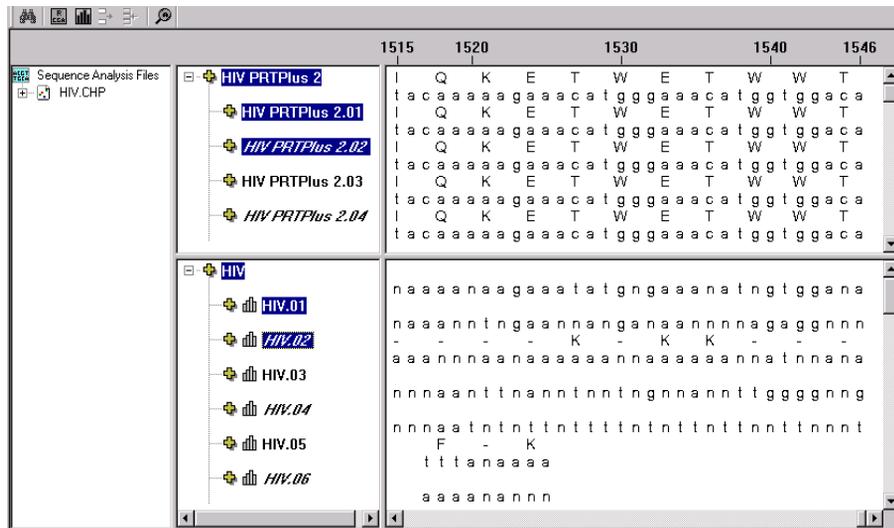


Figure 10.8
SAW, highlighted probe arrays and tiles

2. To open the Analysis Options dialog box, do one of the following:
 - Click the **Analysis Options** toolbar button .
 - Right-click a middle windowpane in the SAW and select **Options** from the shortcut menu.
 - Select **Analysis** → **Options** from the menu bar.

⇒ The Analysis Options dialog box appears and displays the Nucleotides tab (Figure 10.9).

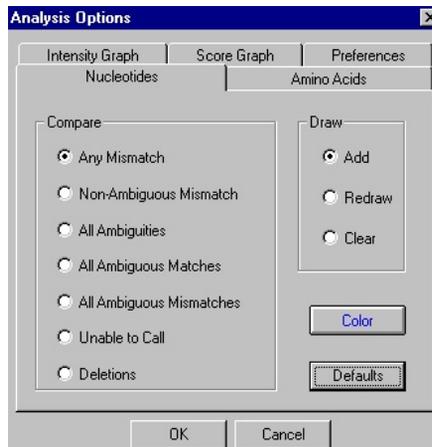


Figure 10.9
Analysis Options, Nucleotides tab

3. Choose a nucleotide compare option (see [Table 10.2](#)).
4. To specify a color for the **Compare** option:
 - a. Click **Color**.
⇒ This displays the color palette ([Figure 9.41](#)).

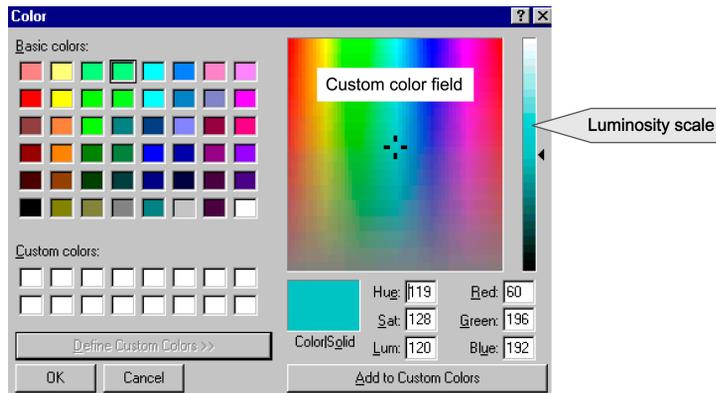


Figure 10.10
Color palette

Selecting an Amino Acid Comparison Option and Highlight Color

1. To select the probe arrays or tile numbers for the comparison:
 - a. Expand the reference and experiment probe array list (click the + sign next to the name)
 - b. Press and hold the **Ctrl** key while you click the probe arrays or tile numbers for the comparison (Figure 10.11)

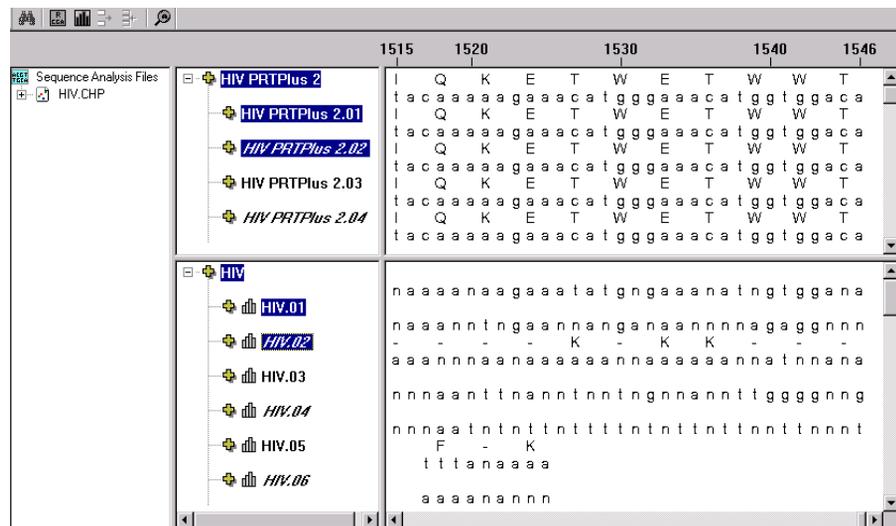


Figure 10.11
SAW, highlighted probe arrays and tiles

2. To open the Analysis Options dialog box, do one of the following:
 - click the **Analysis Options** toolbar button 
 - right-click a middle windowpane in the SAW and select **Options** from the shortcut menu
 - select **Analysis** → **Options** from the menu bar

⇒ The Analysis Options dialog box appears (Figure 10.12).

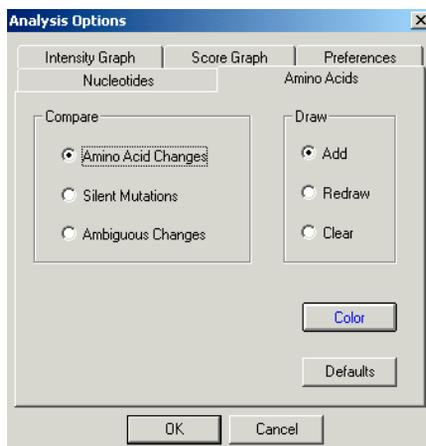


Figure 10.12
Analysis Options, Amino Acids tab

3. Click the Amino Acids tab, then choose a compare option (see [Table 10.2](#)).
4. To specify a color for the **Compare** option:
 - a. Click **Color**.
⇒ This displays the color palette ([Figure 10.13](#)).

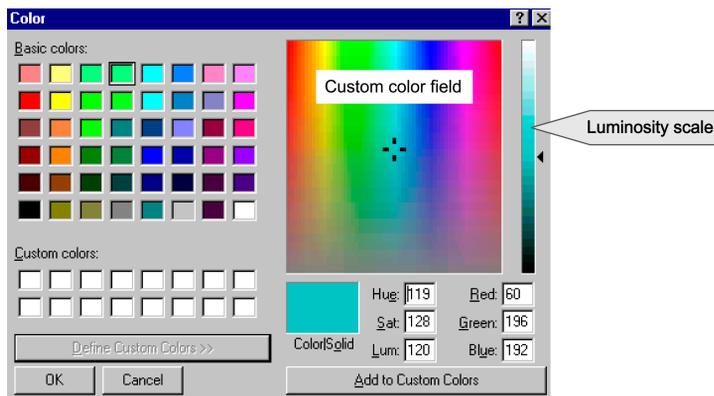


Figure 10.13
Color palette

- b. To select a predefined color, click one of the basic colors.
 - c. To define a custom color, click **Define Custom Colors**, then use the click-and-drag method to move the cross hairs in the custom color field. Adjust the color brightness using the luminosity scale to the right. When finished, click **Add to Custom Colors** to apply the color
 - d. Click **OK** to close the color palette.
 - ⇒ The **Color** button in the Nucleotides tab displays the new color choice. The color choices are saved on a per user basis.
- 5.** In the Analysis Options dialog box, choose a **Draw** option.
- Add** Adds the new comparison to the SAW without removing previous comparisons so that all of the compare option/highlight color combinations are displayed simultaneously in the selected tiles or rows.
- Redraw** Replaces the previous comparison with the new one and displays only the new compare option/highlight color combination in the selected tiles.
- Clear** Removes all amino acid highlights from selected tiles or rows.
- 6.** Repeat steps 3 through 5 to display another compare option in the selected tiles.

Hiding, Reordering, or Removing Sequences

Hiding Tiles, Bases, or Rows

- 1.** Highlight the tile number(s) to be hidden, then click the **Hide** toolbar button  in the SAW or select **View** → **Hide** from the menu bar.
- 2.** To unhide the hidden sequences, click the **Unhide** toolbar button  or select **View** → **Unhide All** from the menu bar.

Reordering a Probe Array Type or Experiment

1. Click and hold the probe array or experiment name.
2. Use the drag-and-drop method to move the probe array or experiment name to a new position in the list.

The mouse pointer is tagged with *actg* as you drag the selected probe array type or experiment name. A red triangle arrow indicates the new location when the mouse button is released (above or below the outlined experiment name or probe array type). For example, in [Figure 10.14](#) the experiment A6 will be moved from the last to the second position in the list.



Note

When a probe array type or experiment name is reordered, the corresponding tiles are also reordered; however, individual tile numbers of probe arrays cannot be reordered.

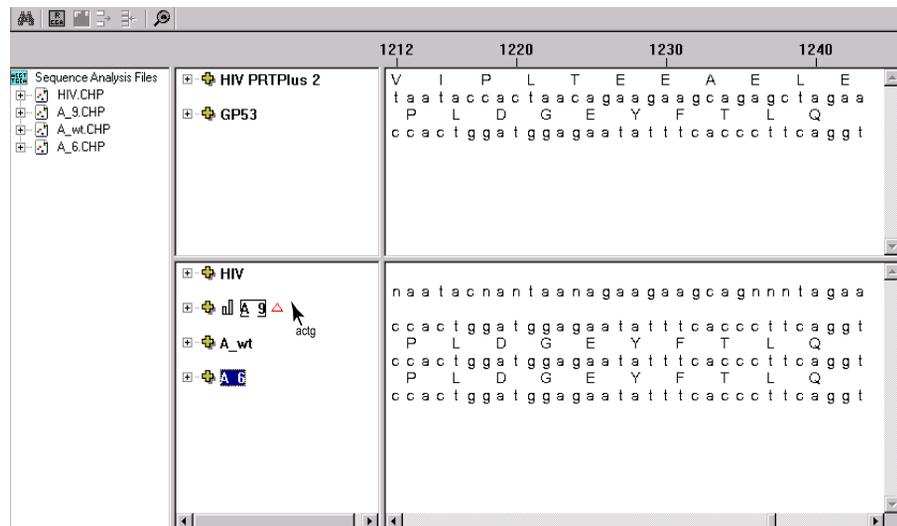


Figure 10.14 Drag a probe array type or experiment name to reorder sequence information in the SAW

Removing Data, Tiles, or Rows

You can remove all of the data, tiles, and rows of a particular probe array or experiment from the SAW (equivalent to closing the *.chp file).

1. Highlight the name of the probe array or *.chp file.
2. Select **Edit** → **Remove** from the menu bar.
3. Reopen the *.chp file to return the data, tiles, or rows to the SAW.

Alignment Bars

You can apply a vertical alignment bar to any nucleotide position(s) in the right SAW windowpanes ([Figure 10.15](#)).

1. To display an alignment bar, double-click the nucleotide position(s) of interest (in the upper or lower right windowpane).
2. Double-click an alignment bar to remove it.
3. To remove all alignment bars from the SAW, select **Edit** → **Clear Alignment Bars** from the menu bar.

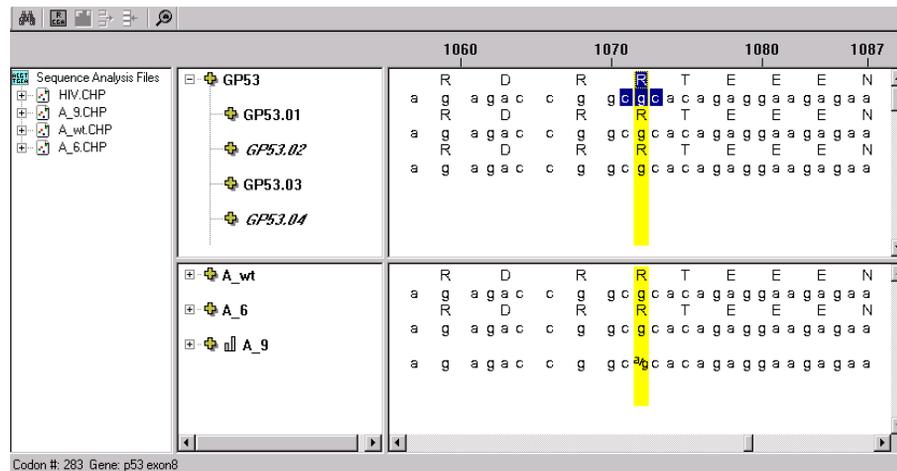


Figure 10.15
Alignment bar in the SAW

Mutation & Polymorphism Analysis Graphs

HIV PRT *Plus* Intensity Graph

The HIV PRT *Plus* intensity graphs display a bar graph of the hybridization intensity data for each base at each interrogated position (**Figure 10.16**).

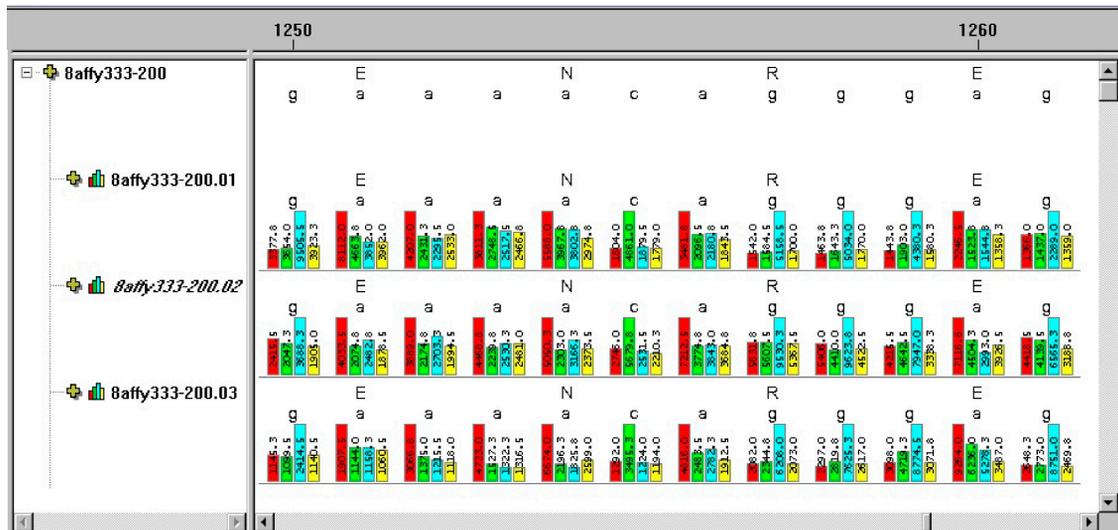


Figure 10.16
HIV PRT *Plus* intensity graphs

Plotting the HIV PRT *Plus* Intensity Graph

- To display (or hide) the intensity graphs for all tile numbers in the SAW, do either of the following:
 - Click the **Graphs** toolbar button .
 - Select **View** → **Intensity** from the menu bar.
- To display (or hide) the intensity graphs for a particular tile number in the SAW, double-click the **Intensity Graph** button  to the left of the tile number.
- To display (or hide) the intensity graph for a particular nucleotide in the SAW, right-click the nucleotide.

HIV PRT *Plus* Intensity Graph Options

You can change some display features of the intensity graph.

- Click the **Analysis Options** toolbar button . Alternatively, select **Analysis** → **Options** from the menu bar.
⇒ The Analysis Options dialog box appears (Figure 10.17).

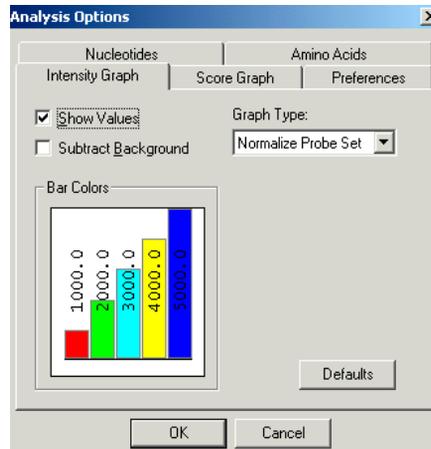


Figure 10.17
Analysis Options, Intensity Graph tab

2. Click the Intensity Graph tab.

Show Values

Choose this option to display the hybridization intensity value for each bar in the Intensity graph.

Subtract Background

Choose this option to display background-subtracted intensity data.

Graph Type

Normalize
Probe Set

For each probe set, the software selects the highest intensity value and displays it in the graph at the maximum bar height (when the background is not subtracted). The intensity bars for the remaining bases of the probe set are scaled relative to the highest intensity bar in the same probe set (Figure 10.18).

Log10 Probe
Set

The software selects the highest intensity value from all of the probe sets in a tile and displays it in the graph at the maximum bar height (when the background is not subtracted). The intensity bars for all of the remaining bases in the tile are scaled relative to the highest intensity bar for the tile. The data are plotted on a log scale (Figure 10.18).

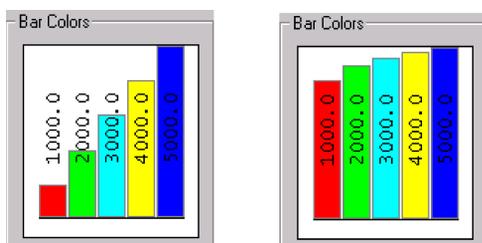


Figure 10.18

HIV PRT *Plus* intensity graphs, normalize probe set option (left) and log10 probe set option (right)

Defaults

Returns the intensity graph options to the factory set defaults.

Bar Colors

You can change the color of a bar in the intensity graph.

1. In the **Bar Colors** box, click a graph bar.
⇒ This displays the color palette (**Figure 10.19**).

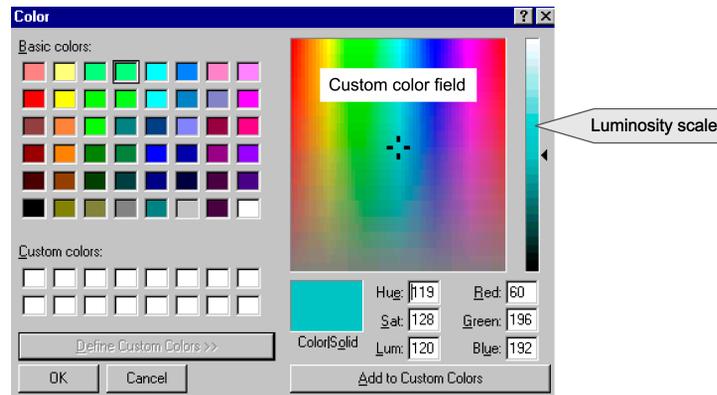


Figure 10.19
Color palette

2. To select a predefined color, click one of the basic colors.
3. To define a custom color:
 - a. Click **Define Custom Colors**, then use the click-and-drag method to move the cross hairs in the custom color field. Adjust the color brightness using the luminosity scale to the right. When finished, click **Add to Custom Colors** to apply the color.
4. Click **OK** to close the color palette.

The color choices are saved on a per user basis.

p53 Score Graph

The Mixture Detection algorithm computes a score for each position interrogated by the GeneChip® p53 probe array. The score is derived from all of the probe sets that interrogate the position. The score graph displays the sum of the scores for a given position (Figure 10.20).

If the probe sets that interrogate the position pass the Mixture Detection algorithm quality control filters and other criteria, the algorithm calls a putative mutant for the position. (See Appendix G for more information about the Mixture Detection algorithm.)

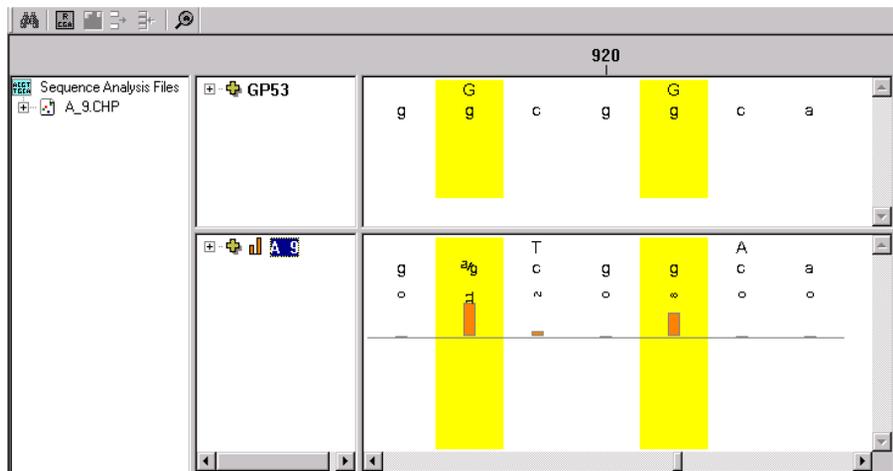


Figure 10.20
p53 score graph

In Figure 10.20, the score graph shows a mutant called at position 918 (a/g, mutant/wildtype), but not at position 921 (g) where the sum of the probe set scores is less than the score cutoff.

✓ Note

The range of score values depends on the number of probe sets that interrogate the position. For example, a score graph for a position analyzed by 14 probe sets has a larger maximum value than a score graph for a position analyzed by two probe sets. Score graphs are not normalized and may only be directly compared when the positions are analyzed by the same number of probe sets.

Plotting the p53 Score Graph

- To display (or hide) the p53 score graphs for all tile numbers in the SAW, do either of the following:
 - Click the **Intensity Graphs** toolbar button .
 - Select **View** → **Intensity** from the menu bar.
- To display (or hide) the Score graphs for a particular tile number in the SAW, double-click the **Score Graph** button  to the left of the tile number.

p53 Score Graph Options

You can change some display features of the p53 score graph.

- Click the **Analysis Options** toolbar button . Alternatively, select **Analysis** → **Options** from the menu bar.
 - ⇒ The Analysis Options dialog box appears (Figure 10.21).
- Click the Score Graph tab.

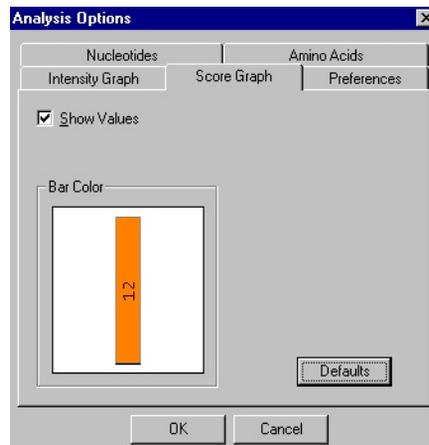


Figure 10.21
Analysis options, score graph

Show Values

Choose this option to display the score associated with each score graph bar.

Defaults

Returns the score graph options to the factory set defaults.

Bar Color

You can change the color of the score graph bar.

1. In the **Bar Color** box, click the graph bar.
⇒ This displays the color palette (**Figure 10.22**).

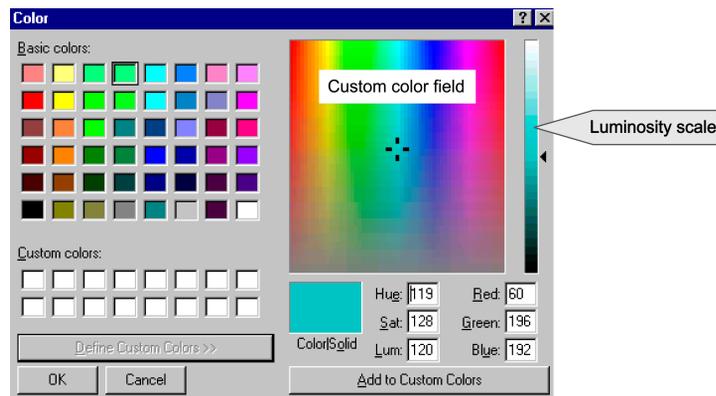


Figure 10.22
Color palette

2. To select a predefined color, click one of the basic colors.
3. To define a custom color:
 - a. Click **Define Custom Colors**, then use the click-and-drag method to move the cross hairs in the custom color field. Adjust the color brightness using the luminosity scale to the right. When finished, click **Add to Custom Colors** to apply the color.
4. Click **OK** to close the color palette.

The color choices are saved on a per user basis.

Navigating the Sequence Analysis Window (SAW)

Go To a Specific Nucleotide or Amino Acid

The Go To command locates and displays a user-specified nucleotide or amino acid in the reference bases of a probe array.

1. Right-click a right-hand windowpane in the SAW and select **Go To** from the shortcut menu. Alternatively, select **Edit** → **Go To** from the menu bar.

⇒ The Go To dialog box appears (Figure 10.23).

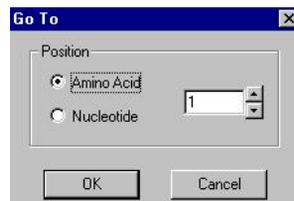


Figure 10.23
Go To dialog box

2. Choose the **Amino Acid** or **Nucleotide** option, enter the position number, then click **OK**.

⇒ The SAW highlights and displays the specified amino acid or nucleotide in the reference bases for the probe array.

Finding Specific Nucleotides or Amino Acids

The Find command locates and displays:

- every highlighted nucleotide or amino acid
- each nucleotide in a user-defined list of nucleotide positions

Moving Between Highlighted Positions

The Find command locates and displays each nucleotide or amino acid that is highlighted according to a comparison option (specified in the Analysis Options dialog box). (See the section *Selecting a Nucleotide Comparison Option and Highlight Color*, on page 219 or *Selecting an Amino Acid Comparison Option and Highlight Color*, on page 223.)

1. To start a search, do either of the following:
 - Click the **Find** toolbar button .
 - Right-click a right hand windowpane in the SAW and select **Find** from the shortcut menu or select **Edit** → **Find** from the menu bar.
- ⇒ The Find dialog box appears ([Figure 10.24](#)).

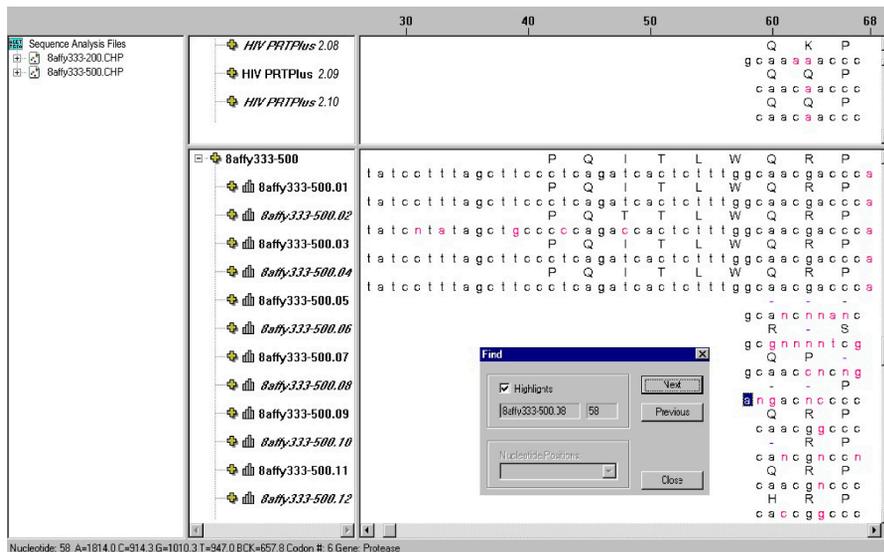


Figure 10.24
SAW, Find dialog box, highlights option

2. In the Find dialog box, choose the **Highlights** option, then click **Next** or **Enter** to view successive highlighted positions.
3. Click **Previous** to view preceding highlighted positions.

Moving Between User-Specified Nucleotide Positions

The Find command locates and displays user-specified nucleotide positions in the reference bases.

1. To start a search, do either of the following:
 - Click the **Find** toolbar button .
 - Right-click a right hand windowpane in the SAW and select **Find** from the shortcut menu or select **Edit** → **Find** from the menu bar.

⇒ The Find dialog box appears (Figure 10.24).
2. Make sure the **Highlights** option is unchecked so that the **Nucleotide Positions** box is available.
3. In the **Nucleotide Positions** box, enter the nucleotide positions of interest following the example format: 15 22 99 101 145 (Figure 10.25).

Up to 255 characters including spaces may be entered. A list of previously specified positions may be selected from the **Nucleotide Positions** drop-down list.

After a list is specified, it is automatically added to the drop-down list. Up to eight of the most recent lists are available in the **Nucleotide Positions** drop-down list.
4. Click **Next** to view successive positions.

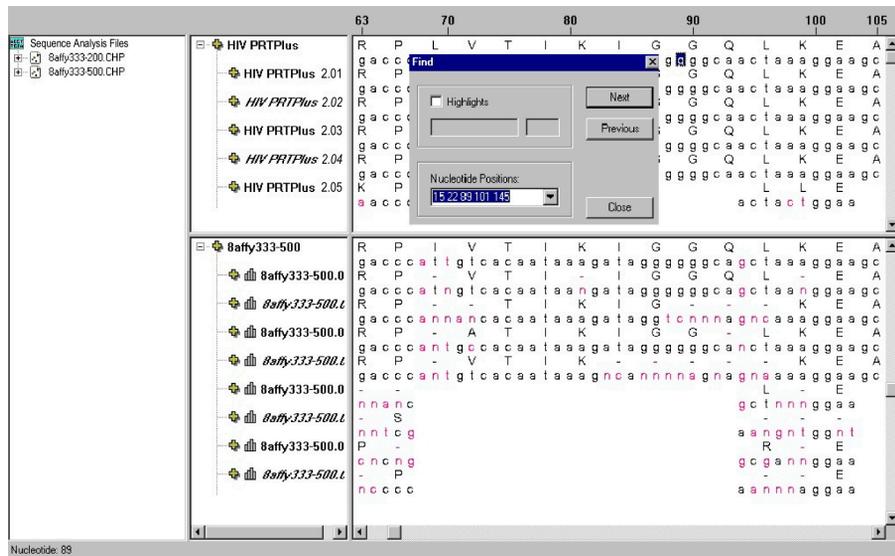


Figure 10.25
SAW, Find dialog box, nucleotide position option

Sequence Analysis Window (SAW) Display Options

You can change some display features of the SAW may be changed.

1. Click the **Analysis Options** toolbar button  in the SAW. Alternatively, select **Analysis** → **Options** from the menu bar.
⇒ The Analysis Options dialog box appears (**Figure 10.26**).
2. Click the Preferences tab.

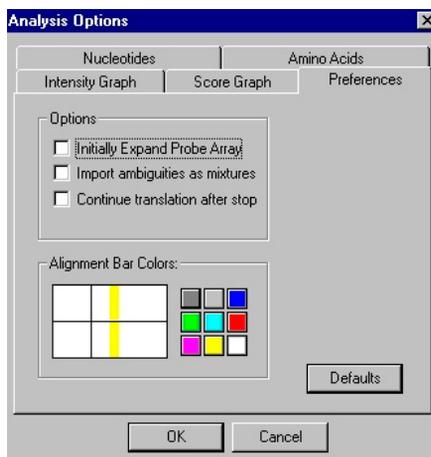


Figure 10.26
Analysis Options, Preferences tab

Options

Initially Expand Probe Array

Select this option to display expanded probe array or experiment information when an output analysis file (*.chp) is opened in the SAW. All of the tile numbers and corresponding bases for probe arrays and experiments are displayed in the SAW.

Importing Ambiguities as Mixtures

If this option is selected, an ambiguity code in an imported text file is converted to a mixture. M is converted to A/C, R is converted to A/G, and W is converted to W/T. The ambiguities V, H, D, B, or N are not converted.

Continuing Translation After Stop

Select this option to continue to display amino acids after a stop codon (TGA, TAA, or TAG) is encountered in the called bases.

Alignment Bar Colors

This option specifies the color of the vertical bar that highlights an amino acid or nucleotide position in the right panes of the SAW. Click a color swatch to choose a different color for the alignment bar.

Defaults

Returns the Preference tab settings to the factory set defaults.

Editing Consensus Base Calls

You can edit the nucleotides of consensus base calls.

1. Click the **Toggle** button until the SAW displays only nucleotide bases or amino acid and nucleotide bases information. Alternatively, select **View** → **Nucleotides** from the menu bar.
2. Click the nucleotide you wish to edit ([Figure 10.27](#)).

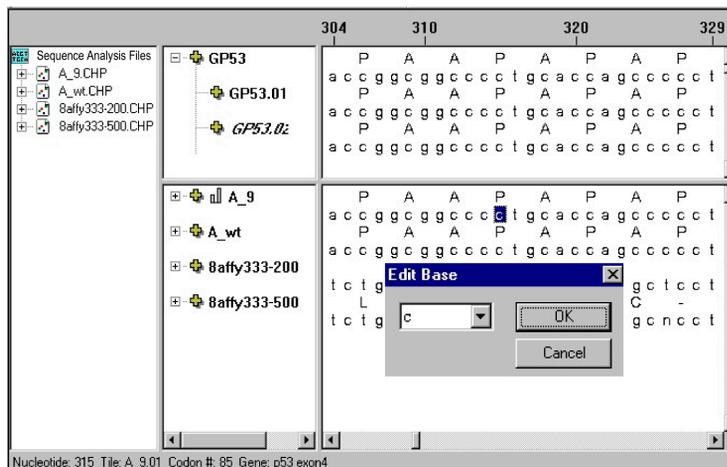


Figure 10.27
SAW, Edit Base dialog box

3. Select **Edit** → **Edit Base** from the menu bar.
⇒ The Edit Base dialog box appears ([Figure 10.27](#)).

4. Choose a nucleotide from the drop-down list to replace the highlighted consensus nucleotide in the SAW, then click **OK**.
⇒ The SAW displays the edited base as an upper case letter to indicate it has been revised.
5. To restore all edited bases to their original call, close the *.chp file without saving changes:
 - a. Highlight the experiment name in the SAW.
 - b. Select **Edit** → **Remove** from the menu bar.
 - c. Select **No** when prompted to save file changes.

The SAW displays the original consensus bases when the file is reopened.
6. To restore a single revised base to the original call:
 - a. Right-click the base of interest.
 - b. Click **Undo** → **Edit Base** in the shortcut menu.
7. To restore several bases to their original call:
 - a. Press and hold the **Ctrl** key while you click the bases of interest
 - b. Select **Edit** → **Undo Edit** from the menu bar. Alternatively, right-click a selected base, then click **Undo Edit** in the shortcut menu.

Importing Files

You can import reference nucleotide data (text format files (ASCII) that use standard IUPAC nucleotide abbreviations) for use as a reference in mutation and polymorphism assays. (See Appendix L for more information about IUPAC nucleotide abbreviations.) The SAW displays imported data as a sense strand and applies a + designation.

Most of the Edit and View commands as well as the Analysis Options are available for imported data. Intensity graphs are not available for imported files since they have no associated intensity data.

1. To import nucleotide data, select **File** → **Import** from the menu bar.
⇒ The Import dialog box appears ([Figure 10.28](#)).

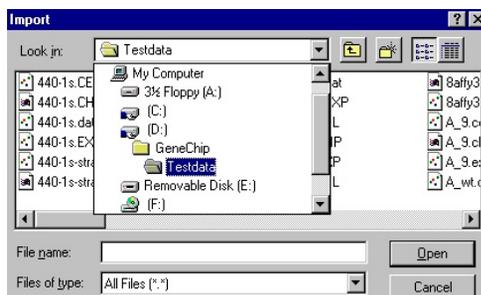


Figure 10.28
Import dialog box

2. Click the **Look in** drop-down list, then choose the disk drive that contains the file for import.
3. Open the directory containing the file and double-click the desired file name (*.txt).
4. Click **OK** when the **File name** field displays the desired file.

The imported sequence is added to the reference (upper) panes of the SAW. It may be necessary to scroll down to view the newly added sequence information if other files are already open in the SAW.

Mutation & Polymorphism Analysis Reports

p53 Report

The p53 report summarizes GeneChip® p53 probe array data and assay results from the analysis output file (*.chp). Use Microarray Suite or a word processing software to open, display, edit, print, or save a report.

Generating the p53 Report

1. In the data file tree, right-click the desired analysis output file (*.chp) and select **Report** from the shortcut menu.
⇒ The Select Report dialog box appears ([Figure 10.29](#)).

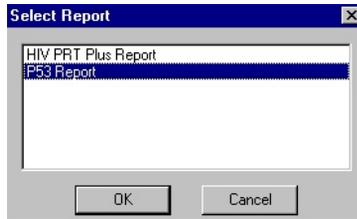


Figure 10.29
Select Report dialog box

- 2.** Choose **P53 Report** and click **OK** to generate the report (**Figure 10.30**).
⇒ The report is displayed in the main display area.
- 3.** To best view the P53 report:
 - a.** Use Arial font (10 point) (select **View** → **Set Font** from the menu bar).
 - b.** Set the tab stops set at 14 (select **View** → **Set Tab Stops** from the menu bar).

Filter Type: Mutations Only
 Date: 10:23AM 07/25/2001

Filename: A_9.CHP
 Probe Array Type: GP53
 Algorithm: Mixture Detection

SampleID	Location	AA Change	Codon Change	Score	SAW Reference
A_9	p53 intron1		g->c	11	4
A_9	p53 exon4	P75L	g->g cct->ctt	0,10,0	284,285,286
A_9	p53 exon4	A78P	cct->cct gca->cca	4,0,0	293,294,295
A_9	p53 exon5	R156I	gca->gca cgc->atc	12,14,0	561,562,563
A_9	p53 exon5	V157D	cgc->cgc gtc->gac	0,9,0	564,565,566
A_9	p53 exon5	M160I	gtc->gtc atg->atc	0,0,19	573,574,575
A_9	p53 exon6		atg->atg g->-	7	817
A_9	p53 exon6	P222H	ccg->cac ccg->cgc	0,12,0	830,831,832
A_9	p53 exon7	C229S	tgt->tca tgt->gta	9,12,0	861,862,864
A_9	p53 exon7	G244T	ggc->acg ggc->gcg	11,2,0	918,919,920
A_9	p53 exon7		t->c t->t	9	956
A_9	p53 exon8		g->-	6	1027
A_9	p53 exon8		g->t g->g	10	1051
A_9	p53 exon8		g->a g->g	10	1072
A_9	p53 exon8		g->c g->g	11	1119

Figure 10.30
 p53 report

Sample ID	The name of the analysis output file (*.chp).
Location	The location of the mutation on the p53 gene.
AA Change	The amino acid change resulting from the codon change (for example, at p53 exon 4, P75L indicates proline changed to leucine at amino acid position 75). No amino acid change results from a base change that occurs in an intron or after a stop codon.
Codon Change	The codon change resulting from the base mutation(s) or change(s) in intron regions or after a stop codon. Two entries for the same location of the p53 gene (for example, cct ctt and cct cct at location p53 exon 4) indicate a mixture of mutant and wildtype base at that location. A dash sign (-) indicates a single base deletion.
Score	The probe set score(s) for a mutation(s) at a given nucleotide location(s). The score is a sum of the variables that increase in value with increasing fractions of mutant base in the target. The score value increases as the amount of mutant signal increases.
SAW Reference	The nucleotide position of a single base mutation or deletion or the three nucleotide positions of a codon that contains one or more mutated bases or deletions.

Viewing User Modifiable Report Settings

The p53 report has user-modifiable settings.

1. Click **Report Settings** in the shortcut bar, then click **p53 Report** .
⇒ The p53 Report dialog box appears ([Figure 10.31](#)).

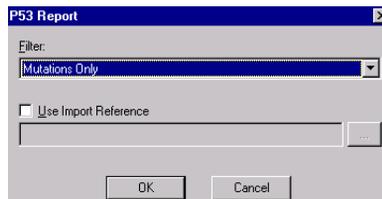


Figure 10.31
P53 Report dialog box

2. Choose the display option for the report (mutations only, silent mutations only, amino acid changes, or mixtures only) from the **Filter** drop-down list.

Selecting an Imported Reference File

1. To use a different wildtype reference file for the report other than the cell intensity file (*.cel) used in the analysis, choose the **Use Import Reference** option.
2. Click the **Browse** button .
⇒ This displays the Open dialog box and the available reference text files (*.txt).

(See the preceding section *Importing Files* for information about importing nucleotide data files into Microarray Suite.)
3. Select a reference file (*.txt) from the Open dialog box, then click **Open**.
⇒ The **Use Import Reference** field in the p53 Report dialog box displays the selected file ([Figure 10.32](#)).

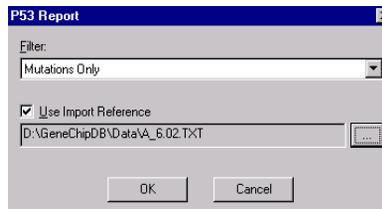


Figure 10.32
p53 Report dialog box, imported reference file selected

4. Click **OK** to close the p53 Report dialog box.

HIV PRT *Plus* Report

The HIV PRT *Plus* Report summarizes GeneChip® HIV PRT *Plus* probe array data and assay results derived from the analysis output file (*.chp) (Figure 10.34).

Use Microarray Suite or a word processing software to open, display, edit, print, or save a report.

Generating the HIV PRT *Plus* Report

1. In the data file tree, right-click the analysis output file (*.chp) for the report.
2. Select **Report** from the shortcut menu.
⇒ The Select Report dialog box appears (Figure 10.33).

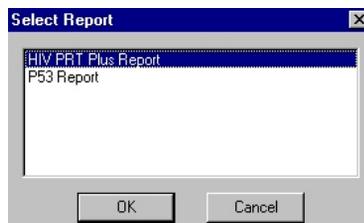


Figure 10.33
Select Report dialog box

- 3.** Choose **HIV PRT Plus Report** and click the **OK** button.
⇒ This generates the report and displays it in the main display area (**Figure 10.34**).
- 4.** To best view the HIV PRT Plus report:
 - a.** Use Arial font (10 point) (select **View** → **Set Font** from the menu bar).
 - b.** Set the tab stops set at 14 (select **View** → **Set Tab Stops** from the menu bar).

Report Type: HIV PRT Plus Report
Date: 10:16AM 07/25/2001

Filename: HIV.CHP
Probe Array Type: HIV PRTPlus 2
Algorithm: Rules
Filter Type: All Codon Changes

SampleID	Gene	AAChange	CodonChange	SAW Reference
HIV	Protease	P1-	cct->ntt	41
HIV	Protease	Q2-	cag->nag	44
HIV	Protease	I3F	atc->ttt	47
HIV	Protease	T4-	act->ant	50
HIV	Protease	L5-	ctt->ntt	53
HIV	Protease	W6-	tgg->tng	56
HIV	Protease	Q7-	caa->naa	59
HIV	Protease	R8-	cga->nga	62
HIV	Protease	P9-	ccc->anc	65
HIV	Protease	L10V	ctc->gtn	68
HIV	Protease	V11-	gtc->nan	71
HIV	Protease	T12-	aca->atn	74
HIV	Protease	K14-	aag->tng	80
HIV	Protease	I15	ata->att	83
HIV	Protease	G17R	ggg->aga	89
HIV	Protease	K20	aag->aaa	98
HIV	Protease	A22-	gct->nnt	104
HIV	Protease	L23I	cta->ata	107
HIV	Protease	T26-	aca->ana	116
HIV	Protease	A28-	gca->gna	122
HIV	Protease	D30E	gat->gaa	128
HIV	Protease	T31	aca->act	131
HIV	Protease	V32E	gta->gaa	134

Figure 10.34
HIV PRT Plus Report

Sample ID	The name of the analysis output file (*.chp).
Gene	The HIV-1 gene that carries the mutation.
AA Change	The amino acid change resulting from the codon change (for example, I3T indicates isoleucine changed to threonine at amino acid position 3). An asterisk (*) indicates a stop codon.
Codon Change	The codon change resulting from the base mutation.
SAW Reference	The nucleotide position of the base mutation.

Viewing User Modifiable Report Settings

The HIV PRT Plus report has user-modifiable settings.

1. Click **Report Settings** in the shortcut bar, then click **HIV PRT Plus Report** .
⇒ The HIV PRT Plus Report dialog box appears ([Figure 10.35](#)).

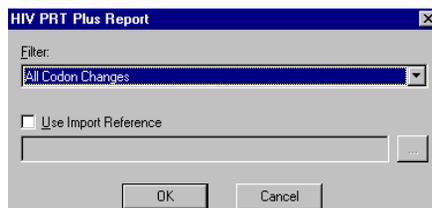


Figure 10.35
HIV PRT Plus Report dialog box

2. Choose the display option for the report (all codon changes, only silent mutations, or only amino acid changes) from the **Filter** drop-down list.

Selecting an Imported Reference File

You may choose a different reference file for the report other than the probe array reference displayed in the SAW.

1. In the HIV PRT Plus Report dialog box, choose the **Use Import Reference** option.
2. Click the **Browse** button .
⇒ This displays the Open dialog box and the available reference text files (*.txt).

(See the preceding section *Importing Files* for information about importing nucleotide data files into Microarray Suite.)
3. Select a reference file (*.txt) from the Open dialog box, then click **Open**.
⇒ The **Use Import Reference** field in the HIV PRT Plus Report dialog box displays the selected file ([Figure 10.36](#)).

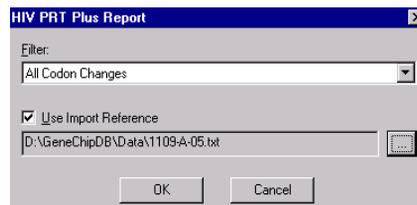
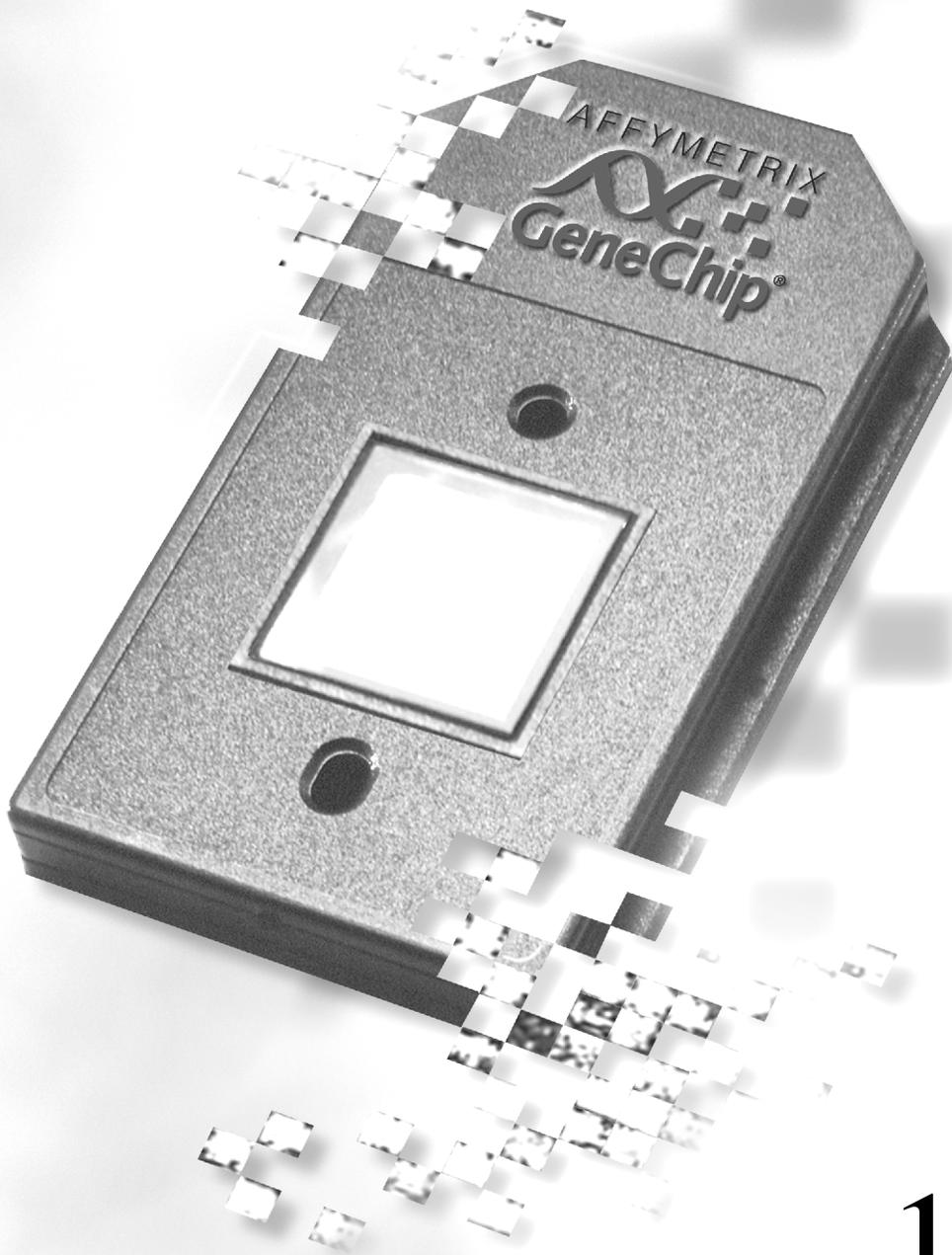


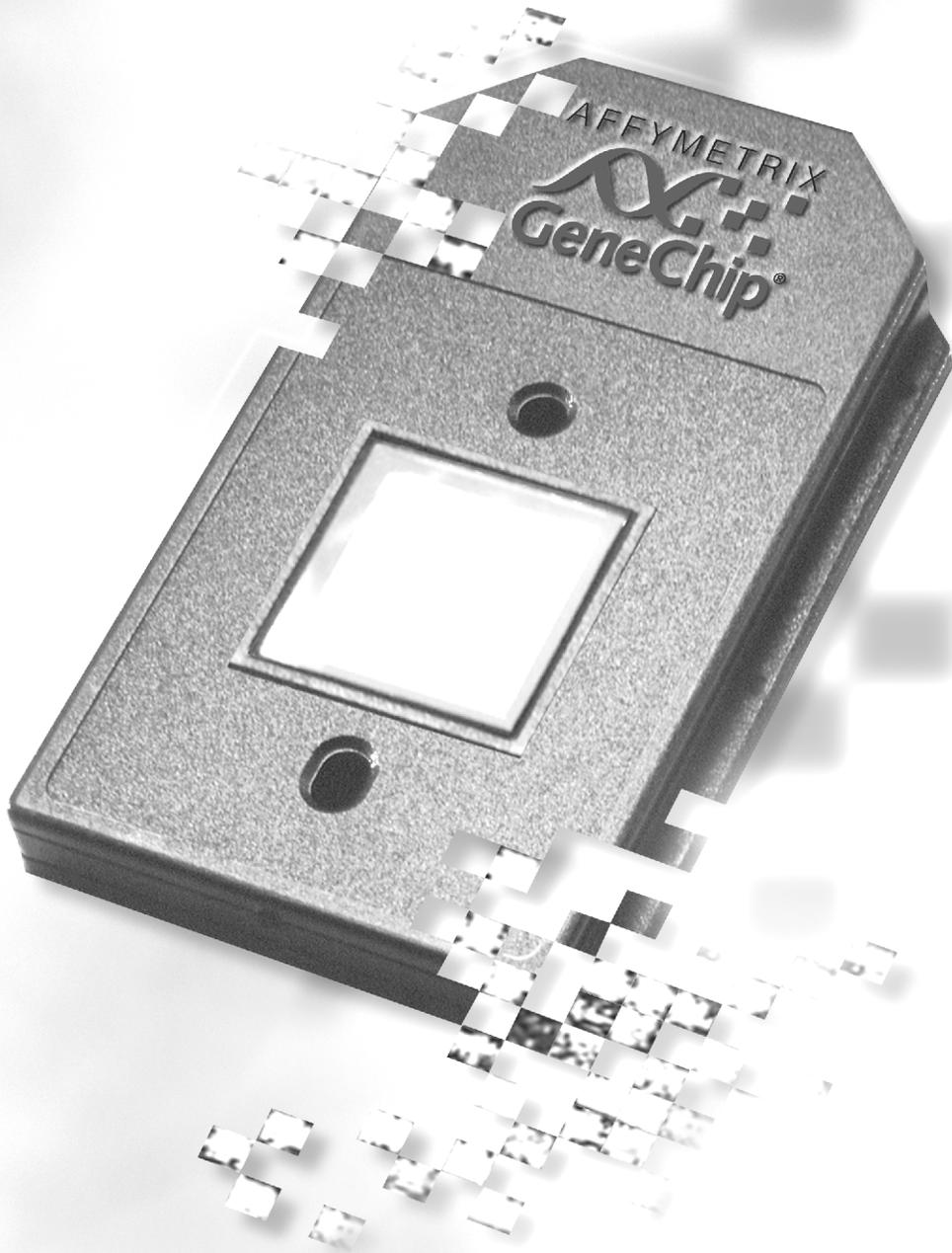
Figure 10.36

p53 Report dialog box, imported reference file selected

4. Click **OK** to close the HIV PRT Plus Report dialog box.



Chapter 11



Genotyping Analysis

GeneChip® probe arrays for genotyping analysis interrogate the target at known mutation sites that define alleles and enable genotype calls. Microarray Suite analyzes the hybridization intensity data to call a genotype for each allele examined.

This chapter describes how to:

- run a genotyping analysis
- view the analysis output (*.chp) in graphical or tabular format in the Nucleotide Analysis window (NAW)
- generate a report

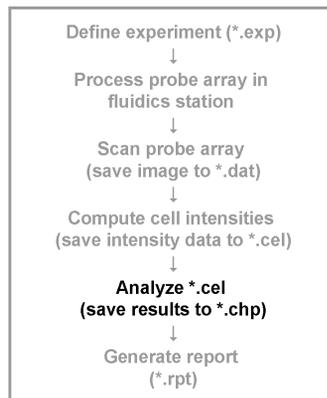


Figure 11.1
Assay & analysis flow chart

Overview

For genotyping analysis, block tiling specifies probe sets that examine known polymorphic sites of a gene and are designed to be complementary to either wildtype or mutant sequences. Alternative Block tiling specifies probe sets that interrogate human single nucleotide polymorphisms (SNP).

Microarray Suite analyzes the hybridization intensity data to identify the alleles and call the genotype. The Nucleotide Analysis window (NAW) displays the results contained in the analysis output or *chip* file (*.chp).

If the NAW is already open, the results are added to the open window and it may be necessary to scroll down to see the newly added results. A second toolbar (see Appendix J) and additional menu bar commands are available in the NAW.

In the NAW, you may view:

- a tabular display of genotype analysis output (*.chp)
- probe set intensity graphs

Genotyping Algorithms

Microarray Suite selects the probe array calling algorithm for an experiment based on the probe array type specified during experiment setup (see [Table 11.1](#)).

Table 11.1
Genotyping analysis algorithms

Genotype Analysis		
Probe Array Calling Algorithm	GeneChip® Probe Array	Regions Genotyped
Block	CYP450	15 known mutations in the CYP2D6 gene and 3 mutations in the CYP2C19 gene that encode the CYP450 enzymes.
Genotyping	HuSNP™	A panel of 1,494 human single nucleotide polymorphisms covering all 22 autosomes and the X chromosome.

Block Algorithm

The Block algorithm analyzes hybridization intensity data from the GeneChip® CYP450 probe array to genotype 15 known mutations in the CYP2D6 gene and three mutations in the CYP2C19 gene that encode the CYP450 enzymes. The algorithm examines hybridization intensity data to produce a call of wildtype, mutant, heterozygous, or ambiguous (N) for each site examined.



Note

The Block algorithm has no user-modifiable analysis settings.

Genotyping Algorithm

The Genotyping algorithm analyzes hybridization intensity data from the GeneChip® HuSNP™ probe array to call the genotype for 1,494 human SNPs (Table 11.2).

Table 11.2

Affymetrix® HuSNP™ Mapping assay, possible genotype calls

Genotyping Algorithm	
Possible Call	Genotype
A	Homozygous A
B	Homozygous B
AB	Heterozygous AB
AB_A	Heterozygous AB or homozygous A
AB_B	Heterozygous AB or homozygous B
No Signal	Insufficient signal to make a call

Viewing the Genotyping Algorithm Settings

1. In the Analysis Settings shortcut bar, click **Genotyping** .
⇒ The Genotyping Analysis Settings dialog box appears (Figure 11.2).

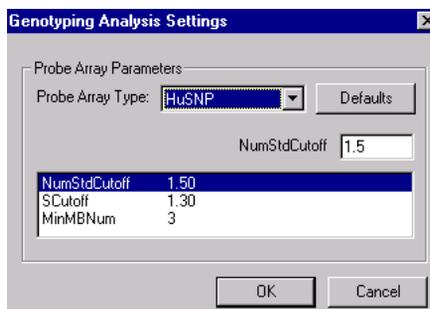


Figure 11.2
Genotyping Analysis Settings

2. Confirm or change the settings as described in Appendix H.

Running a Genotype Analysis

Microarray Suite analyzes a cell intensity file (*.cel) and generates an analysis output file (*.chp).

- To run an analysis, do one of the following:
 - In the data file tree, right-click a cell intensity file (*.cel) and select **Analyze** from the shortcut menu.
 - Select **File** → **Analysis** from the menu bar, then choose the desired *.cel file from the Analyze dialog box that appears.
 - If the image is displayed, click the **Analyze** toolbar button  or select **Run** → **Analysis** from the menu bar.

The status log displays:

- the name of the analysis output file (*.chp)
- the location of the *.chp
- a message indicating when the analysis is completed

If the status bar is not displayed, click the **Status Log** toolbar button  in or select **View** → **Status Bar** from the menu bar.

When the analysis is finished, the Nucleotide Analysis window (NAW) displays the output analysis output file (*.chp) (Figure 11.3). If the NAW is already open, the results are added to the open window. It may be necessary to use the scroll bars at the bottom and right side of the NAW to see the newly added results.

Nucleotide Analysis Window (NAW)

The Nucleotide Analysis window (NAW) displays the tabular and graphic analysis output for genotyping assays (Figure 11.3). The NAW opens when a genotyping analysis is completed or a genotyping analysis output file (*.chp) is opened.

The NAW is divided into two windowpanes. The bottom pane displays genotype calls and related information and the top pane displays intensity graphs of probe set data.

You may resize the windowpanes may be resized (see Appendix K) and view the contents of the entire window using the scroll bars at the right side of the NAW. The status bar at the bottom of the NAW indicates the view option (All Blocks or Mutations Only) and the number of data rows displayed.

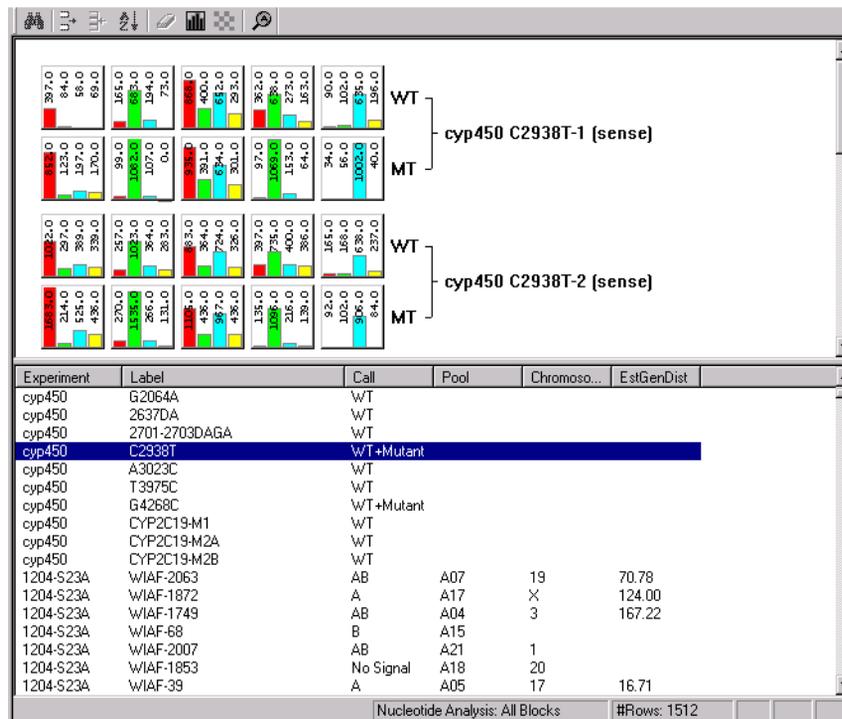


Figure 11.3
NAW, graph pane (top), tabular data (bottom)

Nucleotide Analysis Window (NAW) View Options

Viewing Mutations Only

1. Select **View** → **Mutation Only** from the menu bar.



Note

The table will be empty if there are no mutations.

2. To display all results, select **View** → **All Blocks** from the menu bar.

Locating a Polymorphism

1. Right-click the data table, then click **Select** in the shortcut menu.
 - ⇒ The Select Block dialog box appears ([Figure 11.4](#) and [Figure 11.5](#)).
2. Select a polymorphism label (identifier) from the drop-down list, then click **Select**.
 - ⇒ The NAW highlights and displays the corresponding row ([Figure 11.4](#) and [Figure 11.5](#)).

The screenshot displays the NAW software interface. At the top, there are two gene models for Ks122b WIAF-2525, one for the sense strand and one for the anti-sense strand. Each model shows a series of colored bars representing different probes. Below these models is a data table with the following columns: Experiment, Label, Call, Pool, Chromosome, and EstGenDist. The table contains 10 rows of data. A 'Select Block' dialog box is open over the table, with 'WIAF-2525' selected in the dropdown menu. The 'Select' button is highlighted.

Experiment	Label	Call	Pool	Chromosome	EstGenDist
Ks122b	WIAF-1596	B	A17		
Ks122b	WIAF-2513	B	A07		
Ks122b	WIAF-1464	A	A07		
Ks122b	WIAF-2525	B	A07	8	
Ks122b	WIAF-2490	B	A07		
Ks122b	WIAF-1706	AB	A07	22	
Ks122b	WIAF-1588	AB	A20	11	
Ks122b	WIAF-4501	AB	A17	1	
Ks122b	WIAF-4502	A	A07	2	

Figure 11.4
NAW, HuSNP analysis output file (*.chp), Select Block dialog box

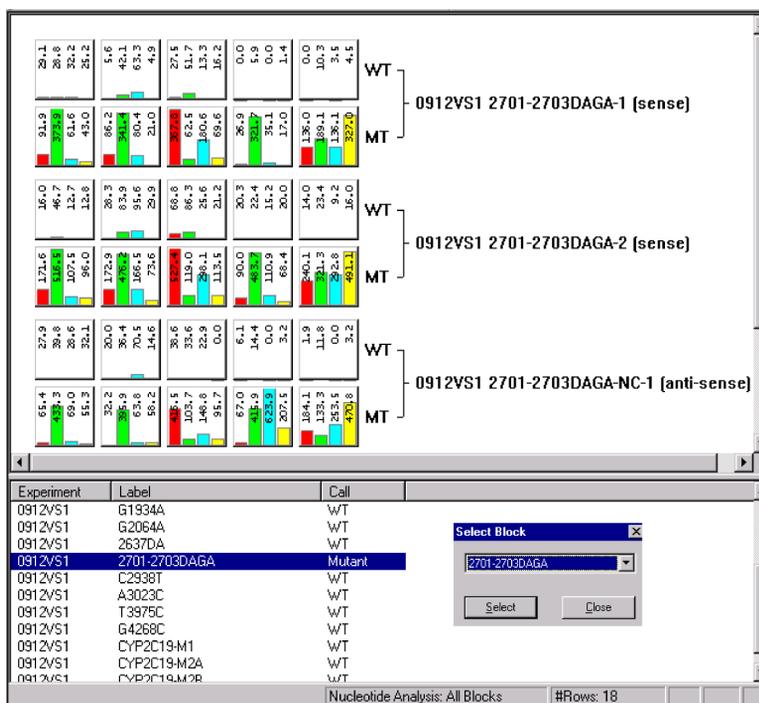


Figure 11.5
NAW, CYP450 analysis output file (*.chp), Select Block dialog box

Hiding Data Table Rows

1. Select the desired row(s) in the data table.
To select non-adjacent rows, press and hold the **Ctrl** key while you click the desired rows. To select adjacent rows, press and hold the **Shift** key while you click the first and the last row in the selection.
2. Right-click the data table, then click **Hide** in the shortcut menu.
3. To restore all hidden rows in the NAW, right-click the table and select **Unhide All** from the shortcut menu. Alternatively, select **View** → **Unhide All** from the menu bar.
⇒ This adds the restored rows to the bottom of the table.

Removing an Experiment from the NAW

1. Highlight any result row of the experiment(s) to be removed.
2. Select **Edit** → **Remove Experiments** from the menu bar.



Note

Selecting **Edit** → **Remove Experiments** is equivalent to closing the *.chp file.

Sorting Information

You can alphabetically sort the NAW table columns that contain text or numerically sort columns that contain numeric data in ascending or descending order.

Single Column Sort

1. To sort a single column in ascending order, click the column header in the NAW.
2. Click the column header again to sort the information in descending order.

Multiple Column Sort

1. Select **Edit** → **Sort** from the menu bar.
⇒ The Sort dialog box appears (Figure 11.6).

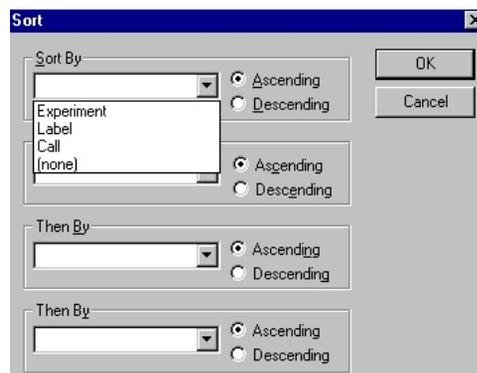


Figure 11.6
Sort dialog box

2. Select the first category for the sort from the **Sort By** drop-down list, then choose a sort order.
3. Select subsequent sort categories from the **Then By** drop-down lists, choose the sort order, and click **OK** when finished.

Find Feature

The Find feature searches the data table.

1. To start a search, do either of the following:
 - Click the **Find** toolbar button .
 - Right-click in the data table, then select **Find** from the shortcut menu or select **Edit** → **Find** from the menu bar.

⇒ The Find dialog box appears (Figure 11.7).



Figure 11.7
Find dialog box

2. Enter the desired text for the search (up to 256 alphanumeric characters including spaces) in the **Find What** field.
3. Click **Find Next** to view the first search result.
4. Continue to click **Find Next** to view each successive search result.

Note

The **Find** command finds all strings that match the text string for the search. For example, using the **Find** command to search for the text string 4a would find G2064A and G1934A as well as other occurrences of 4a.

Genotyping Tabular Data

CYP450 Tabular Data

The CYP450 tabular data includes information from the analysis output file (*.chp) (Figure 11.8). Each row displays the results for one allele.

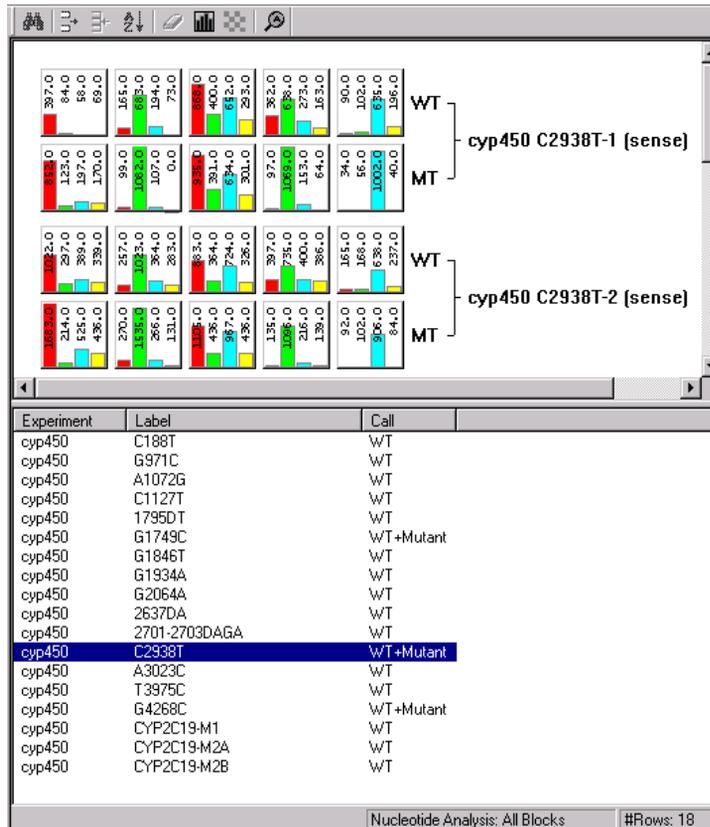


Figure 11.8
Nucleotide Analysis window (NAW), CYP450 analysis output (*.chp)

Experiment	The name of the analysis output (*.chp) file.
Label	The identifier for the polymorphism. (See the international standard for human cytochrome P450 (CYP) allele nomenclature accessible at http://www.imm.ki.se/CYPalleles/ .)
Call	The Block algorithm makes one of three possible calls for each site of the CYP2D6 and CYP2C19 gene examined: homozygous wildtype (WT), homozygous mutant (MT), or heterozygous (WT+Mutant).

HuSNP™ Tabular Data

The HuSNP™ tabular data includes information from the analysis output file (*.chp) (Figure 11.9). Each row of the table displays the analysis result for one SNP.

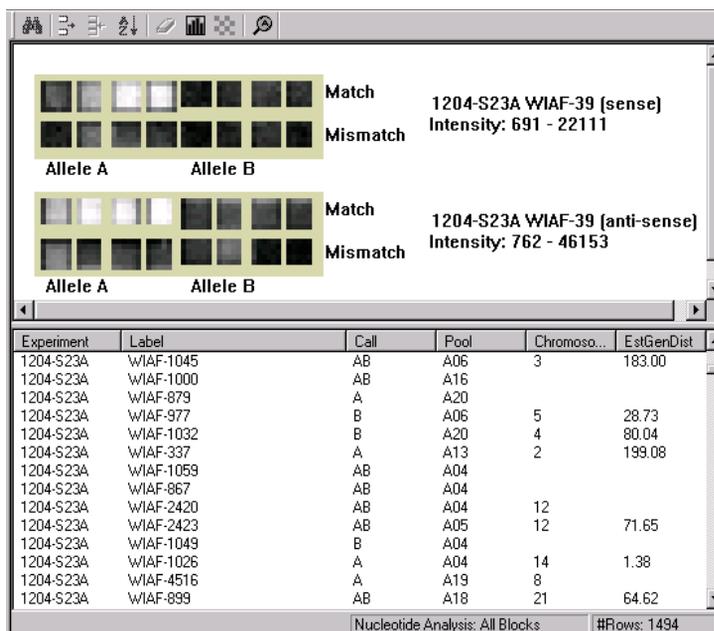


Figure 11.9
NAW, HuSNP analysis output file (*.chp)

Experiment	The name of the analysis output (*.chp) file.
Label	The identifier for the SNP.
Call	The algorithm generates six possible calls: (1) A (homozygous A allele), (2) B (homozygous B allele), (3) AB (heterozygous), (4) AB_A (two possible genotypes AB or A that could not be distinguished), (5) AB_B (two possible genotypes AB or B that could not be distinguished), or (6) No Signal (insufficient data passed the quality tests to perform an analysis).
Pool	The number of the primer pool containing the primer for the SNP (refer to the <i>GeneChip® HuSNP™ Mapping Assay User Manual</i>).
Chromosome	The number or name of the chromosome that contains the SNP.
Estimated Genetic Distance	The estimated chromosomal location of the SNP expressed in centiMorgans. Genetic distances are measured from the top of the chromosome's short arm. A large set of markers including the SNPs were physically mapped onto radiation hybrid panels. The physical locations of the SNP markers were integrated into the genetic map to obtain the Estimated Genetic Distances.

Exporting Genotype Data

You can export genotype data to a tab delimited text file (*.txt).

1. Open a genotype analysis output file (*.chp).
⇒ The NAW displays the file.
2. Choose **Edit** → **Select All** from the menu bar.
3. Select **File** → **Save As** from the menu bar.
⇒ The Save As dialog box appears ([Figure 11.10](#)).



Figure 11.10
Save As dialog box

4. Enter a name for the text file in the **File name** box, then click **Save**.

Displaying Two Columns in the Text File

You can display the text file data in one (default) or two columns ([Table 11.3](#)).



Note

The two-column display option is not recommended for CYP450 data.

Table 11.3

HuSNP genotype calls displayed in separate columns in a text file

GeneChip® HuSNP™ Mapping Assay		
Call	Calls Displayed in Separate Columns	
A	A	A
B	B	B
AB	A	B
AB_A	AB_A	
AB_B	AB_B	
No Signal	No Signal	
A	A	A

1. Click the **Options** toolbar button . Alternatively, select **Analysis** → **Options** from the menu bar.
⇒ The Analysis Options dialog box appears ([Figure 11.11](#)).



Figure 11.11
Analysis Options, Export tab

2. Click the Export tab.
3. Choose the **Save calls in separate columns** option.

Genotyping Graphs

CYP450 Intensity Graphs

The CYP450 intensity graphs ([Figure 11.12](#)) display the cell intensity values (*.cel) for the probe sets specified by the Block tiling strategy. The graph bars represent the four types of probes (1 perfect match and 3 mismatch) that interrogate the wildtype or mutant sequence. Each type of probe in a probe set contains a different nucleotide at the substitution position (a, c, g, or t; and the graph bars follow the same order).

Each row of graphs represents probe sets complementary to the wildtype (WT) or mutant (MT) sequence (except at the substitution position).

From left to right, graphs in the first and second columns represent probe sets that interrogate the target sequence two bases or one base upstream from the mutation site respectively. The graphs in the third column represent probe sets that interrogate the target sequence at the mutation site. The graphs in the fourth and fifth columns represent probe sets that interrogate the target sequence one or two bases downstream from the mutation site respectively.

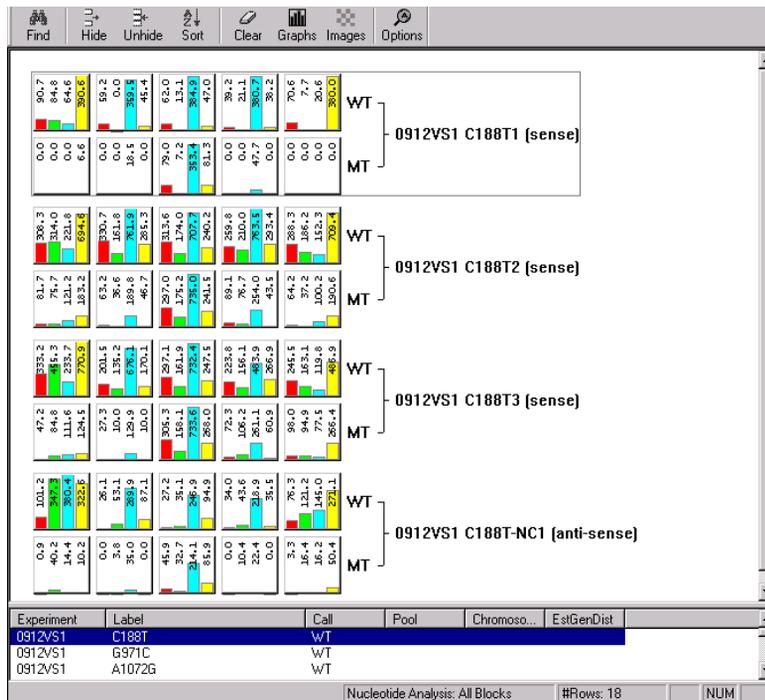


Figure 11.12
Nucleotide Analysis window (NAW), CYP450 intensity graphs for the C188T marker

Plotting CYP450 Intensity Graphs

1. Open a CYP450 analysis output file (*.chp).

⇒ The NAW displays the *.chp.

2. Highlight the desired row(s) in the data table.

To select non-adjacent rows, press and hold the **Ctrl** key while you click the desired rows. To select adjacent rows, press and hold the **Shift** key while you click the first and the last row in the selection.

3. Do one of the following to display the graph(s):

- Click the **Graphs** toolbar button .
- Right-click in the bottom pane of the NAW, then select **Draw Intensity Graphs** from the shortcut menu.
- Select **Graphs** → **Draw Intensity Graphs** from the menu bar.

It may be necessary to use the scroll bar at the right side of the NAW to view all of the intensity graphs.

CYP450 Intensity Graph Options

You can change some display features of the CYP450 intensity graphs.

1. Click the **Analysis Options** toolbar button . Alternatively, select **Analysis** → **Options** from the menu bar.

⇒ The Analysis Options dialog box appears ([Figure 11.13](#)).

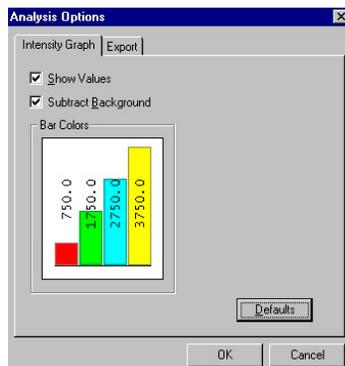


Figure 11.13
Analysis Options, Intensity Graph tab

Show Values

Choose this option to display the hybridization intensity value for each bar in the intensity graph.

Subtract Background

Choose this option to display background-subtracted intensity data.

Bar Colors

The color of a graph bar may be changed.

1. Click a bar in the **Bar Colors** box.
⇒ This displays the color palette (Figure 11.14).

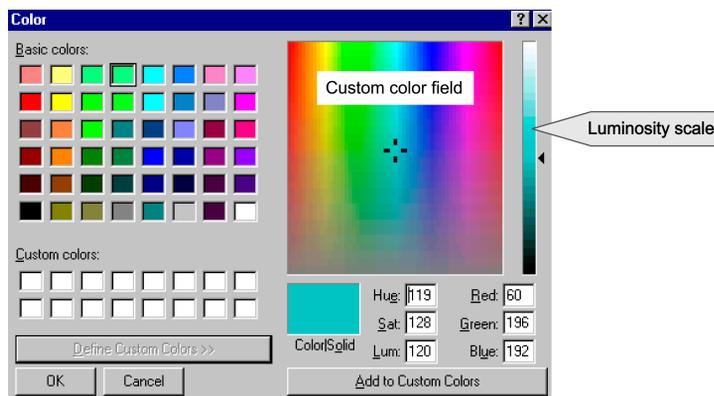


Figure 11.14
Color palette

2. To select a predefined color, click one of the basic colors.
3. To define a custom color, click **Define Custom Colors**, then use the click-and-drag method to move the cross hairs in the custom color field. Adjust the color brightness using the luminosity scale to the right. When finished, click **Add to Custom Colors** to apply the color.
4. Click **OK** to close the color palette.

The color choices are saved on a per user basis.

HuSNP™ Intensity Graphs

The HuSNP intensity graphs display the cell intensity values (*.cel) for the probe sets specified by the Alternative Block tiling strategy.

All of the graphs in a single row represent the probe sets specific for a particular SNP allele (A or B). Each vertical pair of graphs (one pair per column) represents a pair of probe sets or *miniblock* that interrogates the A and B allele of the SNP.

Each graph contains two bars that display the cell intensities of the perfect match (PM) and mismatch (MM) probe. The bars are ordered along the x-axis according to the type of nucleotide at the substitution position (Figure 11.15).

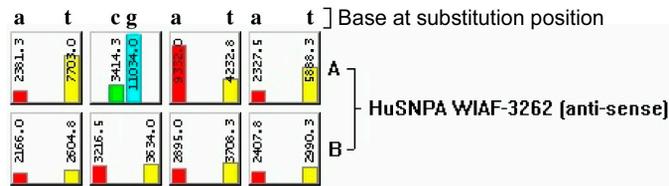


Figure 11.15

HuSNP™ intensity graph (bars ordered according to base type at the substitution position)

Plotting HuSNP Intensity Graphs

1. In the data tree, double-click a HuSNP™ analysis output file (*.chp).
⇒ The NAW displays the file.

2. Highlight the desired row(s) in the data table.

To select non-adjacent rows, press and hold the **Ctrl** key while you click the desired rows. To select adjacent rows, press and hold the **Shift** key while you click the first and the last row in the selection.

3. Do one of the following to display the graph(s):

- Click the **Graphs** toolbar button .
- Right-click in the bottom pane of the NAW, then select **Draw Intensity Graphs** from the shortcut menu.

- Select **Graphs** → **Draw Intensity Graphs** from the menu bar.

The NAW displays the graphs for the probe sets that call the genotype. It may be necessary to use the scroll bar at the right side of the NAW to view all of the intensity graphs.

Measured Images

A measured image displays the probe set hybridization intensity data (*.dat).

Displaying Measured Images

1. In the data table, click a probe set name(s).

To select adjacent names, press and hold the **Shift** key while you click the first and last probe set name. To select non-adjacent names, press and hold the **Ctrl** key while you click the desired probe set names.

2. Do one of the following:

- Click the **Images** toolbar button .
- Right-click the selection and select **Draw Measured Image** from the shortcut menu.
- Select **Graphs** → **Measured Images** from the menu bar.

⇒ The graph pane in the NAW displays the measured image ([Figure 11.16](#)).

When an additional image is generated, the graph pane automatically scrolls to the new image. Use the scroll bar to view all of the measured images created during a session. The graph pane may be resized as described in [Appendix K](#).

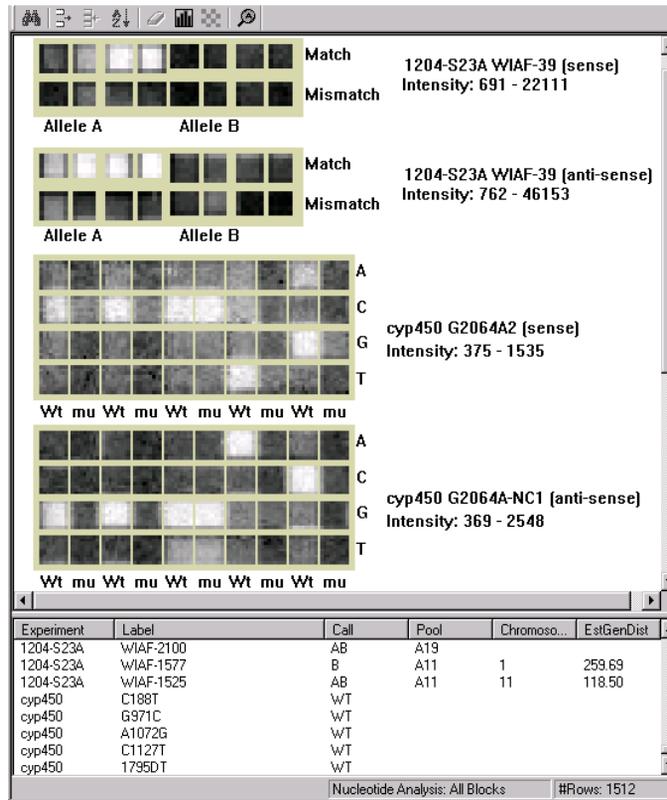


Figure 11.16
Nucleotide Analysis window (NAW), measured images from a HuSNP™ and CYP450 probe array

Removing Graphs and Measured Images

Removing All Graphs or Images

1. Right-click the graph pane.
2. Click **Clear All Graphs** in the shortcut menu. Alternatively, select **Graphs** → **Clear All Graphs** from the menu bar.

Removing Selected Graphs or Images

1. Double-click the graph(s) or image(s)
⇒ This draws an outline around the selected graph or image.
2. Do one of the following:
 - Click the Clear toolbar button .
 - Right-click the graph pane and select **Clear** → **Selected Graphs** from the shortcut menu.
 - Select **Graphs** → **Clear Selected Graphs** from the menu bar.

Genotyping Analysis Reports

Microarray Suite generates a report (*.rpt) from an analysis output file (*.chp).

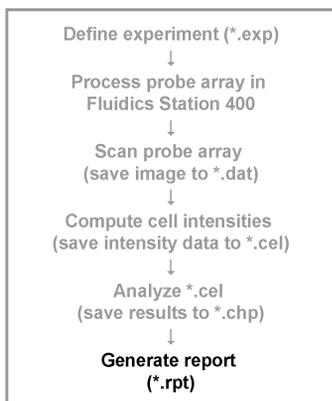


Figure 11.17
Assay & analysis flow chart

CYP450 Report

The CYP450 report ([Figure 11.18](#)) includes the:

- name and location of the analysis output file (*.chp)
- probe array type and experiment name
- genotypes for the CYP2D6 and CYP2C19 gene

To generate the report:

- 1.** In the data tree, right-click a CYP450 analysis output file (*.chp) and select **Report** from the shortcut menu.
⇒ This displays the report in the main display area.
- 2.** To best view the CYP450 report:
 - a.** Use Courier New font (10 point) (select **View** → **Set Font** from the menu bar).
 - b.** Set the tab stops set at 14 (select **View** → **Set Tab Stops** from the menu bar).

```

Algorithm:                               BlockRatioCall
Alg parameters:                           None
cyp450 Gene: CYP2D6
Genotype: CYP2D6*2/CYP2D6*1A

Polymorphism  Call      Discrimination  ALLELE LOOKUP TABLE
Quality      1 2 3 4A 4B 4C 4D 4E 6A 6B 7 8 9 10A 10B 11
C188T        WT          | | | | +| +| +| +| +| | | | | + | + | |
G971C        WT          | | | | | | | | | | | | | | | | +|
A1072G       WT          | | | | +| +| | | | | | | | | | | |
C1127T       WT          | | | | | | | +| | | | | | | | + | |
1795DT       WT          | | | | | | | | | +| +| | | | | | |
G1749C       HET         | +| | | +| | +| +| +| | | +| | + | + | +|
G1846T       WT          | | | | | | | | | | | +| | | | | |
G1934A       WT          | | | | +| +| +| +| +| | | | | | | |
G2064A       WT          | | | | | | | | | | | +| | | | | |
2637DA       WT          | | +| | | | | | | | | | | | | | |
2701-2703DAGA WT        | | | | | | | | | | | | | +| | | |
C2938T       HET         | +| | | | | | +| | | +| | | | | +|
A3023C       WT          | | | | | | | | | | | +| | | | | |
T3975C       WT          | | | | | | +| | | | | | | | | | |
G4268C       HET         | +| | | +| +| +| +| | | +| | + | + | +|

cyp450 Gene: CYP2C19
Genotype: CYP2C19*1

Polymorphism  Call      Discrimination  ALLELE LOOKUP TABLE
Quality      2 3
CYP2C19-M1    WT          | +| |
CYP2C19-M2A   WT          | | |
CYP2C19-M2B   WT          | +|

```

Figure 11.18
CYP450 report

Polymorphism	Describes the polymorphisms detected by the GeneChip® CYP450 probe array (for example, C188T indicates a base change from C to T at nucleotide position 188; DAGA indicates the triple base deletion AGA).
Call	Displays the genotype call for each polymorphism (WT = wildtype, HET = heterozygous, MUT = homozygous mutant, or N = no call).
Discrimination Quality	A measure of how well the cell intensity values discriminate between the wildtype and non-wildtype probes in a probe set. The column displays Low if the discrimination quality is low for WT, HET, or MUT calls.

The two Allele Lookup tables display the mutations that define the alleles of the CYP2D6 and CYP2C19 gene. The tables may be used to confirm the genotype calls made by the software. The first row of the table lists the allele names. In the CYP2D6 Allele Lookup table, 1 is the wildtype, 2 is allele 2 of the CYP2D6 gene (or CYP2D6*2), 3 is allele 3 of the CYP2D6 gene (or CYP2D6*3), and so on.

The column associated with the allele name contains a + for each polymorphic mutation required to call that particular allele. For example, in the CYP450 report shown in [Figure 11.18](#), allele 2 is defined by 3 mutations: G1749C (G changes to C at nucleotide position 1749), C2938T, and G4268C. Since the target contained a mutation at each of these polymorphic locations (the call is heterozygous), the haplotype is CYP2D6*2. The remaining haplotype is wildtype (CYP2D6*1A) and the genotype is CYP2D6*2 / CYP2D6*1A).

Mapping Report

The mapping report for the GeneChip® HuSNP™ probe array (Figure 11.19) includes the percentage of SNPs assigned to a genotype and the percentage of each genotype call for the array as well as per primer pool. This provides both a global and a localized assessment of assay performance.

The mapping report displays the probe array type and the name of the analysis output file (*.chp) examined, the algorithm type and version, and the values specified for the user-modifiable algorithm parameters (NumStdCutoff, Scutoff, and MinMBNum).

In the report, the probe array summary displays the correlation coefficient of hybridization intensity data scanned at 530 and 570 nm, the maximum and median intensity for the array, and the genotype calls for the array.

The pool summary displays the percentage of genotype calls for each primer pool (A01 – A21) in the assay.

To generate the report:

- 1.** In the data tree, right-click a HuSNP™ analysis output file (*.chp) and select **Report** from the shortcut menu.
⇒ This displays the report in the main display area.
- 2.** To best view the HuSNP report:
 - a.** Use Arial font (10 point) (select **View** → **Set Font** from the menu bar).
 - b.** Set the tab stops set at 10 (select **View** → **Set Tab Stops** from the menu bar).

ReportType: Mapping Report
Date: 11:21AM 06/19/2001

Filename: HuSNP4B.CHP
Probe Array Type: HuSNP
Algorithm: Genotyping
Alg version: 4.0
Alg parameters: NumStdCutoff=1.50 SCutoff=1.30 MinMBNum=3

Probe Array Summary:
CorrelationCoefficient=0.999
MaxIntensity=15807.0
MedianIntensity=737.8

%Pass 88.3
%A 28.8
%AB 29.3
%B 28.9
%AB_A 0.6
%AB_B 0.7
%NoSignal 11.7

Pool Summary:

Pool	%Pass	%A	%AB	%B	%AB_A	%AB_B	%NoSignal
A01	97.9	28.7	44.7	24.5	0.0	0.0	2.1
A02	96.5	40.0	29.4	25.9	0.0	1.2	3.5
A03	89.2	27.7	28.9	32.5	0.0	0.0	10.8
A04	85.5	26.3	23.7	34.2	0.0	1.3	14.5
A05	86.1	27.8	25.0	33.3	0.0	0.0	13.9
A06	84.5	31.0	32.4	18.3	1.4	1.4	15.5
A07	80.0	36.0	25.3	18.7	0.0	0.0	20.0
A08	78.8	27.5	26.3	22.5	1.3	1.3	21.3
A09	90.2	28.0	26.8	32.9	1.2	1.2	9.8
A10	96.5	27.1	30.6	37.6	0.0	1.2	3.5
A11	81.4	25.6	24.4	29.1	1.2	1.2	18.6
A12	97.4	39.0	28.6	28.6	1.3	0.0	2.6
A13	95.5	34.1	29.5	30.7	0.0	1.1	4.5
A14	91.4	24.3	31.4	35.7	0.0	0.0	8.6
A15	95.9	27.4	34.2	31.5	2.7	0.0	4.1
A16	92.1	21.1	38.2	31.6	1.3	0.0	7.9
A17	91.0	30.8	34.6	25.6	0.0	0.0	9.0
A18	70.0	14.0	12.0	44.0	0.0	0.0	30.0
A19	92.2	35.3	27.5	25.5	2.0	2.0	7.8
A20	69.2	15.4	28.2	23.1	0.0	2.6	30.8
A21	59.0	25.6	12.8	20.5	0.0	0.0	41.0
Total	88.3	28.8	29.3	28.9	0.6	0.7	11.7

Figure 11.19
Mapping report

%Pass	The percentage of SNPs assigned a genotype.
%A	The percentage of homozygous A genotypes.
%AB	The percentage of heterozygous AB genotypes.
%B	The percentage of homozygous B genotypes.
%AB_A	The percentage of genotypes that could be AB or A (the two possible calls could not be distinguished).
% AB_B	The percentage of genotypes that could be AB or B (the two possible calls could not be distinguished).
% NoSignal	The percentage of genotypes that could not be called.

Genotyping Analysis Viewer Report

The Genotyping (GT) Viewer analyzes Affymetrix® HuSNP™ *.cel files to generate the GT Viewer report that provides a *relative allele signal* (RAS) for each HuSNP™ marker. The RAS, a metric derived from the observed hybridization signals, is a quantitative representation of the two possible alleles. It is expressed as a value between zero and one.

The RAS tends to segregate into three distinct clusters that correspond to the three biallelic genotypes for a marker (AA, AB, BB). A RAS close to zero indicates a predominance of B allele, RAS close to one indicates a predominance of A allele, and an intermediate RAS indicates signal derived from both alleles.

The GT viewer compares the RAS of a control and an experimental sample, enabling you to detect allelic shifts between the two samples. The GT Viewer report displays the change in RAS in the *DeltaRAS* column. The additional *StdUnit* column reports an approximate normalization of delta RAS based on observed standard deviations for RAS values for each SNP. Alternatively, the GT Viewer report can use a custom set of user-defined standard deviation values. Appendix I provides information on how to make a customized table.

Selecting Cell Intensity Files for the GT Viewer Report

The following procedure allows you to select a control and experimental HuSNP™ cell intensity file for the report.

1. In the Microarray Suite shortcut bar, click **GT viewer** . Alternatively, select **Run** → **GT Viewer** from the menu bar.
⇒ This displays the GT viewer ([Figure 11.20](#)).

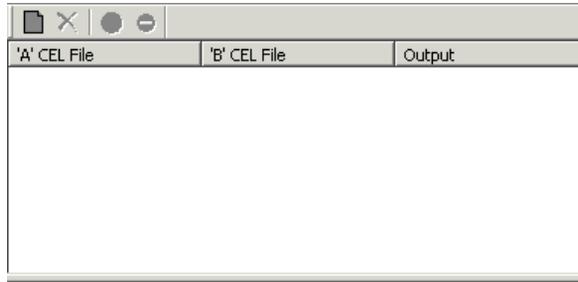


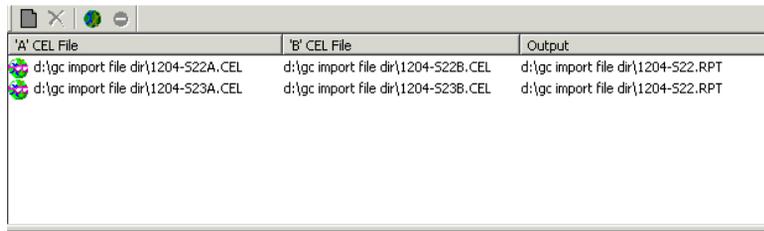
Figure 11.20
GT Viewer

2. Use one of the following methods to select a control and experimental HuSNP™ cell intensity file for the report.

 **Note**

*Add the control sample *.cel to the GT Viewer first, then the experiment *.cel. You need only select the A or B *.cel file. The other member of the *.cel file pair is automatically placed in the GT viewer.*

- Drag the control sample *.cel from the data file tree to the GT Viewer, then drag the experiment *.cel from the data file tree to the GT Viewer.
 - In the GT Viewer, click the **Add** toolbar button , then double-click the control sample *.cel in the Open dialog box that appears. Repeat to select the experiment *.cel.
 - Select **Edit** → **Add Cell File** from the menu bar, then select the control sample *.cel file from the Open dialog box that appears. Repeat to select the experiment *.cel.
- ⇒ The GT Viewer displays the selected *.cel files and the default output name for the genotyping analysis viewer report (*.rpt) ([Figure 11.21](#)).



'A' CEL File	'B' CEL File	Output
d:\jgc import file dir\1204-S22A.CEL	d:\jgc import file dir\1204-S22B.CEL	d:\jgc import file dir\1204-S22.RPT
d:\jgc import file dir\1204-S23A.CEL	d:\jgc import file dir\1204-S23B.CEL	d:\jgc import file dir\1204-S22.RPT

Figure 11.21

GT Viewer, control sample *.cel (top) and experiment *.cel (bottom)

✓ Note

If only one sample is added to the GT Viewer (A.cel and B*.cel for the sample), the Genotyping Analysis Viewer Report displays RAS, but not Delta RAS for each HuSNP™ marker.*

3. To remove a *.cel file pair from the GT Viewer, click the *.cel, then click the **Remove** toolbar button  or select **Edit** → **Remove Cell File** from the menu bar.
4. To change the default output name (*.rpt):
 - a. Double-click the default *.rpt name or select **Edit** → **Rename Output File** from the menu bar.
 - b. Enter a new name for the report (*.rpt).
5. Click the **Analyze** toolbar button .

⇒ The report is automatically displayed in the main display area ([Figure 11.22](#)).
6. To quit an analysis in process, click the **Stop** toolbar button .
7. To best view the Genotyping Analysis Viewer report:
 - a. Use Arial font (10 point) (select **View** → **Set Font** from the menu bar).
 - b. Set the tab stops at 15 (select **View** → **Set Tab Stops** from the menu bar).

CEL Filenames (1st Pair): R1_030801A.CEL, R1_030801B.CEL
 CEL Filenames (2nd Pair): R2_030801A.CEL, R2_030801B.CEL
 Probe Array Type: HuSNP
 Algorithm: Genotyping
 Alg parameters: NumStdCutoff=1.50 SCutoff=1.30 MinMBNum=3

Marker	Chromosome	EstGenDist	R1_030801 Call	RAS1	RAS2	R2_030801 Call	RAS1	RAS2	R1_030801__R2_030801 Delta RAS	StdUnits
WIAF-3819	20	22.00	A	1.000	-1000	A	1.000	-1000	0.000	-1000
WIAF-3821	4	125.80	B	0.000	-1000	B	0.000	-1000	0.000	-1000
WIAF-3818	4	125.80	B	0.032	-1000	B	0.076	-1000	0.043	-1000
WIAF-3271	2	180.30	AB	0.671	-1000	AB	0.726	-1000	0.055	1.658
WIAF-2729			AB	0.731	-1000	AB	0.768	-1000	0.037	0.970
WIAF-548	5	102.50	NoSignal	-9	-1000	NoSignal	-9	-1000	-9	-1000
WIAF-267	5	121.60	B	0.114	-1000	NoSignal	0.191	-1000	0.078	-1000
WIAF-272	10	31.50	NoSignal	0.228	-1000	NoSignal	0.269	-1000	0.041	-1000
WIAF-1918	12	122.80	B	0.089	-1000	B	0.141	-1000	0.053	-1000
WIAF-1904	22	41.80	A	0.880	-1000	A	0.921	-1000	0.041	-1000
WIAF-1362	8	62.50	B	0.052	-1000	B	0.091	-1000	0.039	-1000
WIAF-1634	9	64.10	B	0.000	-1000	B	0.039	-1000	0.039	-1000
WIAF-2460			A	0.991	-1000	A	1.000	-1000	0.009	-1000
WIAF-2184	19	42.20	B	0.000	-1000	B	0.000	-1000	0.000	-1000
WIAF-2189	7	5.20	A	0.963	-1000	A	1.000	-1000	0.037	-1000
WIAF-2180	X	11.10	AB	0.557	-1000	AB	0.584	-1000	0.027	0.789
WIAF-3540	6	47.00	NoSignal	0.948	-1000	A	0.929	-1000	0.018	-1000
WIAF-3542	10	31.80	A	0.693	-1000	A	0.746	-1000	0.053	-1000
WIAF-3537	10	17.30	NoSignal	0.932	-1000	A	0.903	-1000	0.029	-1000
WIAF-3272	1	160.50	B	0.013	-1000	B	0.010	-1000	0.003	-1000
WIAF-3268	13	77.10	B	0.151	-1000	B	0.198	-1000	0.048	-1000
WIAF-2999			NoSignal	0.055	-1000	B	0.080	-1000	0.025	-1000
WIAF-2730			AB	0.487	-1000	AB	0.515	-1000	0.028	1.461

Figure 11.22

Genotyping analysis viewer report

Marker	The identifier for the SNP.
Chromosome	The number or name of the chromosome that contains the SNP.
EstGenDist	The estimated chromosomal location of the SNP expressed in centiMorgans. Genetic distances are measured from the top of the chromosome's short arm. A large set of markers including the SNPs were physically mapped onto radiation hybrid panels. The physical locations of the SNP markers were integrated into the genetic map to obtain the Estimated Genetic Distances.
Call	The algorithm generates six possible calls: (1) A (homozygous A allele), (2) B (homozygous B allele), (3) AB (heterozygous), (4) AB_A (two possible genotypes AB or A that could not be distinguished), (5) AB_B (two possible genotypes AB or B that could not be distinguished), or (6) No Signal (insufficient data passed the quality tests to perform an analysis).

RAS 1	The relative allele signal for the first used block of the SNP.
RAS 2	The relative allele signal for the second used block of the SNP, if a second block exists. Note: a second block may exist, but may not be used for making a genotype call.
Delta RAS	The difference between the control and experimental RAS for each HuSNP™ marker.
StdUnits	The normalized DeltaRAS using the observed RAS standard deviation in heterozygotes. (See Appendix I for information on StdUnits.)

 **Note**

A HuSNP™ marker may have a sense and an anti-sense block tiled on the array. When both sense and antisense blocks are present, the values are reported in the RAS1 and RAS2 columns.

A RAS of -1000 indicates the block does not exist or is not used to determine the call. A RAS of -9 indicates the block failed quality control.



12

Chapter 12



Hybridization Analysis

This chapter explains how to determine the cell intensities and background values for each probe on the GeneChip® GenFlex™ tag array. When the hybridization analysis is completed, the Hybridization Analysis window (HAW) displays the analysis output file (.chp).*

Running a Hybridization Analysis

To run a hybridization analysis, do one of the following:

- Right-click the desired cell intensity file (*.cel) in the data file tree and select **Analyze** from the shortcut menu.
- If the image (*.cel) is displayed, click the **Analyze** toolbar button  or select **Run** → **Analysis** from the menu bar.
- Select **File** → **Analyze** from the menu bar, then double-click the *.cel in the Analyze dialog box that appears.

The status log displays:

- the name of the analysis output file (*.chp)
- the location of the *.chp
- a message indicating when the analysis is completed

If the status bar is not displayed, click the **Status log** toolbar button  or select **View** → **Status Bar** from the menu bar.

When the analysis is finished, the Hybridization Analysis window (HAW) displays the output analysis file (*.chp) (Figure 12.1). If the HAW is already open, the results are added to the open window. It may be necessary to use the scroll bars at the bottom and right side of the HAW to see the newly added results.

		1109-A-05a				
		PM	MM	CPM	CMM	BG
ControlSet00001	GATTCACACTGACCCATGTA	261	68	11	15	115
ControlSet00002	TAAATAGATTGGAGACGCGC	171	30	4	1	119
ControlSet00003	GCATTAGAAGGCTCTGGACTA	94	11	6	6	115
ControlSet00004	ATTGGCATAACGTATTGCGC	93	19	2	6	115
ControlSet00005	CAGGACTGAAGATCGAGTAC	69	22	2	7	115
ControlSet00006	TAGAGTCAGTCATAGCTCGA	209	60	11	2	115
ControlSet00007	TTTATCGTAGCTGGCTGCC	127	21	5	21	115
ControlSet00008	AGGATTAGAACCTACGCACC	122	28	6	3	115
ControlSet00009	GCCGTGAGACCACTGTACTA	161	24	8	14	119
ControlSet00010	GACGCTGAATCCTATTGACA	175	44	13	10	119
ControlSet00011	CGCCTAAGGATCGTGAAGTA	178	61	11	11	115
ControlSet00012	CGACGACGAAGCTGCATGAA	507	138	7	14	115
ControlSet00013	ACTCGAATAACAGCATCTCG	222	53	5	9	115
ControlSet00014	CCCGTAAGCATGGCACAGAT	448	176	13	10	119

Figure 12.1
Hybridization Analysis window (HAW)

Each row in the HAW displays the probe set identifier and the perfect match (PM) sequence.

- PM The background-subtracted intensity for the perfect match probe that is a unique 20-mer oligonucleotide designed to be perfectly complementary to a reference sequence.
- MM The background-subtracted intensity for the mismatch probe that contains a mismatch to the PM at position 10.
- CPM The background-subtracted intensity for the probe complementary to the PM.
- CPM The background-subtracted intensity for the probe complementary to the MM.
- BG Background computed for the array.

Hybridization Analysis Window (HAW)

Exporting Data

You can export hybridization data to a tab delimited text file (*.txt).

1. Open the hybridization analysis output file (*.chp).
2. Select **File** → **Save As** from the menu bar.
⇒ The Export As dialog box appears ([Figure 12.2](#)).

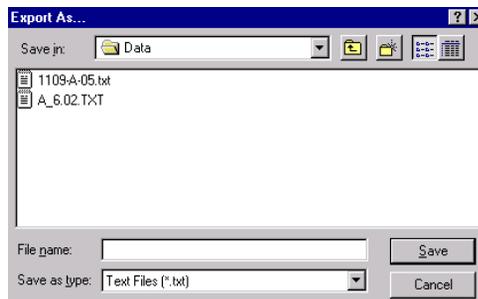


Figure 12.2
Export As dialog box

3. Enter a name for the text file in the **File name** box, then click **Save**.

Copying Data

You can copy all or selected portions of the HAW data table to the system clipboard and paste the copied data into other applications.

1. Highlight the data of interest in the table.
To select non-adjacent rows, press and hold the **Ctrl** key while you click the desired rows. To select adjacent rows, press and hold the **Shift** key while you click the first and the last row in the selection.
2. Click the **Copy** toolbar button . Alternatively, right-click the HAW, then click **Copy** in the shortcut menu or select **Edit** → **Copy** from the menu bar.

Removing Data

To remove experiments from the HAW data table (equivalent to closing the *.chp file), do one of the following:

- Right-click the HAW, then select **Remove Experiments** and the experiment name from the shortcut menu.
- Click the **Remove** toolbar button  or select **Edit** → **Remove Experiments** from the menu bar.
 - ⇒ The Select Experiments to Remove dialog box appears (Figure 12.3).
 - Choose the experiment name, then click **OK**.



Figure 12.3
Select Experiments to Remove dialog box

Find Feature

The Find feature searches the probe set identifier and the probe sequence columns.

1. Do one of the following:
 - Click the **Find** toolbar button .
 - Right-click the HAW, then click **Find** in the shortcut menu.
 - Select **Edit** → **Find** from the menu bar.
- ⇒ The Find dialog box appears (Figure 12.4).



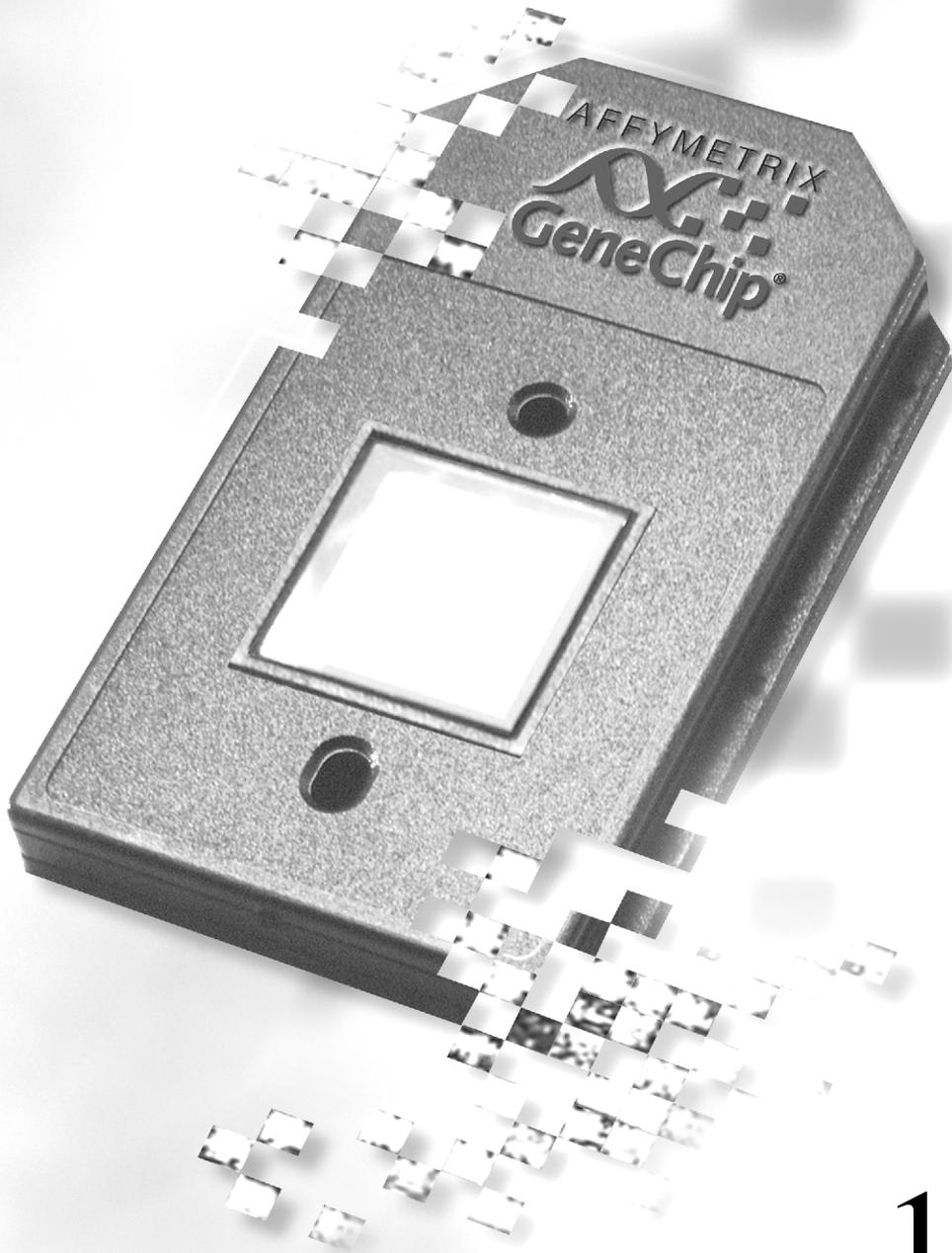
Figure 12.4
Find dialog box

2. Enter the desired text for the search (up to 256 alphanumeric characters) in the **Find What** box, then click **Find Next** to view the first search result.

 **Note**

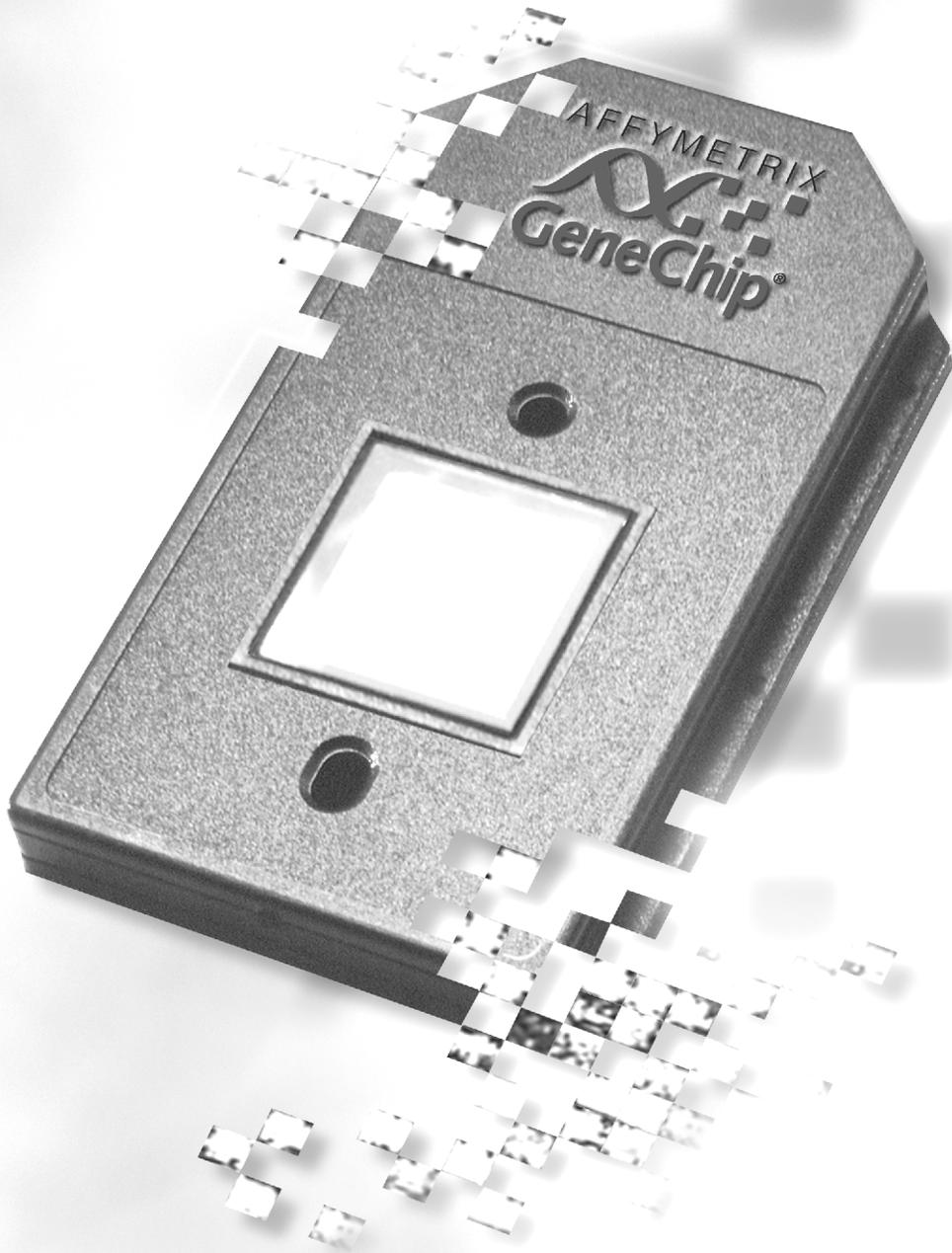
The **Find** command finds all strings that match the text string for the search in the probe set column or the probe sequence column. For example, using the **Find** command to search for the text string 12 would find ProbeSet 00012 and ProbeSet00112 as well as other occurrences of 12.

3. Continue to click **Find Next** to view each successive search result.



13

Chapter 13



Batch Analysis

Batch analysis is a convenient way to analyze the cell intensity data (.cel) from many different experiments and to generate analysis output files (*.chp) with unattended operation.*

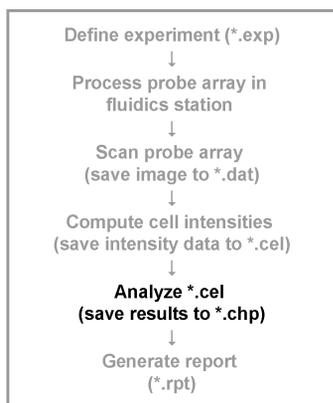


Figure 13.1
Assay & analysis flow chart

Opening the Batch Analysis Window

Do either of the following to open the Batch Analysis window:

- In the Microarray Suite shortcut bar, click **Batch Analysis** .
- Select **Run** → **Batch Analysis** from the menu bar.
 - ⇒ The Batch Analysis window opens ([Figure 13.2](#) and [Figure 13.3](#)).

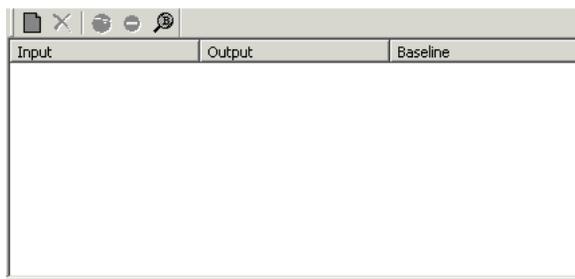


Figure 13.2
Batch analysis window, disk files mode

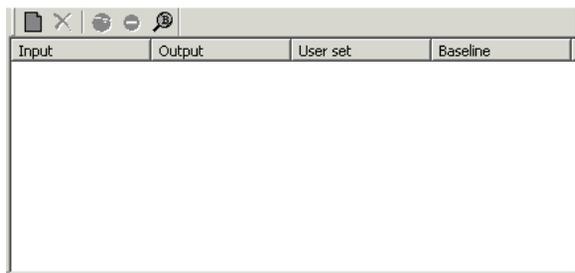


Figure 13.3
Batch Analysis window, LIMS mode

Selecting Files for Batch Analysis (Disk Files Mode)

In disk files mode, cell intensity file (*.cel) or analysis output (*.chp) from different probe array types may be selected for batch analysis. If you select *.chp files, Microarray Suite automatically places the parent *.cel files and baseline *.chp files (if applicable) in the batch analysis window.

Using the Drag-and-Drop Method

1. Drag the cell intensity file (*.cel) or analysis output file (*.chp) from the data file tree to the Batch Analysis window.
2. Repeat the drag-and-drop operation to select the remaining files for the batch analysis (Figure 13.5).

Using the Toolbar or Menu Commands

1. Click the **Add** toolbar button  or select **Edit** → **Add Item** from the menu bar.
⇒ An Open dialog of *.cel files appears (Figure 13.4).

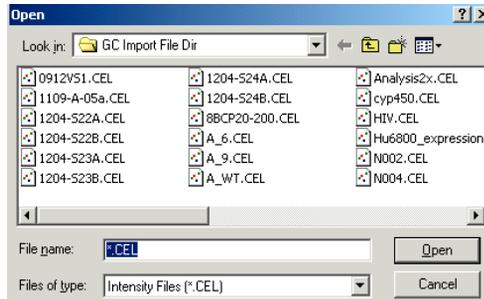


Figure 13.4
Open dialog box

2. Select the *.cel or *.chp files of interest.

To select contiguous files, press and hold the **Shift** key while you click the first and last file in the selection or press and hold the **Ctrl** key while you click noncontiguous files.

3. Click **Open**

⇒ This adds the selected *.cel files to the Batch Analysis window (Figure 13.5).

! CAUTION

Batch analysis will overwrite previous analysis output results (.chp).*

The batch analysis window displays previously analyzed *.cel files in red to indicate that the existing *.chp will be overwritten. To reanalyze experiments without overwriting existing *.chp files, specify a different name for the analysis output (*.chp) from the batch analysis. (See the section *Renaming Results from Batch Analysis*, on page 308.)

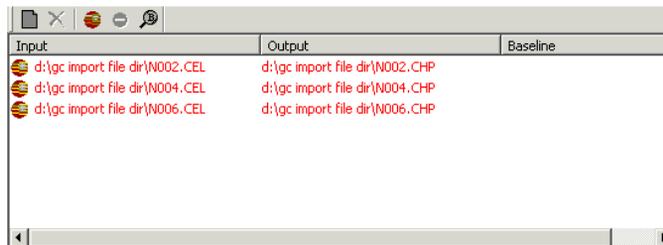


Figure 13.5
Batch Analysis window, disk files mode

Selecting the Baseline for Comparison Expression Analysis

1. In the Batch Analysis window, double-click in the baseline column of the *.cel file for the comparison expression analysis (Figure 13.6). Alternatively, click the *.cel file and choose **Edit** → **Select Baseline** from the menu bar.

⇒ An Open dialog box of *.chp files appears (Figure 13.7).

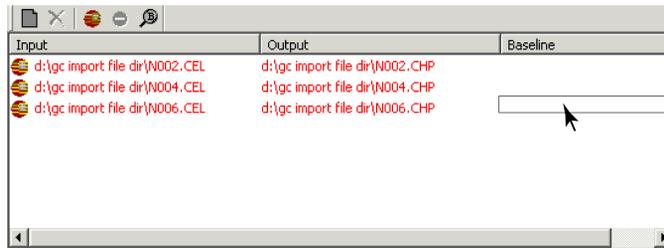


Figure 13.6
Batch Analysis window

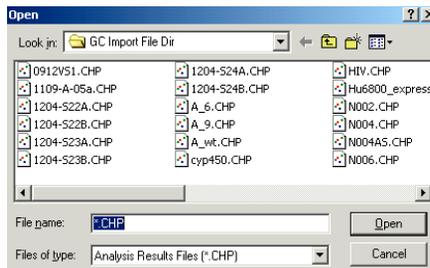


Figure 13.7
Open dialog box

2. Double-click the baseline *.chp file.
 - ⇒ The Batch Analysis window displays the baseline *.chp file (**Figure 13.8**).

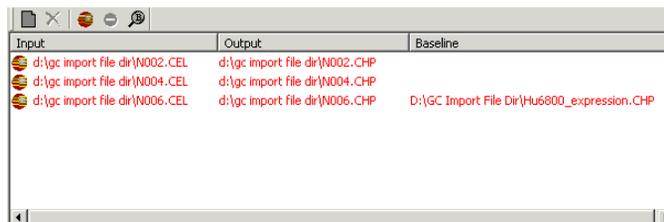


Figure 13.8
Batch Analysis window

3. To clear a baseline from the batch analysis window, right-click the baseline file and select **Clear Baseline** from the shortcut menu. Alternatively, click the baseline file and select **Edit** → **Clear Baseline** from the menu bar.

Selecting Files for Batch Analysis (LIMS Mode)

In LIMS mode, batch analysis is available for expression data only. If you select *.chp files, Microarray Suite automatically places the parent *.cel files and baseline *.chp files (if applicable) in the Batch Analysis window.

Using the Drag-and-Drop Method

1. Drag the cell intensity file (*.cel) or analysis output file (*.chp) from the data file tree to the Batch Analysis window.
2. Repeat the drag-and-drop operation to select the remaining files for the batch analysis (Figure 13.5).

Using the Toolbar or Menu Commands

1. Click the **Add** toolbar button  or select **Edit** → **Add Item** from the menu bar.
⇒ The Add Cell/Chip Files dialog box appears (Figure 13.9).
Click **Details** to display experiment and sample information for each file.

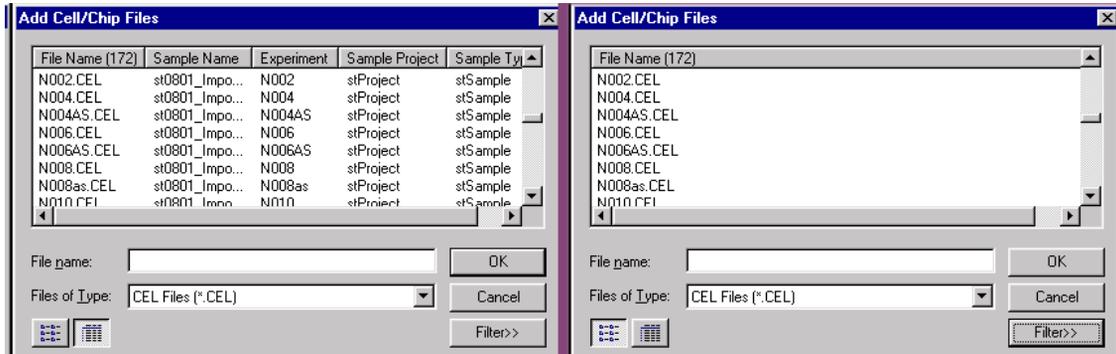


Figure 13.9
Add Cell/Chip Files dialog box, LIMS mode, Details view (left) and List view (right)

2. To filter the files in the Add Cell/Chip Files dialog box:
 - a. Click **Filter**.
⇒ This displays the advanced filter (**Figure 13.10**).
 - b. Make selections from the drop-down lists.
 - c. Click **Apply**.

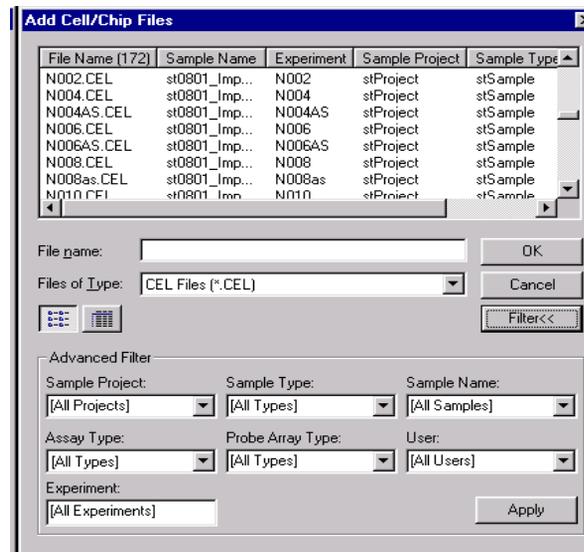


Figure 13.10
Add Cell File dialog box, advanced filter

3. Select the files for the batch analysis, then click **OK**.

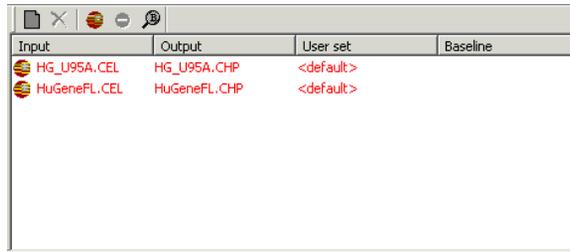
To select adjacent files, press and hold the **Shift** key while you click the first and last file in the selection. To select non-adjacent files, press and hold the **Ctrl** key while you click the files.

⇒ This adds the selected *.cel files to the Batch Analysis window (Figure 13.11).

! CAUTION

Batch analysis will overwrite previous analysis output results (.chp).*

The Batch Analysis window displays previously analyzed *.cel files in red to indicate that the existing *.chp will be overwritten. To reanalyze experiments without overwriting existing *.chp files, specify a different name for the analysis output (*.chp) from the batch analysis. (See the section *Renaming Results from Batch Analysis*, on page 308.)



Input	Output	User set	Baseline
HG_U95A.CEL	HG_U95A.CHP	<default>	
HuGeneFL.CEL	HuGeneFL.CHP	<default>	

Figure 13.11

Batch Analysis window (LIMS mode), *.cel files selected for batch analysis

Selecting a User Set for Expression Analysis

A user set specifies the expression analysis settings (scaling, normalization, probe masks, baseline files, values for user-modifiable expression algorithm parameters). See the Affymetrix® LIMS User Guide for more information about user sets.

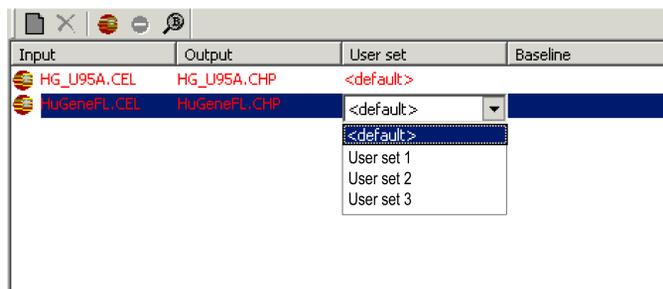
✓ Note

In a batch analysis, the settings for the default user set are specified in the Expression Analysis Settings dialog. The default user set is not the user set specified during experiment setup.

- To select other than the default user set, double-click the **User set** column.
⇒ This displays a drop-down list of available user sets ([Figure 13.12](#)).

✓ Note

If a user set with a baseline file is selected and the field under the baseline column is empty, the baseline file from the user set is used during the analysis.



Input	Output	User set	Baseline
HG_U95A.CEL	HG_U95A.CHP	<default>	
HuGeneFL.CEL	HuGeneFL.CHP	<default>	

Figure 13.12

Batch Analysis window (LIMS mode), drop-down list of User sets

2. Select a user set from the drop-down list. Alternatively, highlight the *.cel file of interest, select **Edit** → **Select User Set** from the menu bar, then make a selection from the drop-down list.

Selecting a Baseline for a Comparison Expression Analysis

1. In the Batch Analysis window, double-click in the baseline column of the *.cel file for the comparison expression analysis (Figure 13.13).

Alternatively, select the *.cel file of interest and choose **Edit** → **Select Baseline** from the menu bar.

⇒ The Select Baseline File dialog box appears (Figure 13.14).

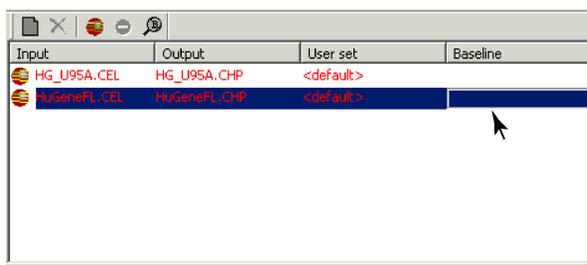


Figure 13.13
Batch Analysis window, LIMS mode

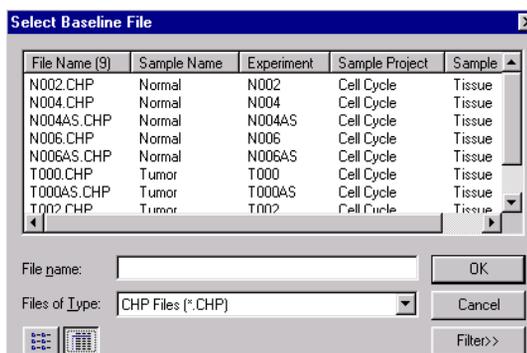
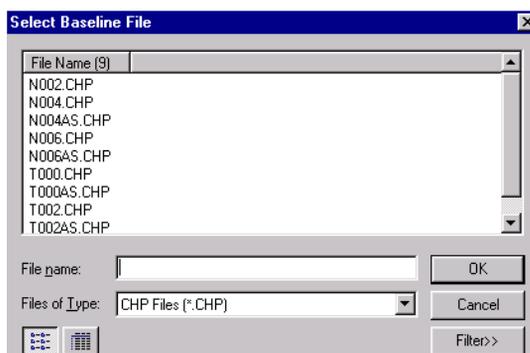


Figure 13.14
Select Baseline File dialog box, List view (left) and Details view (right)

2. Click **Details** to display experiment and sample information for each file.
3. To filter the files in the Add Cell File dialog box:
 - a. Click **Filter**.
⇒ This displays the advanced filter (**Figure 13.15**).
 - b. Make selections from the drop-down lists.
 - c. Click **Apply**.

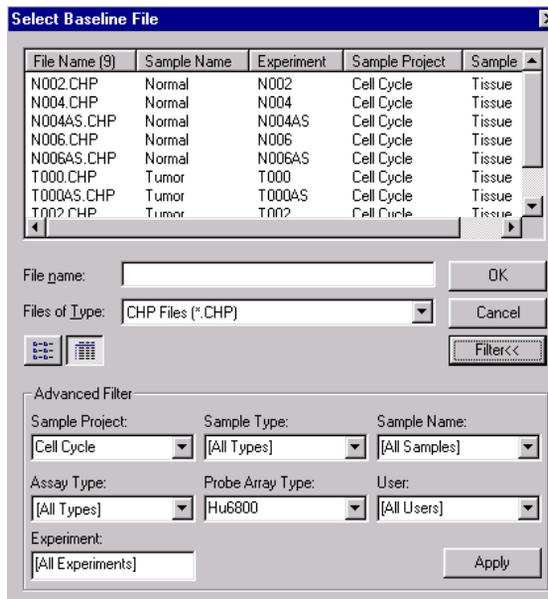
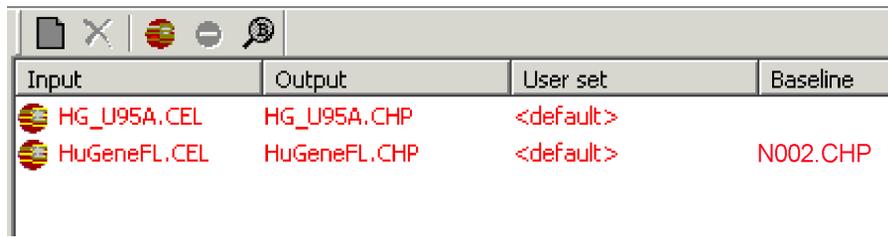


Figure 13.15
Select Baseline dialog box, advanced filter

4. Double-click the *.chp for the baseline.
⇒ The Batch Analysis window displays the baseline *.chp file (**Figure 13.16**).



Input	Output	User set	Baseline
HG_U95A.CEL	HG_U95A.CHP	<default>	
HuGeneFL.CEL	HuGeneFL.CHP	<default>	N002.CHP

Figure 13.16
Batch Analysis window (LIMS mode), baseline file selected

- To clear a baseline from the batch analysis window, right-click the baseline file and select **Clear Baseline** from the shortcut menu.

Alternatively, click the baseline file and select **Edit** → **Clear Baseline** from the menu bar.

Renaming Results from Batch Analysis

To reanalyze experiments without overwriting existing analysis output (*.chp), rename the output *.chp files from the batch analysis.

Renaming Results Automatically

If you select *.cel files for batch analysis, Microarray Suite can automatically rename all of the output *.chp files by adding a user-specified prefix and/or suffix to the file name.

✓ Note

*If you select *.chp files for batch analysis, the prefix or suffix are not applied to the output *.chp file name.*

- Click the **Options** toolbar button  or select **View** → **Options** from the menu bar.
⇒ The Batch Analysis Options dialog box appears ([Figure 13.17](#)).

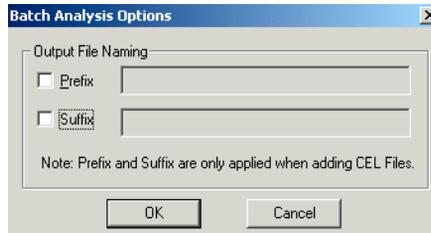


Figure 13.17
Batch Analysis Options dialog box

2. To add a prefix to the *.chp file name, choose the **Prefix** option and enter a prefix for the file name in the associated box.
3. To add a suffix to the *.chp file name, choose the **Suffix** option and enter a suffix for the file name in the associated box.
4. Click **OK**.
 - ⇒ When a *.cel is placed in the Batch Analysis window, the user-specified prefix and/or suffix is automatically added to the output *.chp file name.

Renaming Results Individually

You can individually rename output *.chp files from the batch analysis.

*Editing the Output *.chp File Name*

To edit the default *.chp file name, do either of the following:

- Double-click the *.chp file name in the Batch Analysis window and enter a new name.
- Select the *.cel file of interest in the Batch Analysis window, select **Edit** → **Rename Output File** from the menu bar, and enter a new name.

*Adding a Prefix or Suffix to the Output *.chp File Name*

1. In the Batch Analysis window, use one of the following methods to select one or more output *.chp.
 - To select all files, right-click a file and click **Select All** in the shortcut menu. Alternatively, select **View** → **Select All** from the menu bar.

- To select adjacent files, press and hold the **Shift** key while you click the first and last file in the selection.
 - To select non-adjacent files, press and hold the **Ctrl** key while you click the files.
- 2.** Right-click the selected *.chp(s) and select **Set Prefix/Suffix** from the shortcut menu.

Alternatively, click the selection and select **Edit** → **Set Prefix/Suffix** from the menu bar.

⇒ The Set Prefix/Suffix dialog box appears ([Figure 13.18](#)).



Figure 13.18
Set Prefix/Suffix dialog box

- 3.** Enter a prefix and/or suffix, then click **OK**.
- ⇒ The prefix and/or suffix are added to the selected output *.chp file name.

Removing Files from the Batch Analysis Window

To remove a *.cel file from the Batch Analysis window, do either of the following:

- Select the *.cel file and click the **Remove** toolbar button .
- Select the *.cel file, then select **Edit** → **Remove Item** from the menu bar.

Running a Batch Analysis

1. To run the batch analysis, click the **Analyze** toolbar button . Alternatively, select **Edit** → **Start Analysis** from the menu bar.
2. To stop a batch analysis in progress, click the **Stop** toolbar button . Alternatively, select **Edit** → **Stop Analysis** from the menu bar.

The software completes the current analysis and cancels the remaining files in the batch analysis queue.

Resuming a Batch Analysis

If the batch analysis is interrupted before all analyses are completed, Microarray Suite can resume the run and analyze the remaining *.cel files in the batch.

To resume a batch analysis:

1. Open the Batch Analysis window.
⇒ The Batch Analysis window displays the remaining unanalyzed files.
2. Click the **Analyze** toolbar button . Alternatively, select **Edit** → **Start Analysis** from the menu bar.

Exporting a Batch File

The Batch Analysis window contents can be exported to a text file (tab delimited *.txt). The batch file provides a record of the analysis and can be imported for a subsequent batch analysis.

1. Select **Edit** → **Export Batch** from the menu bar.
⇒ The Export Batch dialog box appears ([Figure 13.19](#)).

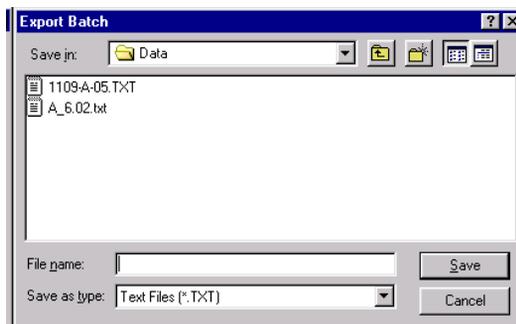


Figure 13.19
Export Batch dialog box

2. Enter a name for the batch file, then click **Save**.
⇒ This creates the batch file (*.txt).

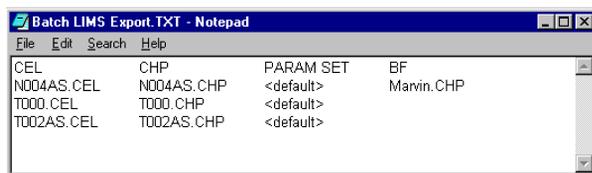
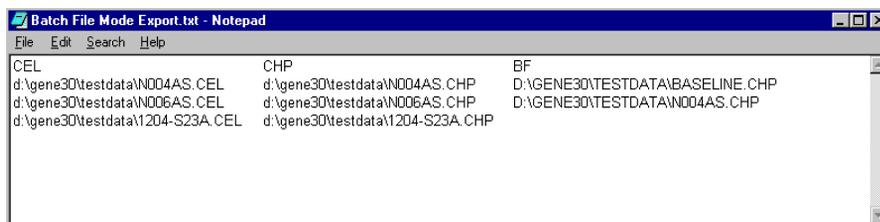


Figure 13.20
Batch file (*.txt), disk files mode (top) and LIMS mode (bottom)

Importing a Batch File

Microarray Suite can import a batch file (tab delimited *.txt) that follows the format shown in [Figure 13.20](#).

To import a batch file:

1. Open the Batch Analysis window.
2. Select **Edit** → **Import Batch** from the menu bar.
⇒ The Import Batch dialog box appears ([Figure 13.21](#)).

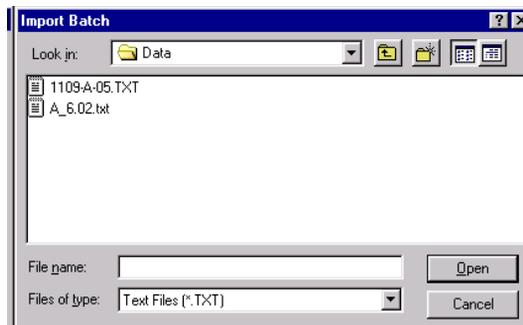
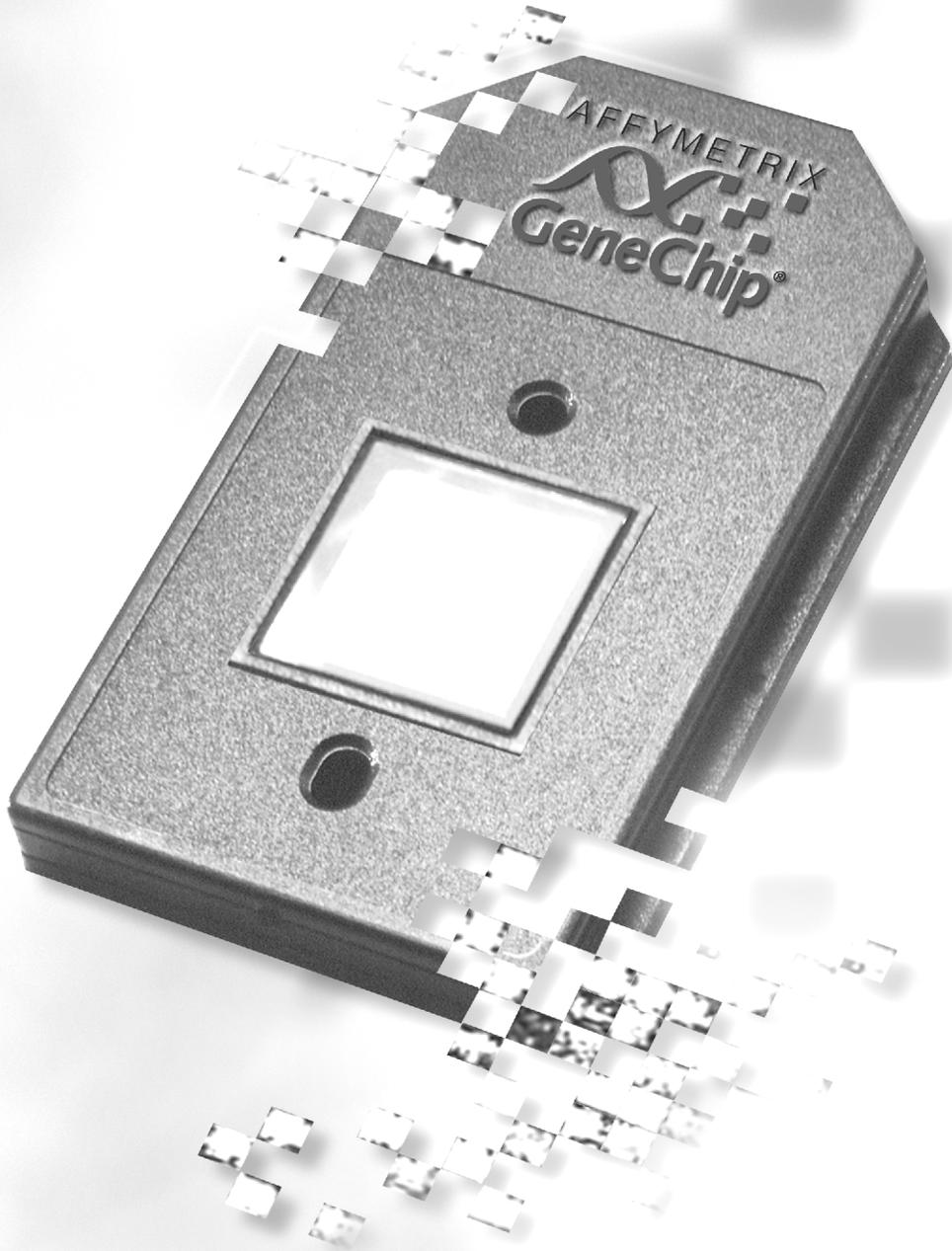


Figure 13.21
Import Batch dialog box

3. Double-click the batch file (*.txt) you wish to import.
⇒ This imports the batch file to the Batch Analysis window.



Chapter 14



In LIMS mode, Microarray Suite can copy sample and experimental data (.exp, *.cel, and *.chp) from the process database to a publish database. This is called publishing. Published data can be queried by the Affymetrix® Data Mining Tool (DMT) software or several other third party analysis tools.*

Note

*In disk files mode, use Affymetrix® MicroDB® to publish data to a database on the local workstation. (See the Affymetrix® MicroDB® User's Guide.) If MicroDB® is installed on the workstation, select **Run** → **MicroDB** from the menu bar to start the program.*

Selecting the Publish Database

1. In the Microarray Suite shortcut bar, click **Publish** . Alternatively, select **Run** → **Publish** from the menu bar.
⇒ This displays the publish window (Figure 14.1).

The top windowpane displays all publish databases. The bottom windowpane displays a status log of tasks.

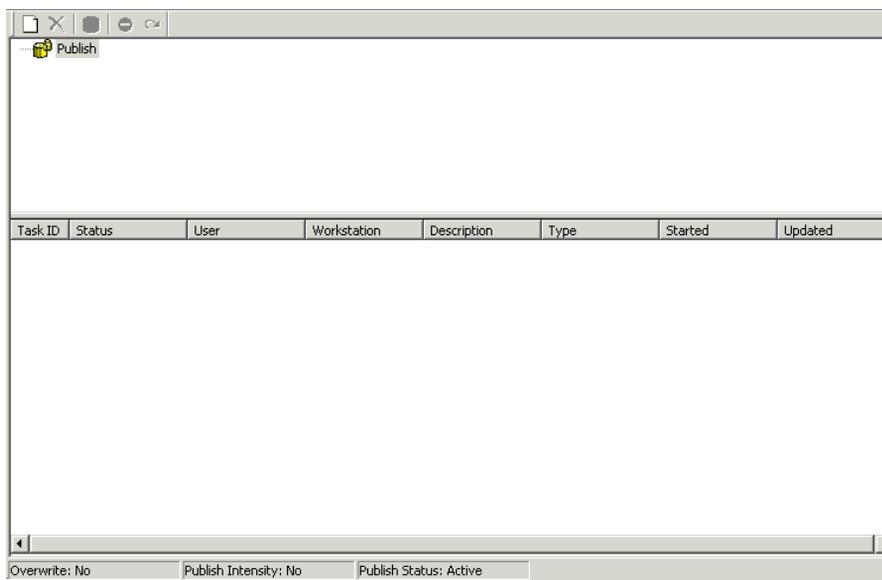


Figure 14.1
Publish window

2. Click the publish database destination for the data.

A yellow icon indicates the currently selected database. A gray icon indicates an unselected database.



Note

Each database requires a login password before publishing can proceed.

Specifying the Task

A task consists of all data and files selected for publishing. Each individual file or experiment to be published is called a task item. [Table 14.1](#) shows the data published for each type of task item.

Table 14.1
Data published for different types of task items

Task Item	Files Published
Experiment *.exp	*.exp All *.cel files associated with the *.exp. All *.chp files associated with the *.exp.
Image data file *.dat	*.exp All *.cel files associated with the *.exp. All *.chp files associated with the *.exp.
Cell intensity file *.cel	The associated *.exp. The *.cel file. All *.chp files associated with the *.cel file.
Analysis output file *.chp	The associated *.exp. The *.chp file. All *.cel files associated with the *.chp.

To select the tasks, use either of the following methods:

Using the Drag-and-Drop Method

1. Drag the desired task item from the data file tree on to a database icon in the top pane of the publish window.
2. Repeat the drag-and-drop operation for the remaining task items to be published ([Figure 14.4](#)).

Using the Add Toolbar Button

1. Click the **Add** toolbar button  or select **Publish** → **Add Item** from the menu bar.
⇒ This displays the Open dialog box ([Figure 14.2](#)).

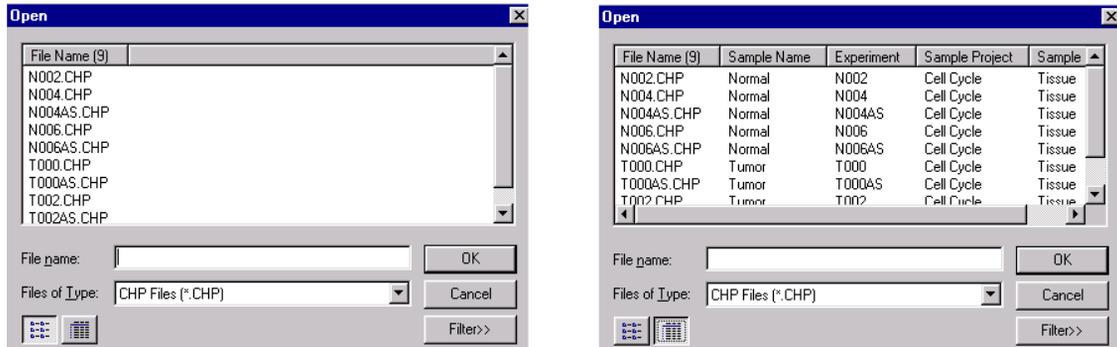


Figure 14.2
Open dialog box, list view (left) and details view (right)

2. Click **Details** to display experiment and sample information for each file.
3. To filter the list of files in the Open dialog box, click **Filter**
⇒ This displays the Filter dialog box (**Figure 14.3**).

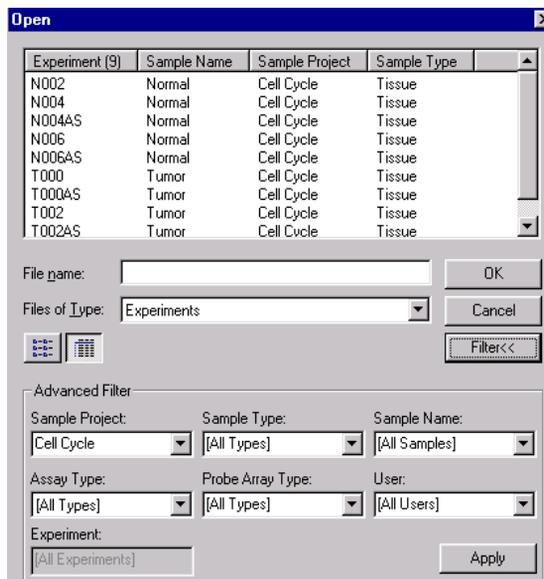


Figure 14.3
Open dialog box, advanced filter

4. Make selections from the drop-down lists and click **Apply** to filter the list.
5. Make a selection from the **Files of Type** drop-down list and select the files to be published.

To select adjacent files, press and hold the **Shift** key while you click the first and last file in the selection. To select non-adjacent files, press and hold the **Ctrl** key while you click the files.

6. Click **OK** to add the files to the selected publish database (Figure 14.4).
7. To remove a task item from the publish window, select the task item in the publish window and click the **Remove** button . Alternatively, choose **Publish** → **Remove Item** from the menu bar.

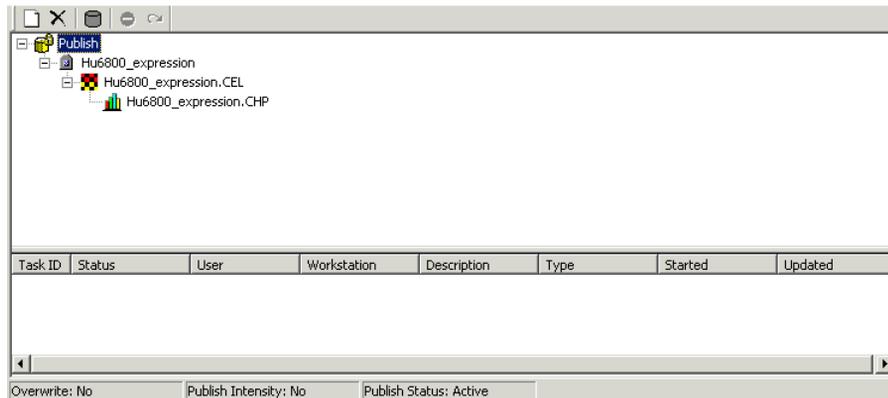


Figure 14.4
Publish window

Publishing Options

You may publish intensity data or overwrite a previously published task item. You must be authorized by the LIMS administrator to overwrite data. The status bar at the bottom of the publish window displays the publishing options (Figure 14.4).

Publishing Intensity Data

Intensity data (*.cel files) are not automatically published. Publishing cell intensity data (*.cel) uses large amounts of computer memory. For example, publishing the analysis output data (*.chp) for a high density probe array requires approximately 2 MB of disk space compared to 30 MB for both the *.chp and *.cel data.



Note

The image data files (.dat) are not published.*

1. To include cell intensity (*.cel) data in the publish database, select **Publish** → **Options** → **Publish Intensities** from the menu bar.
⇒ The **Publish Intensities** menu item is check marked.
2. To not include *.cel data in the publish database, select **Publish** → **Options** → **Publish Intensities** from the menu bar to remove the check mark.

Overwriting Data

A previously published task item can be overwritten. Users must be authorized by the LIMS administrator to overwrite data.

- Select **Publish** → **Options** → **Overwrite Data** from the menu bar.

Publishing, Canceling, or Restarting a Task

Publishing the Task

1. After you have specified the task, click the Publish button . Alternatively, select **Publish** → **Publish** from the menu bar.
 - ⇒ The Publish Login dialog box appears if this is the first time during the session that an item has been published to the database ([Figure 14.5](#)).



Figure 14.5
Publish Database Login

2. Enter the password for the database.
 - ⇒ The task is displayed in the task log and will be published during the time frame specified by the LIMS administrator.

Canceling a Task

You can cancel tasks waiting to be published (the status in the task log is WAIT).

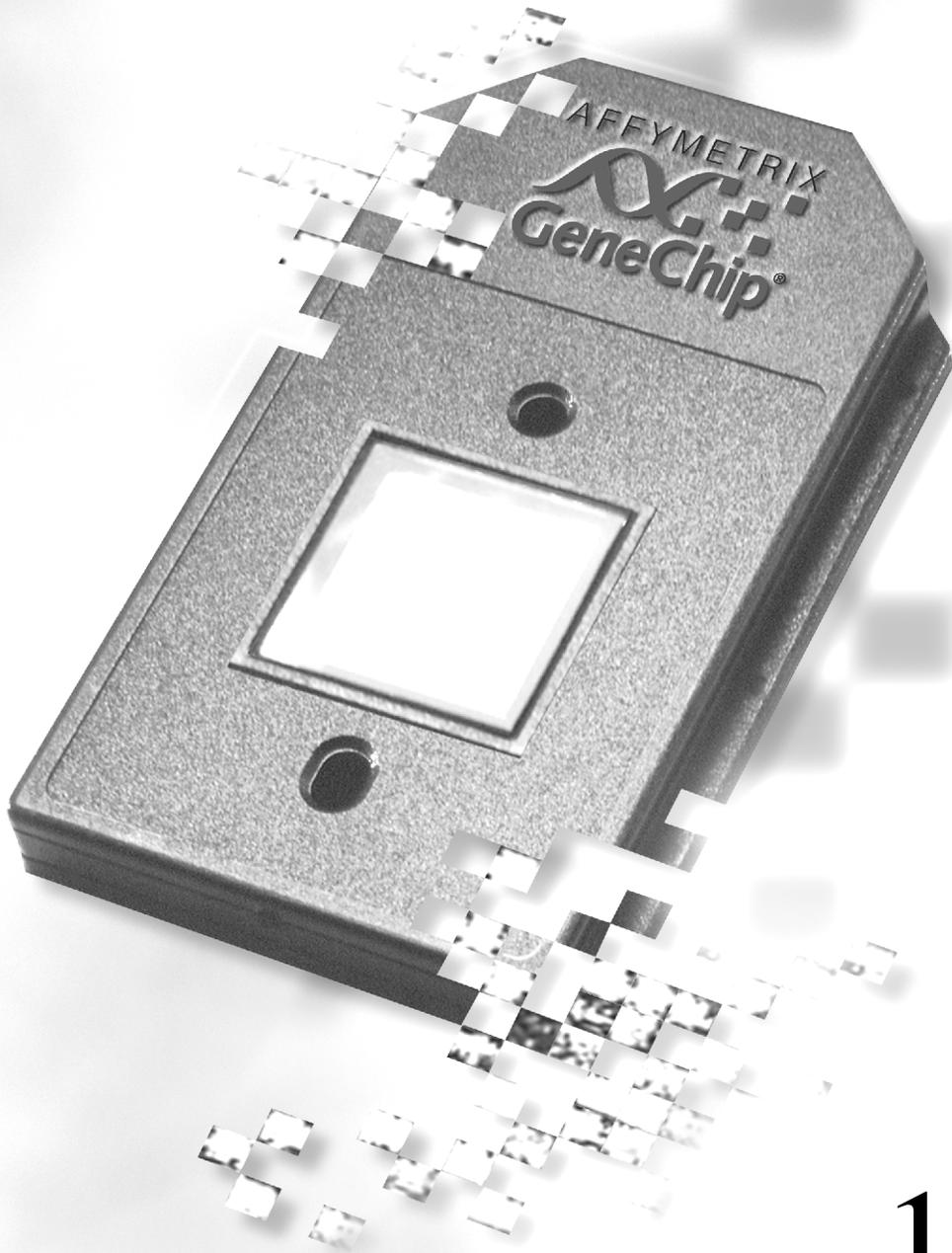
1. Select the task in the task log.
2. Click the **Cancel** toolbar button . Alternatively, select **Publish** → **Monitor** → **Cancel Publish** from the menu bar.
 - ⇒ The task is now canceled (the status is CANCELED).

Restarting a Publishing Task

You can reinstated or *restarted* a canceled task and return it to the publishing queue.

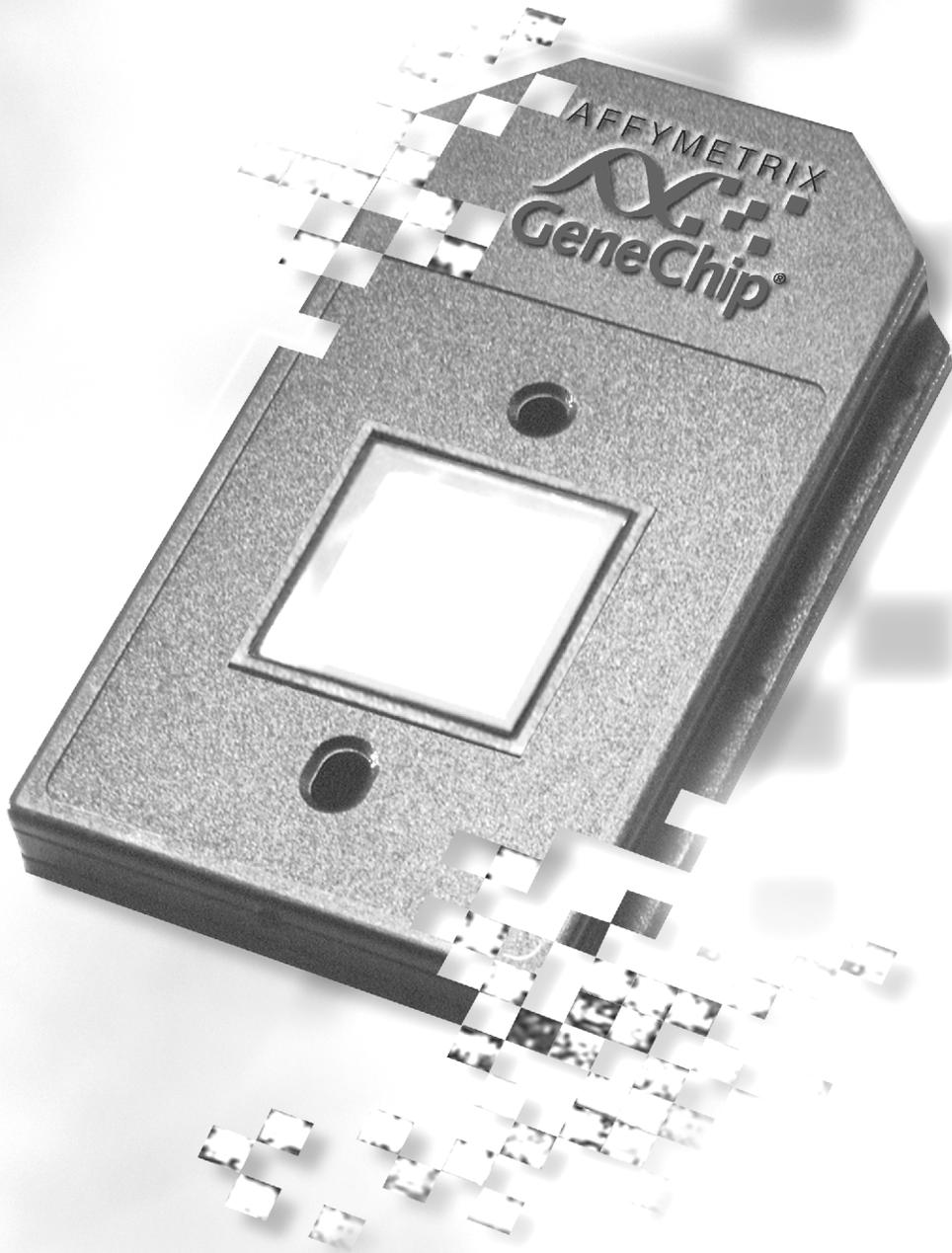
1. Select the canceled task in the task log.
2. Click the **Restart** toolbar button . Alternatively, select **Publish** → **Monitor** → **Restart Publish** from the menu bar.

The task is now restarted (the status is WAIT) and will be published during the time frame specified by the LIMS administrator.



15

Chapter 15



The Microarray Suite software generates a report from the analysis output file (.chp) (Figure 15.1). A report summarizes information about hybridization intensity data and user-modifiable algorithm settings (if any).*

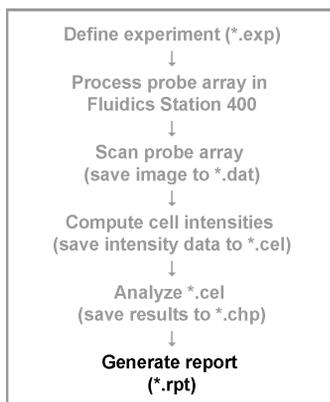


Figure 15.1
Assay & analysis flow chart

This chapter describes the six types of reports:

- Expression
- Mapping
- GT Analysis Viewer
- HIV PRT Plus
- CYP450
- p53

Report Settings

Some reports have user-modifiable settings (see [Table 15.1](#)).

Table 15.1
Reports & user-modifiable settings

Report	User-Modifiable Settings	For More Information, See...
Expression	Yes	Chapter 9
Mapping	No	Chapter 11
GT Viewer	No	Chapter 11
HIV PRT Plus	Yes	Chapter 10
CYP450	No	Chapter 11
p53	Yes	Chapter 10

Expression Report

The expression report summarizes information about the expression analysis settings, algorithm settings, and probe set hybridization intensity data.

1. To generate the report, do either of the following:
 - In the data file tree, right-click the analysis output file (*.chp), then select **Report** from the shortcut menu.
 - Select **File** → **Report** from the menu bar, then double-click an analysis output file (*.chp) in the Report dialog box that appears ([Figure 15.2](#)).

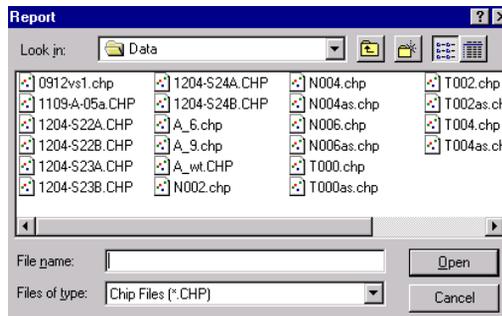


Figure 15.2
Report dialog box

⇒ This displays the expression report in the main display area.
Figure 15.3 and **Figure 15.4** show an example expression report.

2. To best view the expression report:
 - a. Use Arial font (10 point) (select **View** → **Set Font** from the menu bar).
 - b. Set the tab stops set at 10 (select **View** → **Set Tab Stops** from the menu bar).

✓ Note

*A previously generated report will be overwritten. To save a previous report, rename it and save it under the new name. (See the section **Saving an Expression Report.**)*

Report Type: Expression Report
Date: 11:00AM 06/13/2001

Filename: N006AS.CHP
Probe Array Type: Hu6800
Algorithm: Statistical
Probe Pair Thr: 8
Controls: Antisense

Alpha1: 0.04
Alpha2: 0.06
Tau: 0.015
Noise (RawQ): 34.840
Scale Factor (SF): 1.000
Norm Factor (NF): 0.910

Baseline filename: N004AS.CHP
Gamma1H: 0.0025
Gamma1L: 0.0025
Gamma2H: 0.0030
Gamma2L: 0.0030
Perturbation: 1.22
Baseline Noise (RawQ): 36.920
Baseline Scale Factor (SF): 1.000000

Background:				
	Avg: 826.54	Std: 18.25	Min: 788.10	Max: 869.20
Noise:				
	Avg: 26.71	Std: 1.06	Min: 25.00	Max: 30.40
Corner+				
	Avg: 24217	Count: 32		

Figure 15.3
Expression report, comparison analysis, page 1

The following data represents probe sets that exceed the probe pair threshold and are not called "No Call".

Total Probe Sets:	7056	
Number Present:	1878	26.6%
Number Absent:	4988	70.7%
Number Marginal:	190	2.7%
Average Signal (P):	1202.8	
Average Signal (A):	113.7	
Average Signal (M):	294.0	
Average Signal (All):	408.4	
Number Increase:	94	1.3%
Number Decrease:	64	0.9%
Number MIncrease:	6	0.1%
Number MDecrease:	3	0.0%
Number No Change:	6889	97.6%
Number (A/M->P, M/I)	23	0.3%
Number (P->A/M, MD/D)	12	0.2%
Number (P->P, M/I)	63	0.9%
Number (P->P, MD/D)	41	0.6%
Number (A/M->A/M, M/I)	14	0.2%
Number (A/M->A/M, MD/D)	14	0.2%
Number (P->P, NC)	1468	20.6%
Number (A/M->A/M, NC)	4941	70.0%
Number Increase:		
(0<=SLR<1):	73	
(SLR>=1):	27	
(SLR>=2):	15	
(SLR>=3):	2	
(SLR>=4):	0	
Number Decrease:		
(0>=SLR>=-1):	38	
(SLR<=-1):	29	
(SLR<=-2):	9	
(SLR<=-3):	5	
(SLR<=-4):	0	

Housekeeping Controls:

Probe Set	Sig(5)	Det(5)	Sig(M)	Det(M)	Sig(3)	Det(3)	Sig(all)	Sig(3/5)
HUMISGF3A/M97935	20.4	A	154.6	A	195.3	P	123.42	9.56
HUMRGE/M10098	10.6	A	30.7	A	34.9	A	25.42	3.29
HUMGAPDH/M33197	15088.9	P	14182.4	P	13367.5	P	14219.58	0.89
HSAC07/A00351	16005.2	P	26001.9	P	14967.8	P	18991.62	0.94
HUMTFRR/M11507	120.7	A	55.8	A	231.0	P	135.85	1.91
M27830	23.9	A	229.0	A	99.4	A	117.43	4.16

Spike Controls:

Probe Set	Sig(5)	Det(5)	Sig(M)	Det(M)	Sig(3)	Det(3)	Sig(all)	Sig(3/5)
BIOB	53.0	A	23.4	A	82.0	A	52.78	1.55
BIOC	481.1	P			293.9	P	387.48	0.61
BIODN	970.7	P			3778.4	P	2374.53	3.89
CREX	5405.5	P			7817.0	P	6611.24	1.45
DAPX	2146.1	P	2199.2	P	2794.3	P	2379.86	1.30
LYSX	2217.5	P	4278.5	P	3546.6	P	3347.55	1.60
PHEX	2495.5	P	3087.7	P	3865.5	P	3149.54	1.55
THRX	3749.3	P	2322.5	P	3804.9	P	3292.25	1.01
TRPNX	18.9	A	8.9	A	14.5	A	14.12	0.77

Other Controls:

Probe Set	Sig	Det
AFFX-BIOB-5_AT	53.0	A
AFFX-BIOB-M_AT	23.4	A
AFFX-BIOB-3_AT	82.0	A
AFFX-BIOB-5_ST	44.6	A
AFFX-BIOB-M_ST	39.8	A
AFFX-BIOB-3_ST	17.8	A

Figure 15.4

Expression report, comparison analysis, page 2

Expression Report Components

Statistical Expression Algorithm

Probe Pair Threshold	The minimum number of probe pairs a probe set must have in order for the probe set data to be included in the calculation of the report statistics
Alpha1	<p>The significance level for the detection p-value in an absolute analysis. Alpha1 is a user-modifiable parameter that is set in the Parameters tab of the Expression Analysis Settings (see Appendix D).</p> <p>If the probe set detection p-value $<$ alpha1, the call is present.</p>
Alpha2	<p>The second significance level for the detection p-value in an absolute analysis. Alpha2 is a user-modifiable parameter set in the Parameters tab of the Expression Analysis Settings (see Appendix D).</p> <p>If the probe set detection p-value \geq alpha2, the call is absent. If $\text{alpha1} \leq \text{detection p-value} < \text{alpha2}$, the call is marginal.</p>
Tau	Tau is a user-modifiable parameter that is set in the Parameters tab of the Expression Analysis Settings (see Appendix D). Ideally, tau should be set to a value that is a little larger than the median of the discrimination scores of the probe sets whose targets are absent to avoid false detected calls.
Noise	The degree of pixel-to-pixel variation among the probe cells used to calculate the background (see the Appendix C).
Scale Factor (SF)	The scale factor specified in the Scaling tab of the Expression Analysis Settings dialog box or computed by the algorithm. (See Appendix D).

TGT Value	The user-specified target signal for scaling of the experiment probe array. The target intensity is set in the Scaling tab of the Expression Analysis Settings dialog box (see Appendix D).
Norm Factor (NF)	The normalization factor specified in the Normalization tab of the Expression Analysis Settings dialog box or computed by the algorithm (see Appendix D).
Gamma1H	The small significance level for change calls at high intensities. Gamma1H is a user-modifiable parameter that is set in the Parameters tab of the Expression Analysis Settings (see Appendix D).
Gamma2H	The large significance level for change calls at high intensities. Gamma2H is a user-modifiable parameter that is set in the Parameters tab of the Expression Analysis Settings (see Appendix D).
Gamma1L	The small significance level for change calls at low intensities. Gamma2L is a user-modifiable parameter that is set in the Parameters tab of the Expression Analysis Settings (see Appendix D).
Gamma2L	The large significance level for change calls at low intensities. Gamma2L is a user-modifiable parameter that is set in the Parameters tab of the Expression Analysis Settings (see Appendix D).
Perturbation	A user-modifiable expression algorithm parameter that is set in the parameters tab of the Expression Analysis Settings (see Appendix D). Perturbation influences the p-value computed for a probe set in a comparison analysis.
Noise (Raw Q)	The degree of pixel-to-pixel variation among the probe cells used to calculate the background in the baseline probe array. (See Appendix C.)
Baseline Scale Factor (SF)	The scale factor specified for the baseline probe array in the Scaling tab of the Expression Analysis Settings dialog box or computed by the algorithm. (See Appendix D.)

Background	Minimum, maximum, average, and standard deviation of the background intensity of the noise calculated for the probe array.
Noise	The minimum, maximum, average, and standard deviation of the noise calculated for the probe array.
Corner+	The average cell intensity for the sense probe cells used in the grid alignment process.
Corner-	The average cell intensity of the antisense probe cells used in the grid alignment process.
Central+	The average cell intensity for the nine probe cells that compose the cross at the center of a sense probe array.
Central-	The average cell intensity for the nine probe cells that compose the cross at the center of an antisense probe array.
Total Probe Sets	The number of probe sets on the array that exceed the probe pair threshold and are not called "No Call."
Average Signal	The average signal for all probe sets that exceed the probe pair threshold and are not called <i>No Call</i> .
Controls	The expression report includes the signal and call data for the probe sets that correspond to the housekeeping or spike control transcripts. Separate signal and call data are reported for the probe pairs specific to the 5', middle (M'), and 3' regions of the control transcripts.
Sig(all)	The average signal for all control probe sets.
Sig(3'/5')	For a probe set, Sig(3')/Sig(5').

Mapping Report

The mapping report for the GeneChip® HuSNP™ probe array (Figure 15.5) includes the percentage of SNPs assigned to a genotype and the percentage of each genotype call for the array as well as per primer pool. This provides both a global and a localized assessment of assay performance.

1. To generate the report, right-click a HuSNP™ analysis output file (*.chp) in the data tree and select **Report** from the shortcut menu.
⇒ This displays the report in the main display area.
2. To best view the HuSNP report:
 - a. Use Arial font (10 point) (select **View** → **Set Font** from the menu bar).
 - b. Set the tab stops set at 10 (select **View** → **Set Tab Stops** from the menu bar).

ReportType: Mapping Report
 Date: 11:21AM 06/19/2001

Filename: HuSNP4B.CHP
 Probe Array Type: HuSNP
 Algorithm: Genotyping
 Alg version: 4.0
 Alg parameters: NumStdCutoff=1.50 SCutoff=1.30 MinMNum=3

Probe Array Summary:
 CorrelationCoefficient=0.999
 MaxIntensity=15807.0
 MedianIntensity=737.8

%Pass 88.3
 %A 28.8
 %AB 29.3
 %B 28.9
 %AB_A 0.6
 %AB_B 0.7
 %NoSignal 11.7

Pool Summary:

Pool	%Pass	%A	%AB	%B	%AB_A	%AB_B	%NoSignal
A01	97.9	28.7	44.7	24.5	0.0	0.0	2.1
A02	96.5	40.0	29.4	25.9	0.0	1.2	3.5
A03	89.2	27.7	28.9	32.5	0.0	0.0	10.8
A04	85.5	26.3	23.7	34.2	0.0	1.3	14.5
A05	86.1	27.8	25.0	33.3	0.0	0.0	13.9
A06	84.5	31.0	32.4	18.3	1.4	1.4	15.5
A07	80.0	36.0	25.3	18.7	0.0	0.0	20.0
A08	78.8	27.5	26.3	22.5	1.3	1.3	21.3
A09	90.2	28.0	26.8	32.9	1.2	1.2	9.8
A10	96.5	27.1	30.6	37.6	0.0	1.2	3.5
A11	81.4	25.6	24.4	29.1	1.2	1.2	18.6
A12	97.4	39.0	28.6	28.6	1.3	0.0	2.6
A13	95.5	34.1	29.5	30.7	0.0	1.1	4.5
A14	91.4	24.3	31.4	35.7	0.0	0.0	8.6
A15	95.9	27.4	34.2	31.5	2.7	0.0	4.1
A16	92.1	21.1	38.2	31.6	1.3	0.0	7.9
A17	91.0	30.8	34.6	25.6	0.0	0.0	9.0
A18	70.0	14.0	12.0	44.0	0.0	0.0	30.0
A19	92.2	35.3	27.5	25.5	2.0	2.0	7.8
A20	69.2	15.4	28.2	23.1	0.0	2.6	30.8
A21	59.0	25.6	12.8	20.5	0.0	0.0	41.0
Total	88.3	28.8	29.3	28.9	0.6	0.7	11.7

Figure 15.5
 Mapping report

Mapping Report Components

% Pass	The percentage of SNPs assigned a genotype.
% A	The percentage of homozygous A genotypes.
% AB	The percentage of heterozygous AB genotypes.
% B	The percentage of homozygous B genotypes.
% AB_A	The percentage of genotypes that could be AB or A (the two possible calls could not be distinguished).
% AB_B	The percentage of genotypes that could be AB or B (the two possible calls could not be distinguished).
% NoSignal	The percentage of genotypes that could not be called.

CYP 450 Report

The CYP450 report includes the:

- name and location of the analysis output file (*.chp)
 - probe array type and experiment name
 - genotypes for the CYP2D6 and CYP2C19 gene
1. To generate the report, right-click a CYP450 analysis output file (*.chp) in the data tree and select **Report** from the shortcut menu.
 - ⇒ This displays the report in the main display area (Figure 15.6).
 2. To best view the CYP450 report:
 - a. Use Courier New font (10 point) (select **View** → **Set Font** from the menu bar).
 - b. Set the tab stops set at 14 (select **View** → **Set Tab Stops** from the menu bar).

```

Algorithm:                               BlockRatioCall
Alg parameters:                           None
cyp450 Gene:CYP2D6
Genotype:CYP2D6*2/CYP2D6*1A

Discrimination    ALLELE LOOKUP TABLE
Quality           1 2 3 4A 4B 4C 4D 4E 6A 6B 7 8 9 10A 10B 11
C188T             WT           | | | | +| +| +| +| +| | | | | | + | + | |
-----
G971C             WT           | | | | | | | | | | | | | | | | | +|
-----
A1072G           WT           | | | | +| +| | | | | | | | | | | | |
-----
C1127T           WT           | | | | | | | +| | | | | | | | | + | |
-----
1795DT           WT           | | | | | | | | | | +| +| | | | | | | |
-----
G1749C           HET           | +| | | +| | +| +| +| | | | +| | + | + | +|
-----
G1846T           WT           | | | | | | | | | | | | +| | | | | | |
-----
G1934A           WT           | | | | +| +| +| +| +| | | | | | | | | |
-----
G2064A           WT           | | | | | | | | | | | +| | | | | | | |
-----
2637DA           WT           | | +| | | | | | | | | | | | | | | | |
-----
2701-2703DACA   WT           | | | | | | | | | | | | | | +| | | | |
-----
C2938T           HET           | +| | | | | | | +| | | | +| | | | | +|
-----
A3023C           WT           | | | | | | | | | | | | | +| | | | | | |
-----
T3975C           WT           | | | | | | +| | | | | | | | | | | | |
-----
G4268C           HET           | +| | | +| +| +| +| +| | | | +| | + | + | +|
-----

cyp450 Gene:CYP2C19
Genotype:CYP2C19*1

Discrimination    ALLELE LOOKUP TABLE
Quality           2 3
CYP2C19-M1       WT           |+| |
-----
CYP2C19-M2A      WT           | | |
-----
CYP2C19-M2B      WT           | |+|
-----

```

Figure 15.6
CYP450 report

CYP450 Report Components

Polymorphism	Describes the polymorphisms detected by the GeneChip® CYP450 probe array (for example, C188T indicates a base change from C to T at nucleotide position 188; DAGA indicates the triple base deletion AGA).
Call	Displays the genotype call for each polymorphism (WT = wildtype, HET = heterozygous, MUT = homozygous mutant, or N = no call).
Discrimination Quality	A measure of how well the cell intensity values discriminate between the wildtype and non-wildtype probes in a probe set. The column displays Low if the discrimination quality is low for WT, HET, or MUT calls.

The two Allele Lookup tables display the mutations that define the alleles of the CYP2D6 and CYP2C19 gene. The tables may be used to confirm the genotype calls made by the software. The first row of the table lists the allele names. In the CYP2D6 Allele Lookup table, 1 is the wildtype, 2 is allele 2 of the CYP2D6 gene (or CYP2D6*2), 3 is allele 3 of the CYP2D6 gene (or CYP2D6*3), and so on.

The column associated with the allele name contains a + for each polymorphic mutation required to call that particular allele. For example, in the CYP450 report shown in [Figure 15.6](#), allele 2 is defined by 3 mutations: G1749C (G changes to C at nucleotide position 1749), C2938T, and G4268C. Since the target contained a mutation at each of these polymorphic locations (the call is heterozygous), the haplotype is CYP2D6*2. The remaining haplotype is wildtype (CYP2D6*1A) and the genotype is CYP2D6*2 / CYP2D6*1A).

HIV PRT *Plus* Report

The HIV PRT *Plus* Report summarizes GeneChip® HIV PRT *Plus* probe array data and assay results derived from the analysis output file (*.chp).

1. To generate the HIV PRT *Plus* report, right-click the analysis output file (*.chp) in the data file tree and select **Report** from the shortcut menu.
⇒ The Select Report dialog box appears ([Figure 15.7](#)).



Figure 15.7
Select Report dialog box

2. Click **HIV PRT Plus Report** and click the **OK** button.
⇒ This generates the report and displays it in the main display area ([Figure 15.8](#)).
3. To best view the HIV PRT *Plus* report:
 - a. Use Arial font (10 point) (select **View** → **Set Font** from the menu bar).
 - b. Set the tab stops set at 14 (select **View** → **Set Tab Stops** from the menu bar).

Report Type: HIV PRT Plus Report
Date: 10:16AM 07/25/2001

Filename: HIV.CHP
Probe Array Type: HIV PRTPlus 2
Algorithm: Rules
Filter Type: All Codon Changes

SampleID	Gene	AAChange	CodonChange	SAW Reference
HIV	Protease	P1-	cct->ntt	41
HIV	Protease	Q2-	cag->nag	44
HIV	Protease	I3F	atc->ttt	47
HIV	Protease	T4-	act->ant	50
HIV	Protease	L5-	ctt->ntt	53
HIV	Protease	W6-	tgg->tng	56
HIV	Protease	Q7-	caa->naa	59
HIV	Protease	R8-	cga->nga	62
HIV	Protease	P9-	ccc->anc	65
HIV	Protease	L10V	ctc->gtn	68
HIV	Protease	V11-	gtc->nan	71
HIV	Protease	T12-	aca->atn	74
HIV	Protease	K14-	aag->tng	80
HIV	Protease	I15	ata->att	83
HIV	Protease	G17R	ggg->aga	89
HIV	Protease	K20	aag->aaa	98
HIV	Protease	A22-	gct->nnt	104
HIV	Protease	L23I	cta->ata	107
HIV	Protease	T26-	aca->ana	116
HIV	Protease	A28-	gca->gna	122
HIV	Protease	D30E	gat->gaa	128
HIV	Protease	T31	aca->act	131
HIV	Protease	V32E	gta->gaa	134

Figure 15.8
HIV PRT Plus Report

HIV PRT Plus Report Components

Sample ID	The name of the analysis output file (*.chp).
Gene	The HIV-1 gene that carries the mutation.
AA Change	The amino acid change resulting from the codon change (for example, I3T indicates isoleucine changed to threonine at amino acid position 3). An asterisk (*) indicates a stop codon.
Codon Change	The codon change resulting from the base mutation.

SAW Reference The nucleotide position of the base mutation.

p53 Report

The p53 report summarizes GeneChip® p53 probe array data and assay results from the analysis output file (*.chp).

1. To generate the p53 report, right-click the analysis output file (*.chp) in the data tree and select **Report** from the shortcut menu.
⇒ The Select Report dialog box appears ([Figure 15.9](#)).



Figure 15.9
[Select Report dialog box](#)

2. Click **P53 Report** and click **OK** to generate the report ([Figure 15.10](#)).
⇒ The report is displayed in the main display area.
3. To best view the P53 report:
 - a. Use Arial font (10 point) (select **View** → **Set Font** from the menu bar).
 - b. Set the tab stops set at 14 (select **View** → **Set Tab Stops** from the menu bar).

Filter Type: Mutations Only
Date: 10:23AM 07/25/2001

Filename: A_9.CHP
Probe Array Type: GP53
Algorithm: Mixture Detection

SampleID	Location	AA Change	Codon Change	Score	SAW Reference
A_9	p53 intron1		g->c	11	4
A_9	p53 exon4	P75L	g->g cct->ctt	0,10,0	284,285,286
A_9	p53 exon4	A78P	cct->cct gca->cca	4,0,0	293,294,295
A_9	p53 exon5	R156I	gca->gca cgc->atc	12,14,0	561,562,563
A_9	p53 exon5	V157D	cgc->cgc gtc->gac	0,9,0	564,565,566
A_9	p53 exon5	M160I	gtc->gtc atg->atc	0,0,19	573,574,575
A_9	p53 exon6		atg->atg g->-	7	817
A_9	p53 exon6	P222H	ccg->cac	0,12,0	830,831,832
A_9	p53 exon7	C229S	ccg->cgc tgt->tca	9,12,0	861,862,864
A_9	p53 exon7	G244T	tgt->gta ggc->acg	11,2,0	918,919,920
A_9	p53 exon7		ggc->gcg t->c	9	956
A_9	p53 exon8		t->t g->-	6	1027
A_9	p53 exon8		g->t	10	1051
A_9	p53 exon8		g->g g->a	10	1072
A_9	p53 exon8		g->g g->c	11	1119
A_9	p53 exon8		g->g		

Figure 15.10
p53 report

P53 Report Components

Sample ID	The name of the analysis output file (*.chp).
Location	The location of the mutation on the p53 gene.
AA Change	The amino acid change resulting from the codon change (for example, at p53 exon 4, P75L indicates proline changed to leucine at amino acid position 75). No amino acid change results from a base change that occurs in an intron or after a stop codon.
Codon Change	The codon change resulting from the base mutation(s) or change(s) in intron regions or after a stop codon. Two entries for the same location of the p53 gene (for example, cct ctt and cct cct at location p53 exon 4) indicate a mixture of mutant and wildtype base at that location. A dash sign (-) indicates a single base deletion.
Score	The probe set score(s) for a mutation(s) at a given nucleotide location(s). The score is a sum of the variables that increase in value with increasing fractions of mutant base in the target. The score value increases as the amount of mutant signal increases.
SAW Reference	The nucleotide position of a single base mutation or deletion or the three nucleotide positions of a codon that contains one or more mutated bases or deletions.

Genotyping Viewer Analysis Report

The Genotyping (GT) Viewer analyzes Affymetrix® HuSNP™ *.cel files to generate the GT Viewer report that provides a *relative allele signal* (RAS) for each HuSNP™ marker. The RAS, a metric derived from the observed hybridization signals, is a quantitative representation of the two possible alleles. It is expressed as a value between zero and one.

1. For information on how to generate the GT Viewer report, see [Selecting Cell Intensity Files for the GT Viewer Report](#), on page 282.
2. To best view the Genotyping Analysis Viewer report:
 - a. Use Arial font (10 point) (select **View** → **Set Font** from the menu bar).
 - b. Set the tab stops at 15 (select **View** → **Set Tab Stops** from the menu bar).

CEL Filenames (1st Pair): R1_030801A.CEL, R1_030801B.CEL
 CEL Filenames (2nd Pair): R2_030801A.CEL, R2_030801B.CEL
 Probe Array Type: HuSNP
 Algorithm: Genotyping
 Alg parameters: NumStdCutoff=1.50 SCutoff=1.30 MinMNum=3

Marker	Chromosome	EstGenDist	R1_030801			R2_030801			R1_030801_R2_030801	
			Call	RAS1	RAS2	Call	RAS1	RAS2	Delta RAS	StdUnits
WIAF-3819	20	22.00	A	1.000	-1000	A	1.000	-1000	0.000	-1000
WIAF-3821	4	125.80	B	0.000	-1000	B	0.000	-1000	0.000	-1000
WIAF-3818	4	125.80	B	0.032	-1000	B	0.076	-1000	0.043	-1000
WIAF-3271	2	180.30	AB	0.671	-1000	AB	0.726	-1000	0.055	1.658
WIAF-2729			AB	0.731	-1000	AB	0.768	-1000	0.037	0.970
WIAF-548	5	102.50	NoSignal	-9	-1000	NoSignal	-9	-1000	-9	-1000
WIAF-267	5	121.60	B	0.114	-1000	NoSignal	0.191	-1000	0.078	-1000
WIAF-272	10	31.50	NoSignal	0.228	-1000	NoSignal	0.269	-1000	0.041	-1000
WIAF-1918	12	122.80	B	0.089	-1000	B	0.141	-1000	0.053	-1000
WIAF-1904	22	41.80	A	0.880	-1000	A	0.921	-1000	0.041	-1000
WIAF-1362	8	62.50	B	0.052	-1000	B	0.091	-1000	0.039	-1000
WIAF-1634	9	64.10	B	0.000	-1000	B	0.039	-1000	0.039	-1000
WIAF-2460			A	0.991	-1000	A	1.000	-1000	0.009	-1000
WIAF-2184	19	42.20	B	0.000	-1000	B	0.000	-1000	0.000	-1000
WIAF-2189	7	5.20	A	0.963	-1000	A	1.000	-1000	0.037	-1000
WIAF-2180	X	11.10	AB	0.557	-1000	AB	0.584	-1000	0.027	0.789
WIAF-3540	6	47.00	NoSignal	0.948	-1000	A	0.929	-1000	0.018	-1000
WIAF-3542	10	31.80	A	0.693	-1000	A	0.746	-1000	0.053	-1000
WIAF-3537	10	17.30	NoSignal	0.932	-1000	A	0.903	-1000	0.029	-1000
WIAF-3272	1	160.50	B	0.013	-1000	B	0.010	-1000	0.003	-1000
WIAF-3268	13	77.10	B	0.151	-1000	B	0.198	-1000	0.048	-1000
WIAF-2999			NoSignal	0.055	-1000	B	0.080	-1000	0.025	-1000
WIAF-2730			AB	0.487	-1000	AB	0.515	-1000	0.028	1.461

Figure 15.11
Genotyping analysis viewer report

GT Viewer Analysis Report Components

Marker	The identifier for the SNP.
Chromosome	The number or name of the chromosome that contains the SNP.
EstGenDist	The estimated chromosomal location of the SNP expressed in centiMorgans. Genetic distances are measured from the top of the chromosome's short arm. A large set of markers including the SNPs were physically mapped onto radiation hybrid panels. The physical locations of the SNP markers were integrated into the genetic map to obtain the Estimated Genetic Distances.
Call	The algorithm generates six possible calls: (1) A (homozygous A allele), (2) B (homozygous B allele), (3) AB (heterozygous), (4) AB_A (two possible genotypes AB or A that could not be distinguished), (5) AB_B (two possible genotypes AB or B that could not be distinguished), or (6) No Signal (insufficient data passed the quality tests to perform an analysis).
RAS 1	The relative allele signal for the first used block of the SNP.
RAS 2	The relative allele signal for the second used block of the SNP, if a second block exists. Note: a second block may exist, but may not be used for making a genotype call.
Delta RAS	The difference between the control and experimental RAS for each HuSNP™ marker.
StdUnits	The number of heterozygote standard deviations that the experiment RAS value is from the control RAS value, assuming a normal distribution of RAS values.

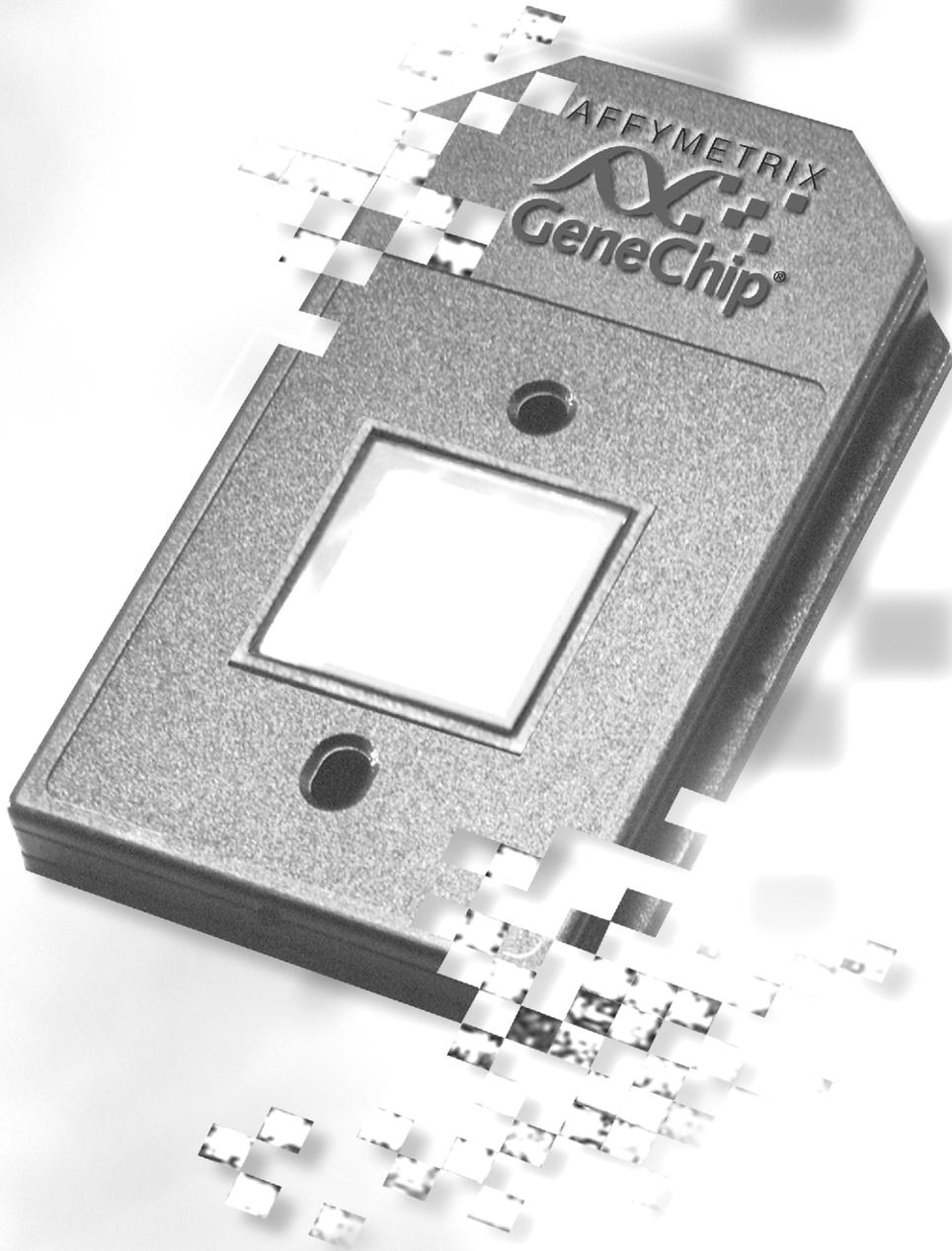
 Note

A HuSNP™ marker may have a sense and an anti-sense block tiled on the array. Not every marker has both sense and anti-sense blocks, and not every block is used to determine the call.

A RAS of -1000 indicates the block does not exist or is not used to determine the call. A RAS of -9 indicates the block failed quality control. If both blocks are present and used in the analysis, a new RAS value is derived from RAS1 and RAS2 (See Appendix I for more information.)



Chapter 16



This chapter describes how to change the factory set defaults for:

- image settings
- file locations
- fluidics station configuration
- analysis prompts
- experiment data storage

Viewing Default Settings

- Select **Tools** → **Defaults** from the menu bar.
⇒ This displays the Defaults dialog box ([Figure 16.1](#)).

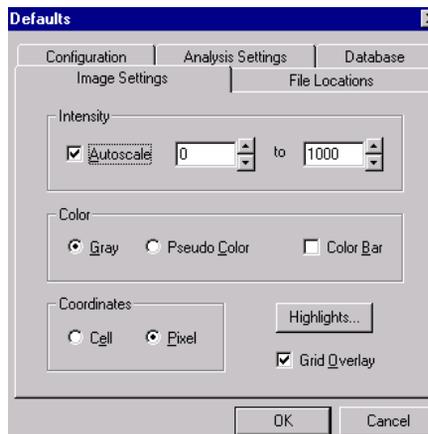


Figure 16.1
Defaults dialog box, Image Settings tab

Image Settings

The image settings ([Figure 16.1](#)) affect the display of image data (*.dat) and cell intensity (*.cel) files in the Image window. (See [Chapter 8](#) for more information about the Image window.)



Note

*The image settings specified in the Defaults dialog box (except for the color option) are applied to subsequently opened *.dat or *.cel files and do not affect an active, open image.*

Intensity Range

The Agilent GeneArray® Scanner software has a dynamic intensity range from 0 to 65,000 and provides 256 colors from black through white for image display. The intensity range of an image is divided into 256 bins and each bin is assigned a color or grayscale.

You may enter a new lower or upper limit for the intensity range associated with an image and apply this subset of the dynamic range to the image. Lowering the upper limit increases the image brightness and raising the lower limit decreases the brightness.

Using this *image scaling* process, you can adjust the display of a data or cell intensity file for optimum viewing. Alternatively, Microarray Suite can automatically scale the image using the minimum and maximum pixel intensities of the image.

1. To automatically scale an image, choose the **Autoscale** option in the Defaults dialog box ([Figure 16.1](#)) and click **OK**.
2. In the **Image Settings** dialog box for the current *.dat or *.cel file, you may click **Defaults** to return the intensity parameters (Autoscale option and Intensity range) to the settings in the Image Settings tab of the **Defaults** dialog box. (See [Making Adjustments to the Current Image Only, on page 122](#) for more information about the Image Settings dialog box).

Color

You can view the pixel intensity values in gray scale or pseudo color (generated by applying rainbow colors to the intensity scale). Choose the **Color Bar** option (Figure 16.1) to display a gray scale or pseudo color bar at the top of the Image window.

Coordinates

When you click the image, a red cross hairs  marks the location. The status bar at the lower left of the main window displays the intensity data and x,y coordinates of a probe cell or pixel indicated by the cross hairs.

If you choose the **Cell** option in the Defaults dialog box (Figure 16.1), a pop-up tool tip and the status bar display the x,y coordinates of the cell, the cell intensity, and the pixel intensity (Figure 16.2).

If you choose the **Pixel** option, the pop-up tool tip and the status bar display the x,y coordinates of the pixel and the pixel intensity (Figure 16.2).

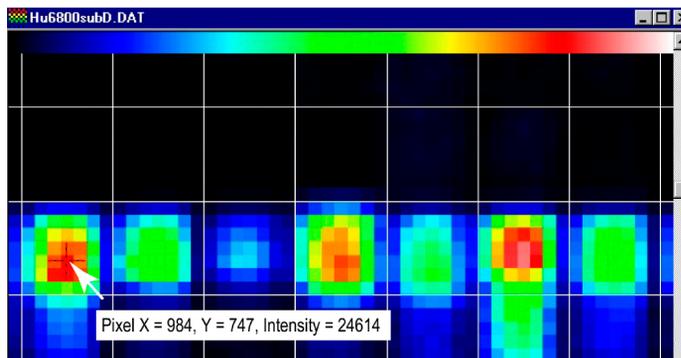
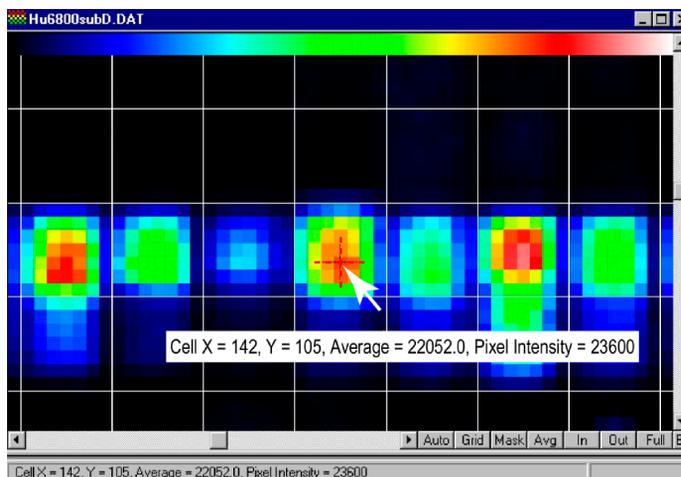


Figure 16.2
Image data file (*.dat), probe cell coordinates (top) and pixel coordinates (bottom)

Highlight Colors

You may specify new highlighting colors for the grid, masked cells, outlier cells, or highlighted probe array cells (features).

1. Click **Highlights** in the Defaults dialog box (Figure 16.1).
 ⇒ The Highlight Colors dialog box appears (Figure 16.3).

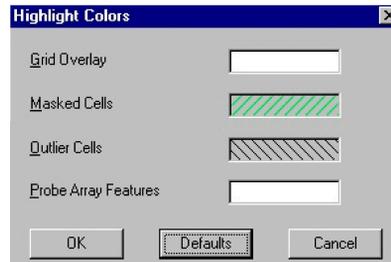


Figure 16.3
Highlight Colors dialog box

2. To change the color of a particular item (for example, the grid overlay):
 - a. Click the associated color box in the Highlight Colors dialog box.
 ⇒ This displays the Color palette (Figure 16.4).

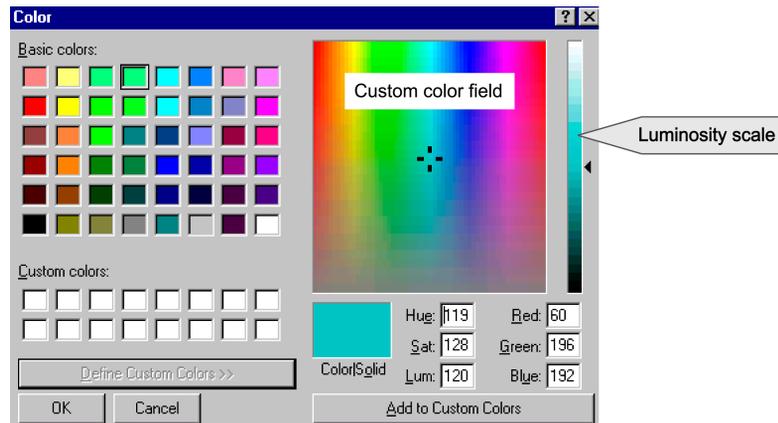


Figure 16.4
Color palette

- b. To select a predefined color, click one of the basic colors.
- c. To define a custom color, click **Define Custom Colors**, then use the click-and-drag method to move the cross hairs in the custom color field. Adjust the color brightness using the luminosity scale to the right. When finished, click **Add to Custom Colors** to apply the color.
 - ⇒ This applies the new color to the selected item.
- d. Click **OK** to close the Color palette.

Grid Overlay

Choose the **Grid Overlay** option ([Figure 16.1](#)) to automatically display the grid on a *.dat or *.cel file. Remove the check mark to toggle the grid overlay off. Alternatively, click **Grid** in the Image window task bar (or press the **G** key) to toggle the grid on or off.

Database

The Database tab ([Figure 16.5](#)) specifies how Microarray Suite manages experiment data files (*.exp, *.dat, *.cel, *.chp).

Choose:

- **Disk Files** option to store the experiment data files on the workstation hard drive (using Microarray Suite in *disk files mode*).
- **LIMS** option to store the experiment data files on the server (using Microarray Suite in *LIMS mode*).

Note

The Experiment Data Storage options are only available if all windows are closed and no instruments are active.

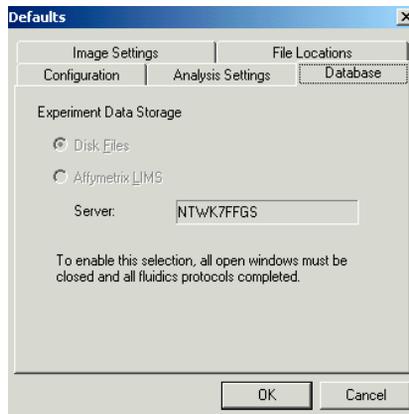


Figure 16.5
Defaults, Database tab

File Locations

The experiment data storage option you select in the Database tab (disk files or LIMS) determines the file locations. The File Locations tab defines the location of the fluidics protocols and the probe information files that Microarray Suite uses to analyze data. It also specifies where to save experiment data.

If the experiment data are stored on disk files, there are three file locations (Figure 16.6). If the data are stored on a network drive, there are four locations (Figure 16.8).

The default file locations are stored on a per user basis and some may be changed. Changes made by one user (identified by the logon name) do not affect the file locations of other users.

Experiment Data Stored on the Workstation

If you choose the **Disk Files** option in the Database tab (Figure 16.5), the experiment data are stored on the workstation hard drive (Figure 16.6).

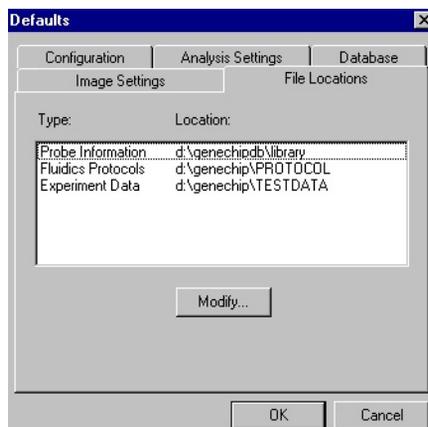


Figure 16.6
Defaults, File Locations tab, disk files mode

Modifying a File Location

1. Highlight the desired file type and click **Modify**.
⇒ The Modify Location dialog box appears (Figure 16.7).

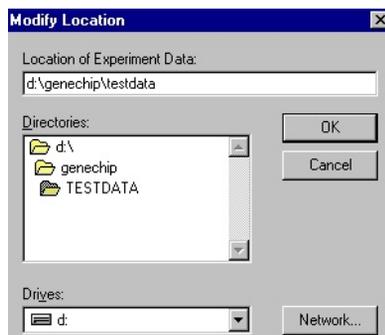


Figure 16.7
Modify Location dialog box

2. Choose a drive from the **Drives** drop-down list and a directory path from the **Directories** box.
3. Click **Network** to map network drives to drive letters.

Experiment Data Stored on the LIMS Server

If you choose the **Affymetrix LIMS** option in the Database tab (Figure 16.5), the experiment data are stored on the server (Figure 16.8). A scanned image is stored on the local hard drive during image capture. After the image is acquired, the image data file (*.dat) is copied to the server.



Note

Only the local scan data and protocols directory may be modified when using Microarray Suite in the LIMS mode.

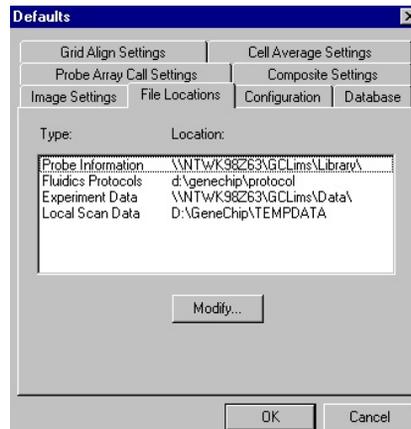


Figure 16.8
Defaults, File Locations tab, LIMS mode

1. To modify the Fluidics Protocols or Local Scan Data location, highlight the desired file type and click **Modify**.
⇒ The Modify Location dialog box appears (Figure 16.9).

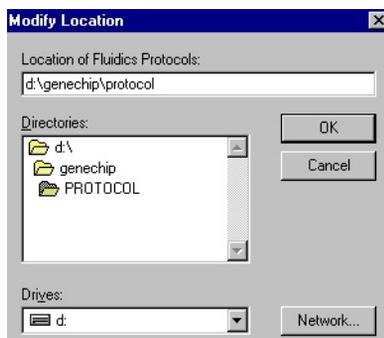


Figure 16.9
Modify Location dialog box (LIMS mode)

2. Choose a drive from the **Drives** drop-down list and a directory path from the **Directories** box. Click **OK**.

The local scan data must be set when the experiment data is pointing to a network drive.

3. Click **Network** to map network drives to drive letters.
4. Click **OK**.

Configuration

The Configuration tab (Figure 16.10) specifies the number of fluidics stations installed on the system. Up to eight fluidics stations may be simultaneously controlled by one workstation.

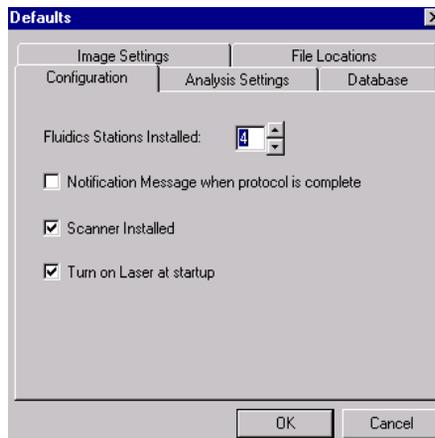


Figure 16.10
Defaults, Configuration tab

Notification Message when
protocol is complete

Choose this option to display a notification
of completion at the end of a protocol

Scanner Installed

Choose this option after the scanner is
installed.

Turn on Laser at startup

Confirm this option is chosen (default) so
that the Agilent GeneArray® Scanner
automatically turns on when Microarray
Suite is started.

Analysis Settings

The Analysis Settings tab (Figure 16.11) displays two options that are available when an analysis is run.

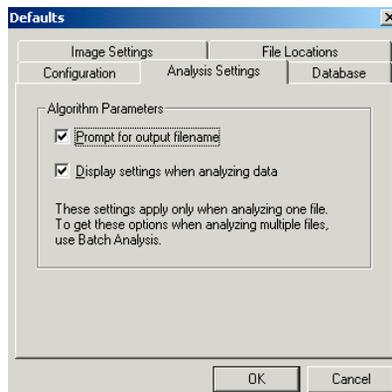


Figure 16.11
Defaults, Analysis Settings tab

Prompt for Output File Name

1. Choose the **Prompt for output file name** option (Figure 16.11) to display a Save Results As dialog box (Figure 16.12) at the start of an analysis.

The default name for the analysis output file (*.chp) is the same as the experiment information file name (*.exp). The Save Results As dialog box displays the default name for the *.chp file.



Figure 16.12
Save Results As dialog box

2. Click **OK** to keep the default file name. Alternatively, enter a new name and click **OK**.

Display Settings When Analyzing Data

Choose the **Display settings when analyzing data** option (Figure 16.11) to display the Expression Analysis Settings dialog box (Figure 16.13) at the start of an analysis.

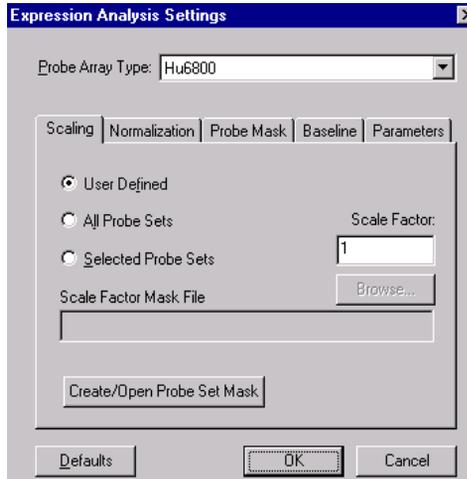
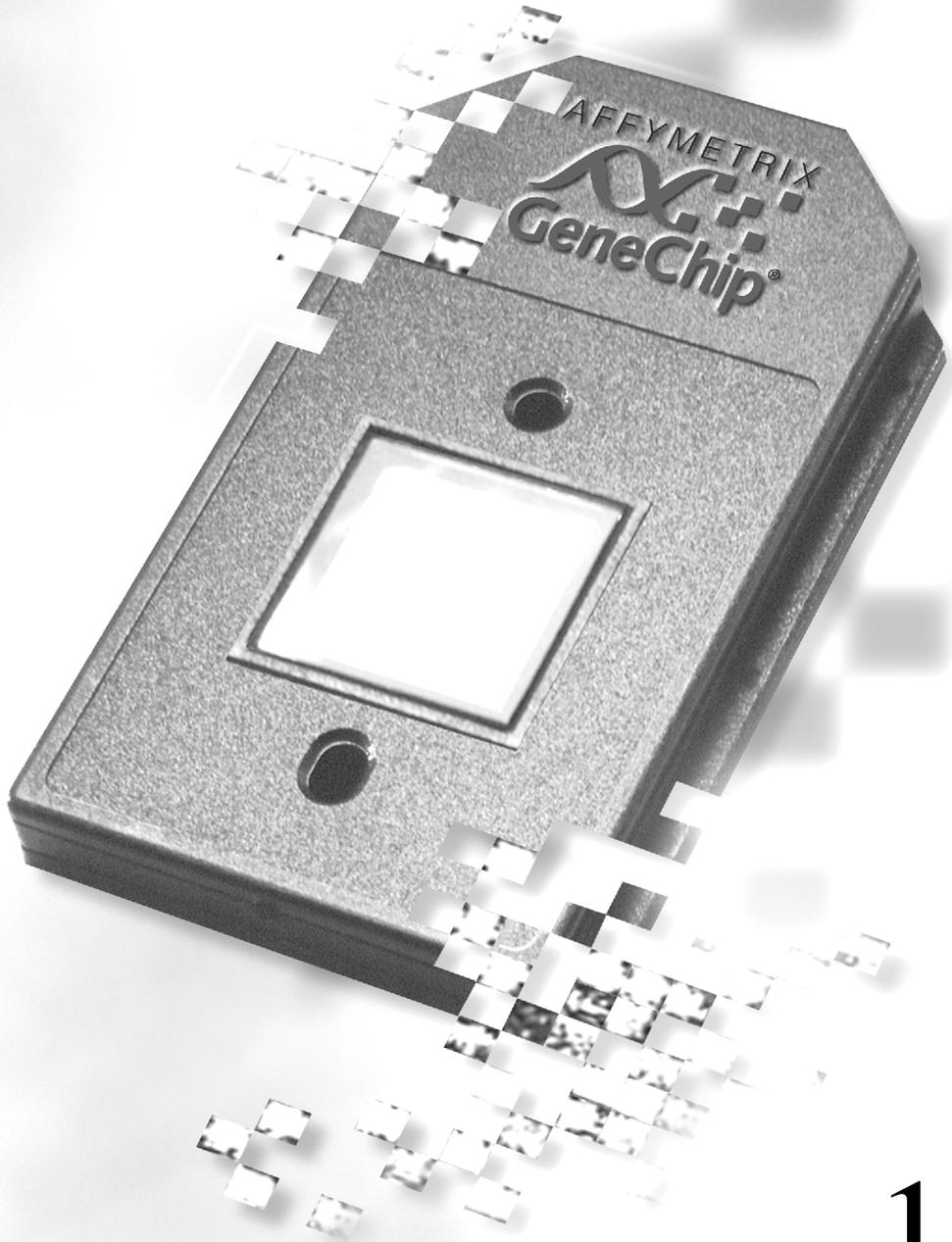


Figure 16.13
Expression Analysis Settings dialog box



Chapter 17

This chapter describes how to print:

- image data files (*.dat)
- cell intensity files (*.cel)
- analysis output files (*.chp)
- reports (*.rpt)

Printing an Image Data or Cell Intensity File

1. When the Image window is open, click the **Print** toolbar button  or select **File** → **Print** from the menu bar.
⇒ The Print dialog box appears ([Figure 17.1](#)).
No print range options are available.

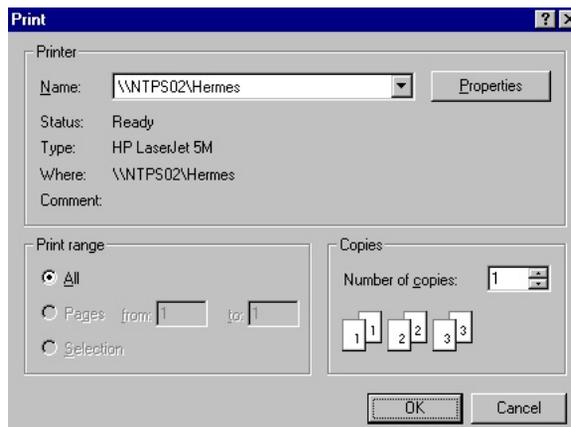


Figure 17.1
Print dialog box in the Image window

2. Click **OK** to print the image (*.dat or *.cel).

Printing an Analysis Output File (*.chp)

Expression Analysis

You may print tabular data and graphs from the Expression Analysis window (EAW).

1. When the EAW is open, click the **Print** toolbar button  or select **File** → **Print** from the menu bar.
⇒ The Print dialog box appears (Figure 17.2).

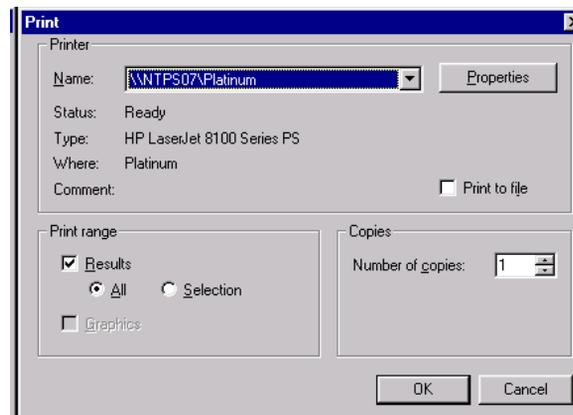


Figure 17.2
Print dialog box, EAW

The **All** and **Selection** options apply only to the Experiment Information, Metrics, and Pivot tables. If the **Graphics** option is chosen, the displayed graphs are automatically printed.

2. Choose **Results** and:
 - **All** to print the entire table in a tab (Experiment Information, Metrics, or Pivot).
 - **Selection** to print only highlighted table rows.

Mutation and Polymorphism Analysis

You may print the summary data in the left pane of the SAW as well as the probe array data (the called and reference bases). Displayed Intensity graphs are automatically printed.

1. When the Sequence Analysis window (SAW) is open, click the **Print** toolbar button  or select **File** → **Print** from the menu bar.
⇒ The Print dialog box appears ([Figure 17.3](#)).

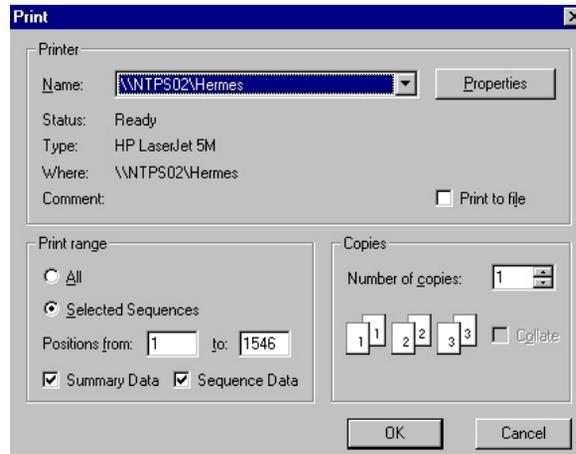


Figure 17.3
Print dialog box, SAW

2. Choose:
 - **All** to print all of the called and reference bases.
 - **Selected Sequences** and specify a nucleotide position range to print only selected bases.

Genotype Analysis

You may print tabular data and graphs from the Nucleotide Analysis window (NAW).

1. When the NAW is open, click the **Print** toolbar button  or select **File** → **Print** from the menu bar.
⇒ The Print dialog box appears ([Figure 17.4](#)).

The **All** and **Selection** options apply only to the data table.

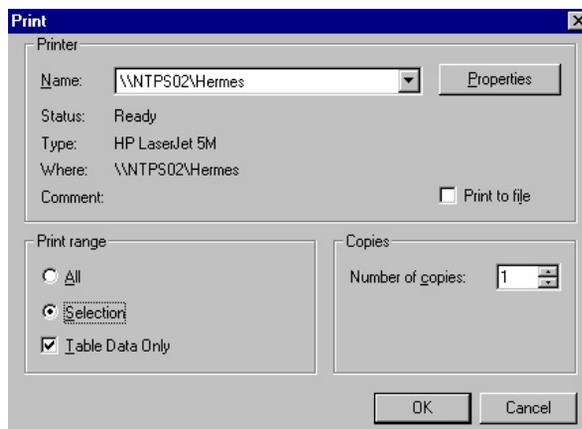


Figure 17.4
Print dialog box, NAW

2. Choose:
 - **All** to print the entire table
 - **Selection** to print only highlighted table rows
3. Choose the **Table Data Only** option to print tabular data only. Remove the check mark to print graphs as well as tabular data.

Hybridization Analysis

You may print all or selected entries in the hybridization analysis data table.

1. When the HAW is open, click the **Print** toolbar button  or select **File** → **Print** from the menu bar.
⇒ The Print dialog box appears (**Figure 175**).

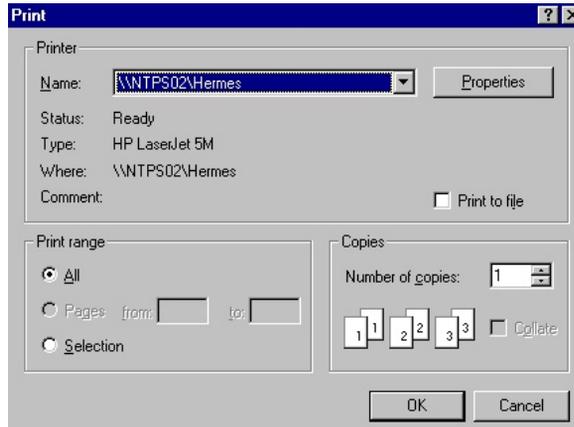


Figure 175
Print dialog box, HAW

2. Choose:
 - **All** to print the entire analysis output.
 - **Selection** to print highlighted cells in the data table.

Printing a Report (*.rpt)

You may print all or selected pages from a report.

1. When a report is open in the main display area, click the **Print** toolbar button  or select **File** → **Print** from the menu bar.
⇒ The Print dialog box appears (**Figure 17.6**).

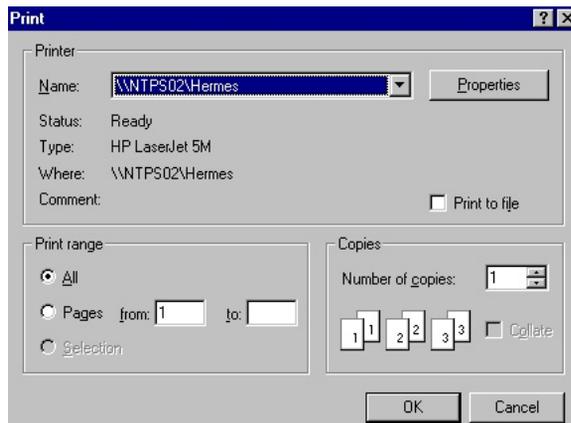
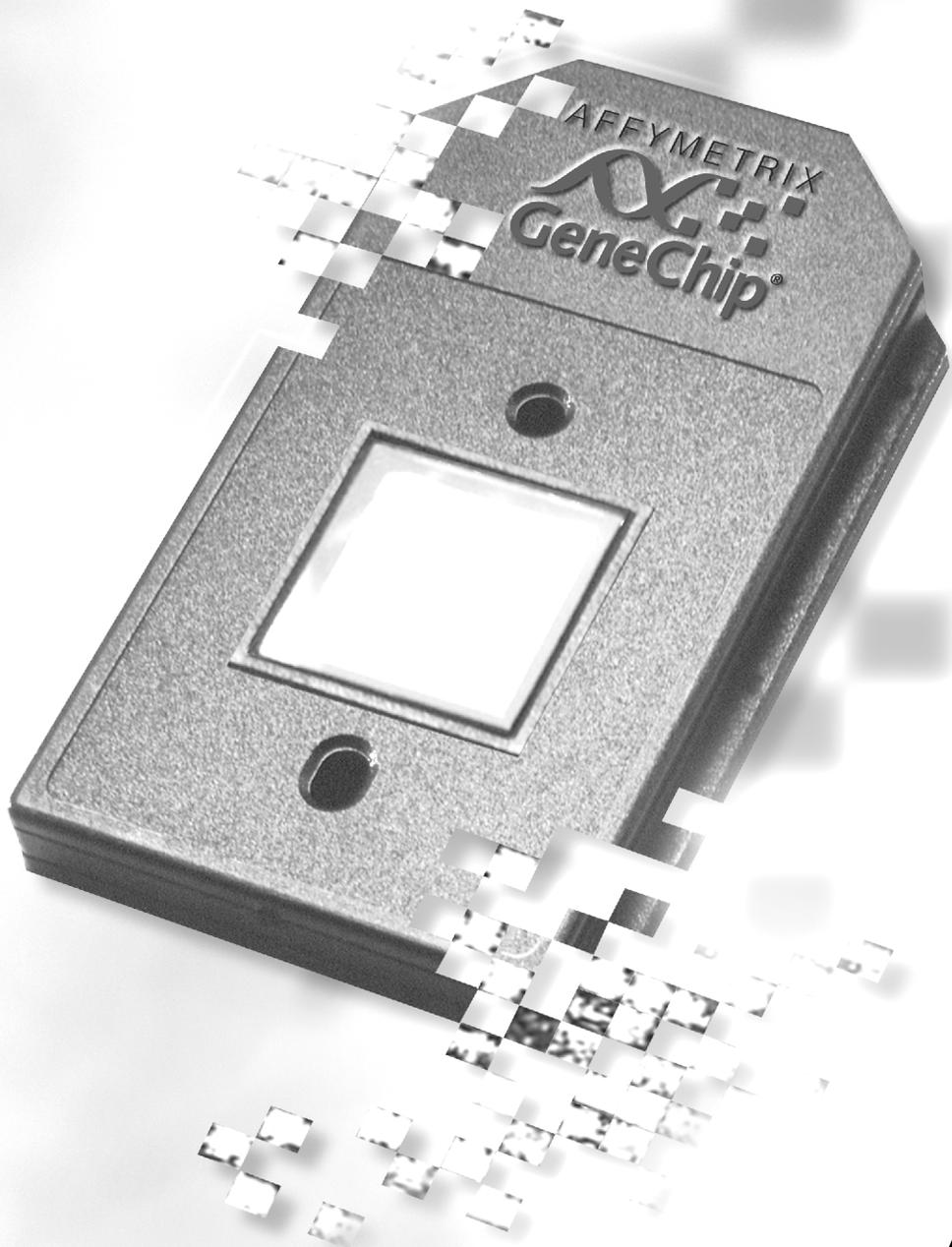


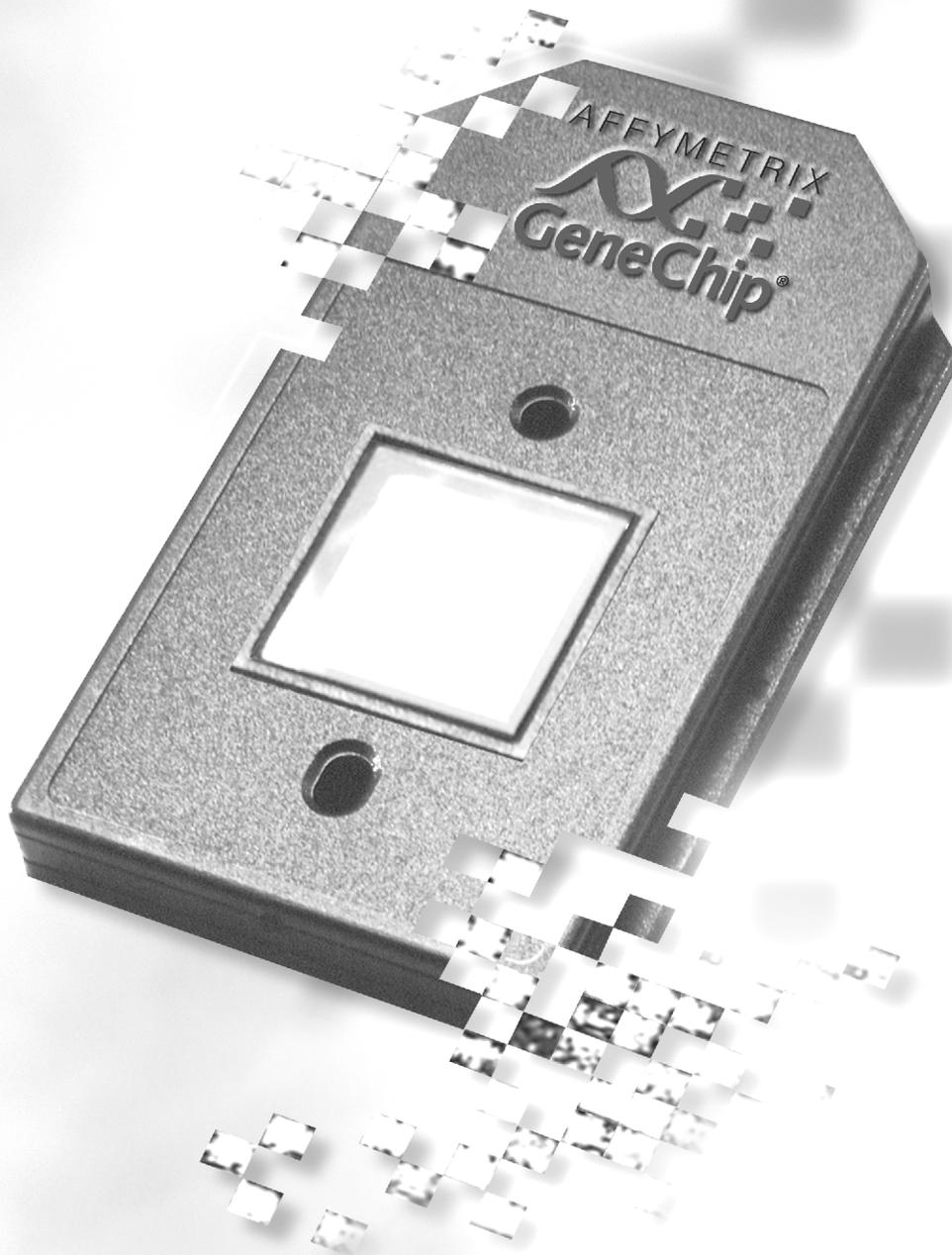
Figure 17.6
Print dialog box, report (*.rpt)

2. Choose:
 - **All** to print all pages of a report.
 - **Pages** and enter a page range to print selected pages of a report.



A

Appendix A



Microarray Suite Instrument Installation

The instrument installation only needs to be installed on an instrument system. Appendix A sections are:

- Fluidics Station and Agilent Scanner Installation
- Fluidics Station Installation (only)
- Agilent Scanner Installation (only)

The instrument installation needs to be run if there has been a change to any of the drivers, otherwise upgrading the software to Microarray Suite 5.0 is the only installation upgrade required.

✓ NOTE

The instrument installation is for Microsoft® Windows NT® 4.0 Workstation only.

Fluidics Station and Agilent Scanner Installation

✓ NOTE

The instrument installation is for Microsoft® Windows NT® 4.0 Workstation only.

If the fluidics station software needs to be installed, the correct driver must also be installed. The two possible types of cards are: Cimetrics and Sealevel.

1. On an instrument workstation, the Fluidics Station driver needs to be installed if a **Sealevel card** is configured in the workstation. The workstation may come configured with a Sealevel card or a Cimetrics card (on older workstations).

Review [Figure A.1](#) to determine if a Sealevel card is installed in your system.

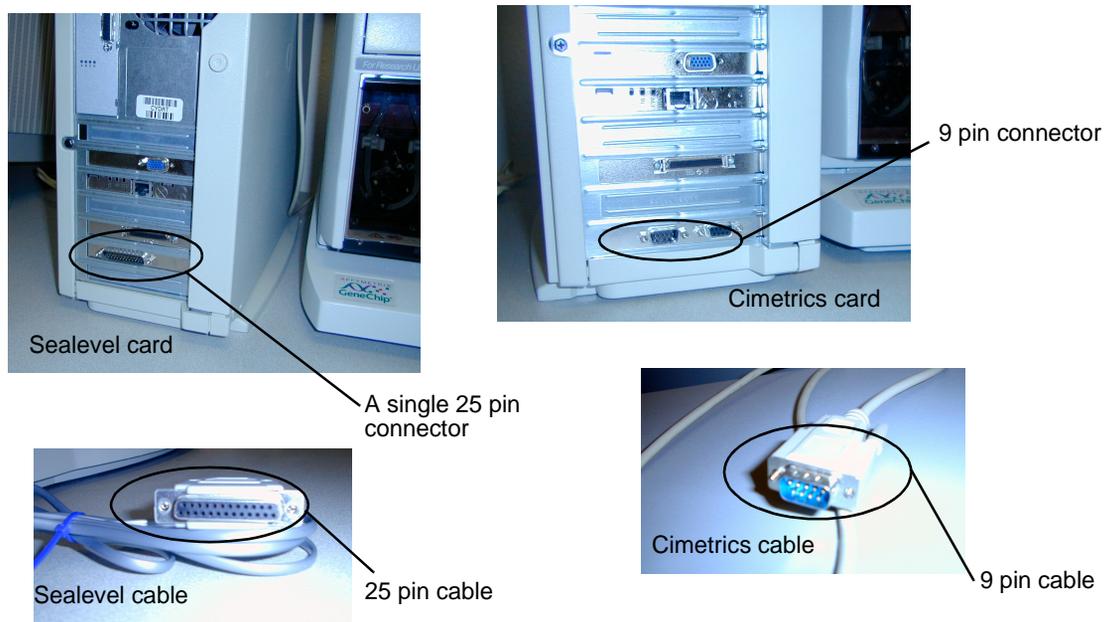


Figure A.1
Determining Sealevel card vs. Cimetrics card

2. Launch **Windows NT[®] Explorer**.
3. Browse to the Microarray Suite 5.0 CD.
4. Double-click the **Instrument** Folder.
5. Double-click **setup.exe** within the Instrument folder.
⇒ The **Welcome** window appears.
6. Click **Next**.
7. Several consecutive **Software License Agreement** windows appear. Click **Yes** in each window to accept the terms of the agreement.
⇒ The Choose Destination Location window appears (**Figure A.2**).

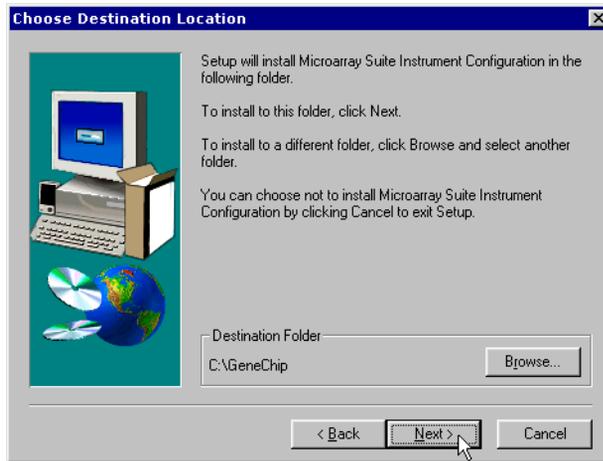
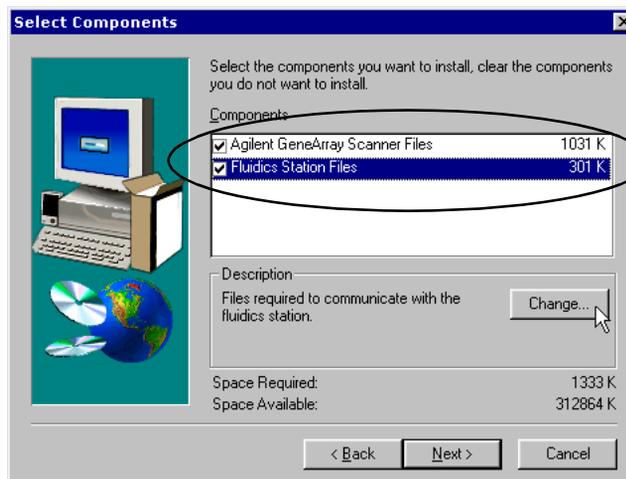


Figure A.2
Choose Destination Location window

8. Click **Browse** and select the **Destination** to install the instrument driver (select the **same location** where you installed Microarray Suite).
9. Click **Next**.
⇒ The Select Components window appears (**Figure A.3**).



Select both the
Scanner Files and
Fluidics Station Files

Figure A.3
Select Components for Instruments

10. Select both the **GeneArray Scanner Files** option and the **Fluidics Station Files** option (Figure A.3).
11. Highlight the **Fluidics Station Files** option and click **Change...** to verify that the correct driver is selected (Cimetrics or Sealevel).
12. If using Sealevel, enter the Port # for the COM serial port for the Sealevel serial card. Enter **2** (Figure A.4).

✓ NOTE

If the workstation is the **Dell GX110**, select **COM Port 3**.

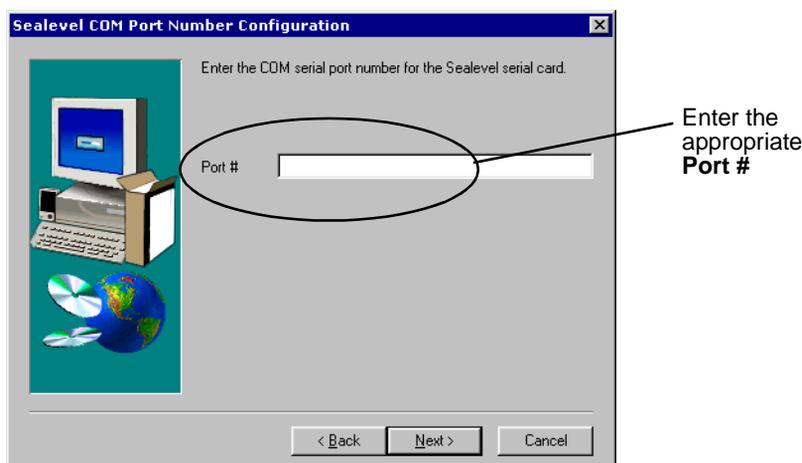


Figure A.4
Select the Port Number (Sealevel card only)

13. Click **Next**.
⇒ The **Start Copying Files** window appears (Figure A.5).
This is a summary of the information selected.

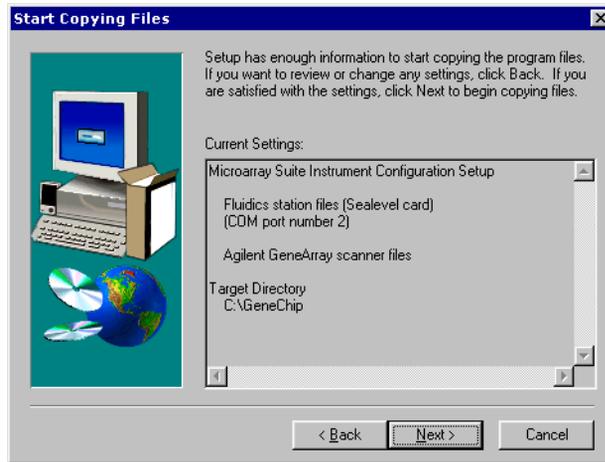


Figure A.5
Start Copying Files Window

14. Review the information and click **Next** to continue.

⇒ Program files are copied to your system. When complete an Adaptec ASPI installation window appears (**Figure A.6**).

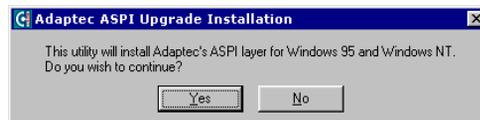


Figure A.6
Adaptec ASPI Upgrade installation window

15. Click **Yes** to install.

⇒ A second Adaptec ASPI Upgrade window appears (**Figure A.7**).

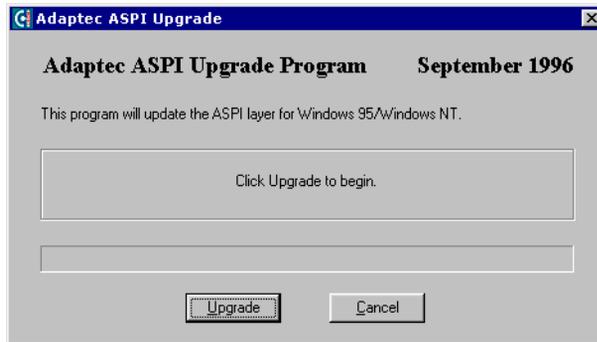


Figure A.7
Adaptec ASPI Upgrade installation window

16. Click **Upgrade** to continue.

⇒ Program files are installed/upgraded. When complete, an information window appears ([Figure A.8](#)).

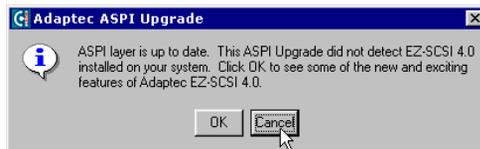


Figure A.8
Adaptec ASPI Upgrade installation window

17. Click **Cancel**.

⇒ The Setup Complete window appears ([Figure A.9](#)).



Figure A.9
Restart Computer Window

- 18.** Select, **Yes, I want to restart my computer now.**
- 19.** Click **Finish.**
 - ⇒ The system reboots.

Fluidics Station Installation (only)

✓ NOTE

The instrument installation is for Microsoft® Windows NT® 4.0 Workstation only.

If the fluidics station software needs to be installed, the correct driver must also be installed. The two possible types of cards are: Cimetrics and Sealevel.

- 1.** On an instrument workstation, the Fluidics Station driver needs to be installed if a **Sealevel card** is configured in the workstation. The workstation may come configured with a Sealevel card or a Cimetrics card (on older workstations).

Review [Figure A.10](#) to determine if a Sealevel card is installed in your system.

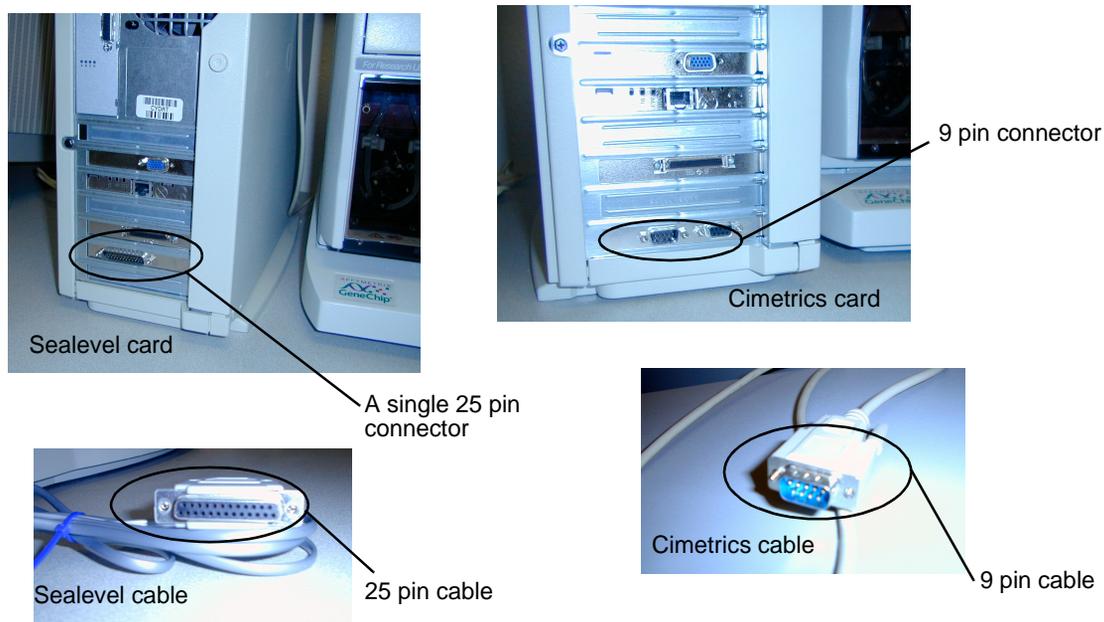


Figure A.10
Determining Sealevel card vs. Cimetrics card

- 2.** Launch **Windows NT[®] Explorer**.
- 3.** Browse to the Microarray Suite 5.0 CD.
- 4.** Double-click the **Instrument** Folder.
- 5.** Double-click **setup.exe** within the Instrument folder.
⇒ The **Welcome** window appears.
- 6.** Click **Next**.
- 7.** Several consecutive **Software License Agreement** windows appear. Click **Yes** in each window to accept the terms of the agreement.
⇒ The Choose Destination Location window appears (**Figure A.11**).

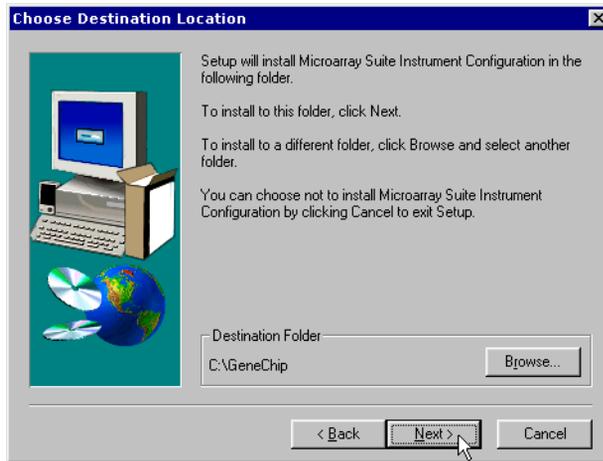


Figure A.11
Choose Destination Location window

8. Click **Browse** and select the **Destination** to install the instrument driver (select the **same location** where you installed Microarray Suite).
9. Click **Next**.
⇒ The Select Components window appears (**Figure A.12**).

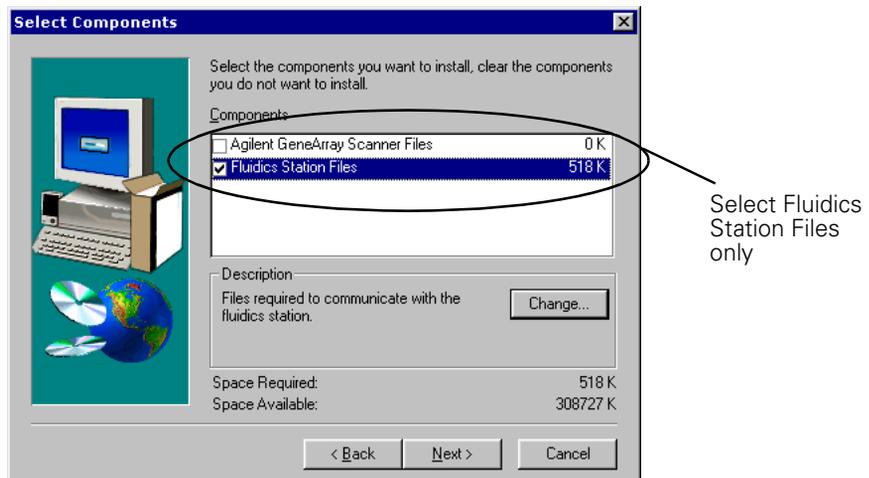


Figure A.12
Select Components for Instruments

10. Select the **Fluidics Station Files**, (Figure A.12).
This should be the only option selected.
11. Highlight the **Fluidics Station Files** option and click **Change...** to verify that the correct driver is selected (Cimetrics or Sealevel).
12. If using Sealevel, enter the Port # for the COM serial port for the Sealevel serial card. Enter 2, (Figure A.13).

✓ NOTE

*If the workstation is the **Dell GX110**, select **COM Port 3**.*

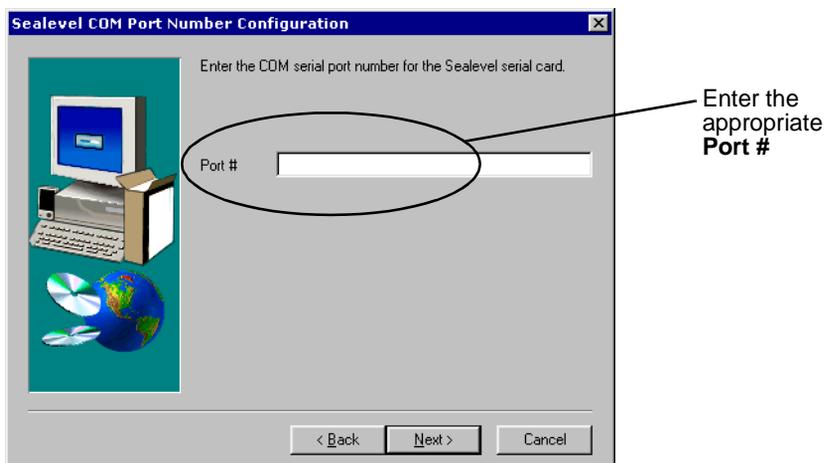


Figure A.13
Select the Port Number (Sealevel card only)

13. Click **Next**.
⇒ The **Start Copying Files** window appears (Figure A.14).
This is a summary of the information that was selected by the installer.

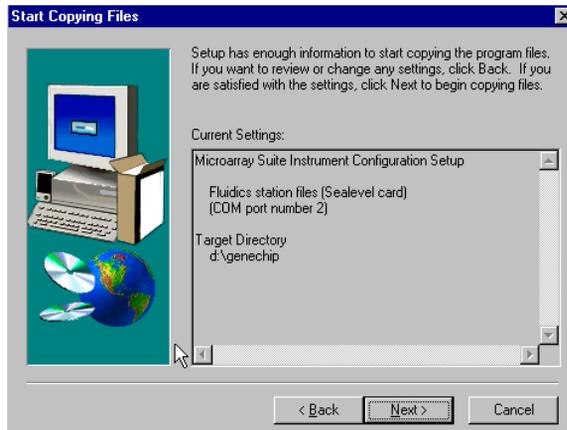


Figure A.14
Start Copying Files Window

14. Review the information and click **Next** to continue.
15. When the Setup Complete window appears (Figure A.15), select **Yes, I want to restart my computer now.**



Figure A.15
Restart Computer Window

16. Click **Finish**.
⇒ The system reboots.

Agilent Scanner Installation (only)

✓ NOTE

The instrument installation is for Microsoft® Windows NT® 4.0 Workstation only.

1. Launch **Windows NT® Explorer**.
2. Browse to the Microarray Suite 5.0 CD.
3. Double-click the **Instrument** Folder.
4. Double-click **setup.exe** within the Instrument folder.
⇒ The **Welcome** window appears.
5. Click **Next**.
6. Several consecutive **Software License Agreement** windows appear. Click **Yes** in each window to accept the terms of the agreement.
⇒ The Choose Destination Location window appears ([Figure A.16](#)).



Figure A.16
Choose Destination Location window

7. Click **Browse** and select the **Destination** to install the instrument driver (select the **same location** where you installed Microarray Suite).
8. Click **Next**.
⇒ The Select Components window appears ([Figure A.17](#)).

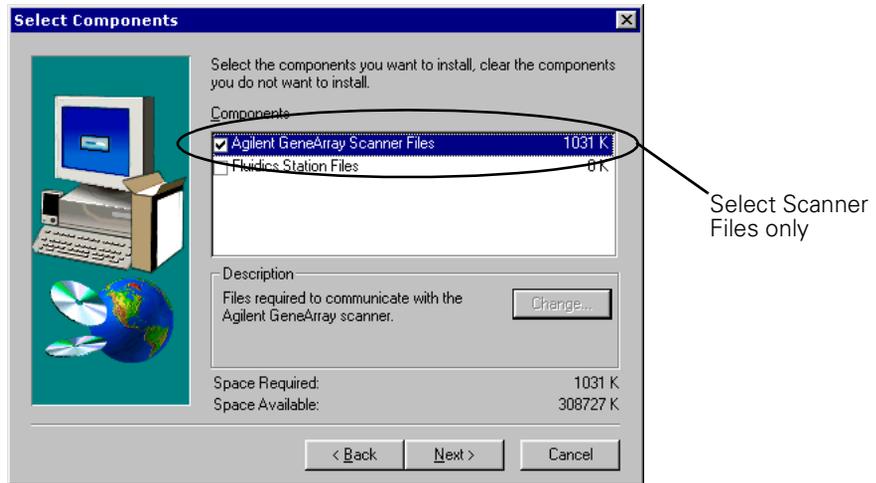


Figure A.17
Select Components for Instruments

9. Click **Next**.
⇒ The **Start Copying Files** window appears ([Figure A.18](#)).
This is a summary of the information that was selected by the installer.

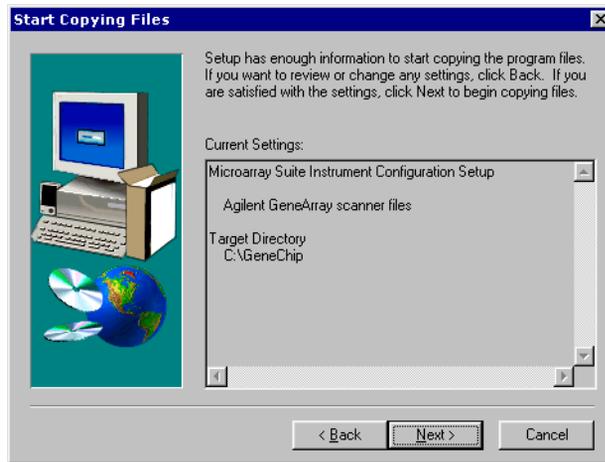


Figure A.18
Start Copying Files Window

- 10.** Review the information and click **Next** to continue.
 - ⇒ Program files are copied to your system. When complete an Adaptec ASPI installation window appears (**Figure A.19**).

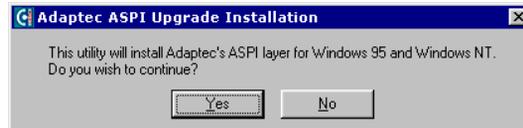


Figure A.19
Adaptec ASPI Upgrade installation window

- 11.** Click **Yes** to install.
 - ⇒ A second Adaptec ASPI Upgrade window appears (**Figure A.20**).

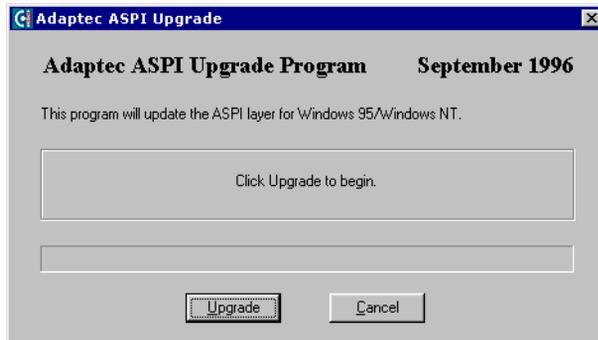


Figure A.20
Adaptec ASPI Upgrade installation window

12. Click **Upgrade** to continue.

⇒ Program files are installed/upgraded. When complete, an information window appears (Figure A.21).

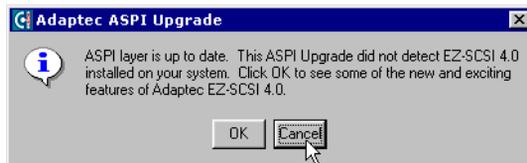


Figure A.21
Adaptec ASPI Upgrade installation window

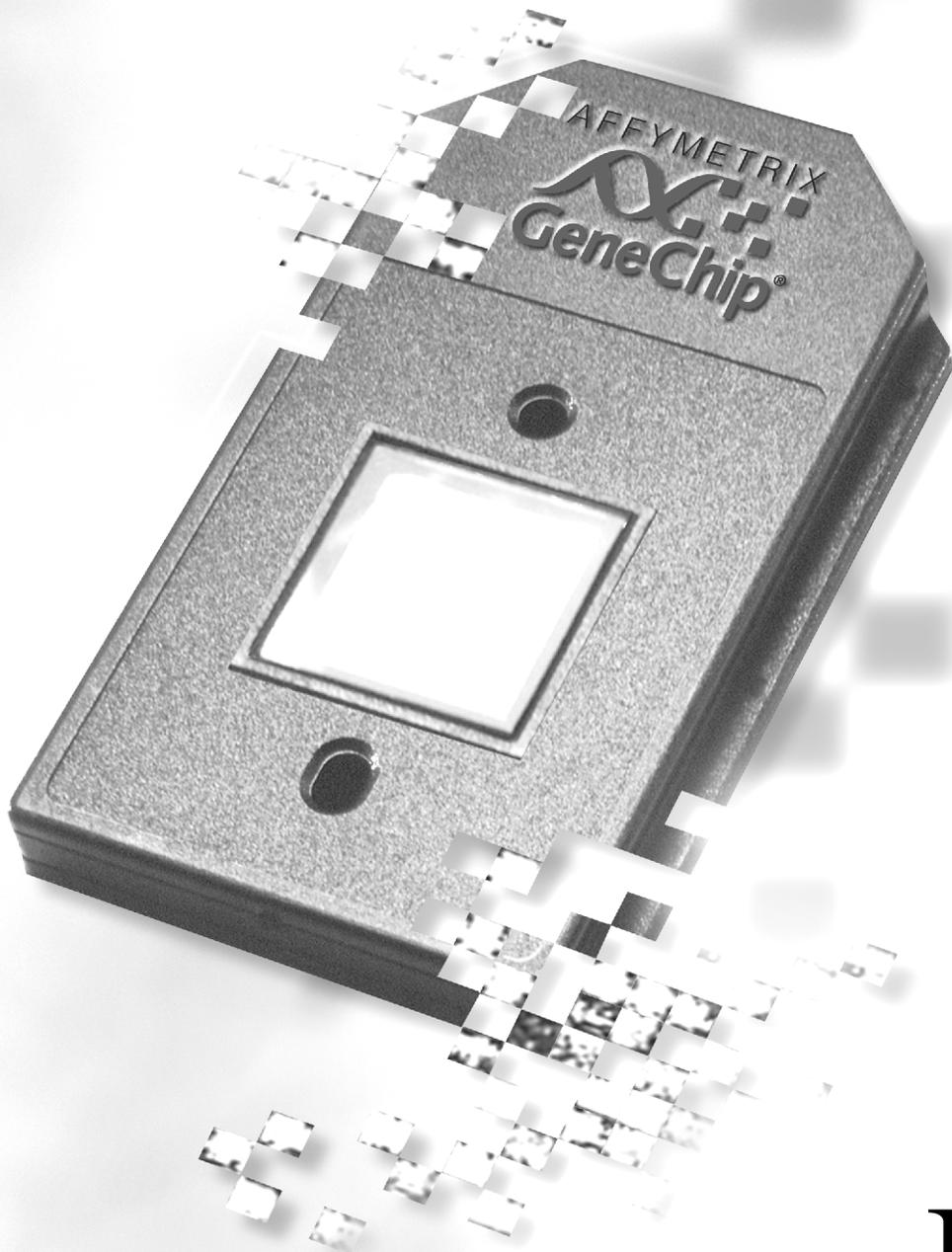
13. Click **Cancel**.

⇒ The Setup Complete window appears (Figure A.22).



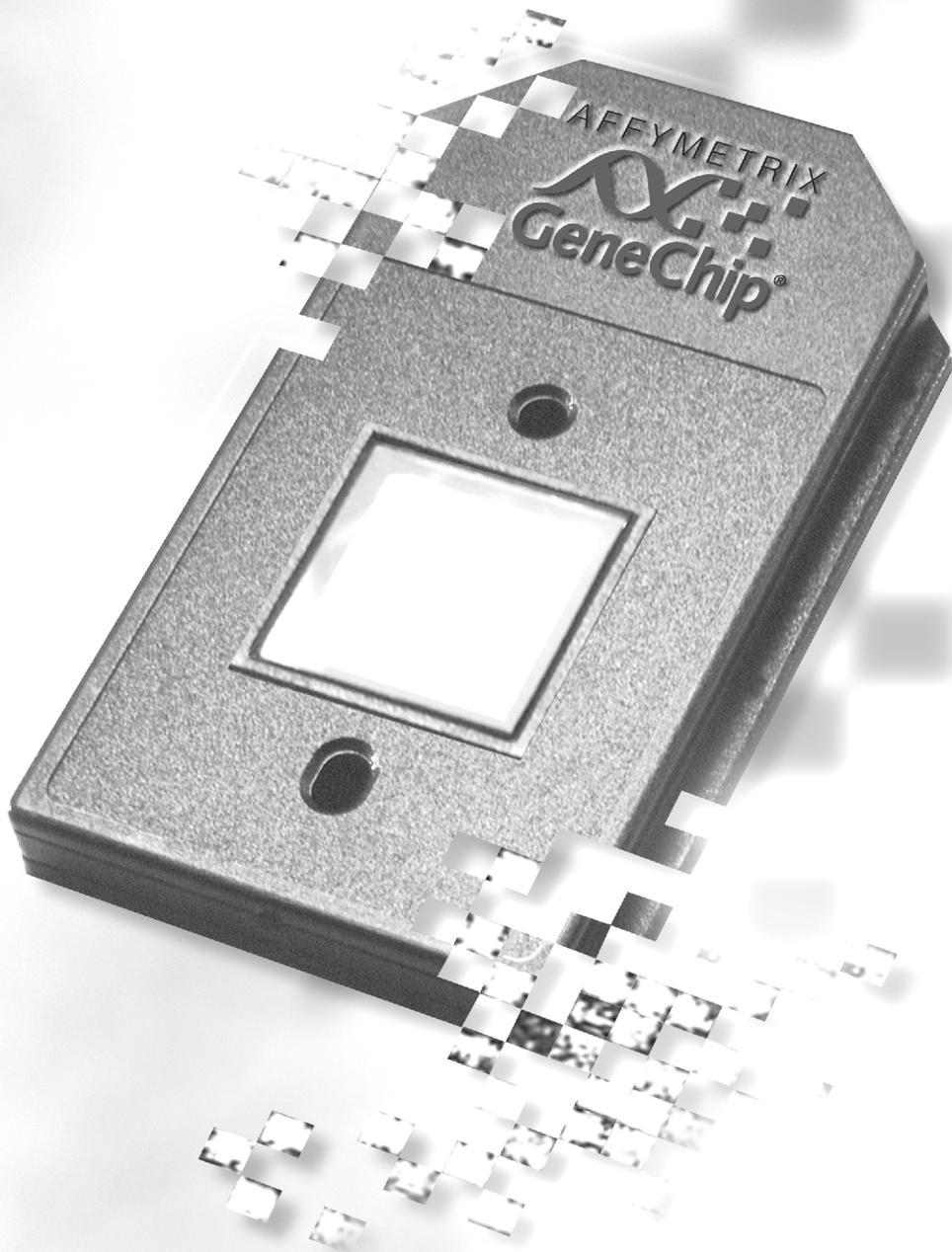
Figure A.22
Restart Computer Window

- 14.** Select, **Yes, I want to restart my computer now.**
- 15.** Click **Finish.**
 - ⇒ The system reboots.



B

Appendix B





Microarray Suite File Types

Probe Information (Library) Files

The probe information or *library* files contain information about the probe array design characteristics, probe utilization and content, and scanning and analysis parameters. These files are unique for each probe array type.

The default path for the probe information files is C:\GeneChip\Library.

■ To view the file location:

1. Select **Tools** → **Defaults** from the menu bar.
2. Click the File Locations tab in the Defaults dialog box that appears.

Library files include user-defined mask files (see [Table B.1](#)).

Table B.1
User-defined mask files

Mask File Name	File Extension	Description
Probe Mask	*.msk	A user-specified list of probe pairs that are excluded from an analysis.
Cross Hybridization Probe Mask	*.msk	A type of probe mask that specifies probe pairs that include a PM or MM probe cell whose intensity exceeds a user-specified limit. These probe pairs will be excluded from the analysis.
Hybridization Probe Mask	*.msk	A type of probe mask that specifies probe pairs where $PM - MM < \text{Difference Threshold}$ or $PM/MM < \text{Ratio Threshold}$. These probe pairs will be excluded from the analysis.
Spike Probe Mask	*.msk	A type of probe mask that specifies probe pairs where: $(PM - MM)_{\text{spike}} - (PM - MM)_{\text{unspike}} < \text{Difference Threshold}$ or $(PM - MM)_{\text{spike}} / (PM - MM)_{\text{unspike}} < 1 + \text{Ratio Threshold}$. These probe pairs will be excluded from the analysis.

Fluidics Protocol Files

The fluidics protocol files contain the instrument control instructions used by the GeneChip® Fluidics Station 400. The default path for these files is C:\GeneChip\Protocol.

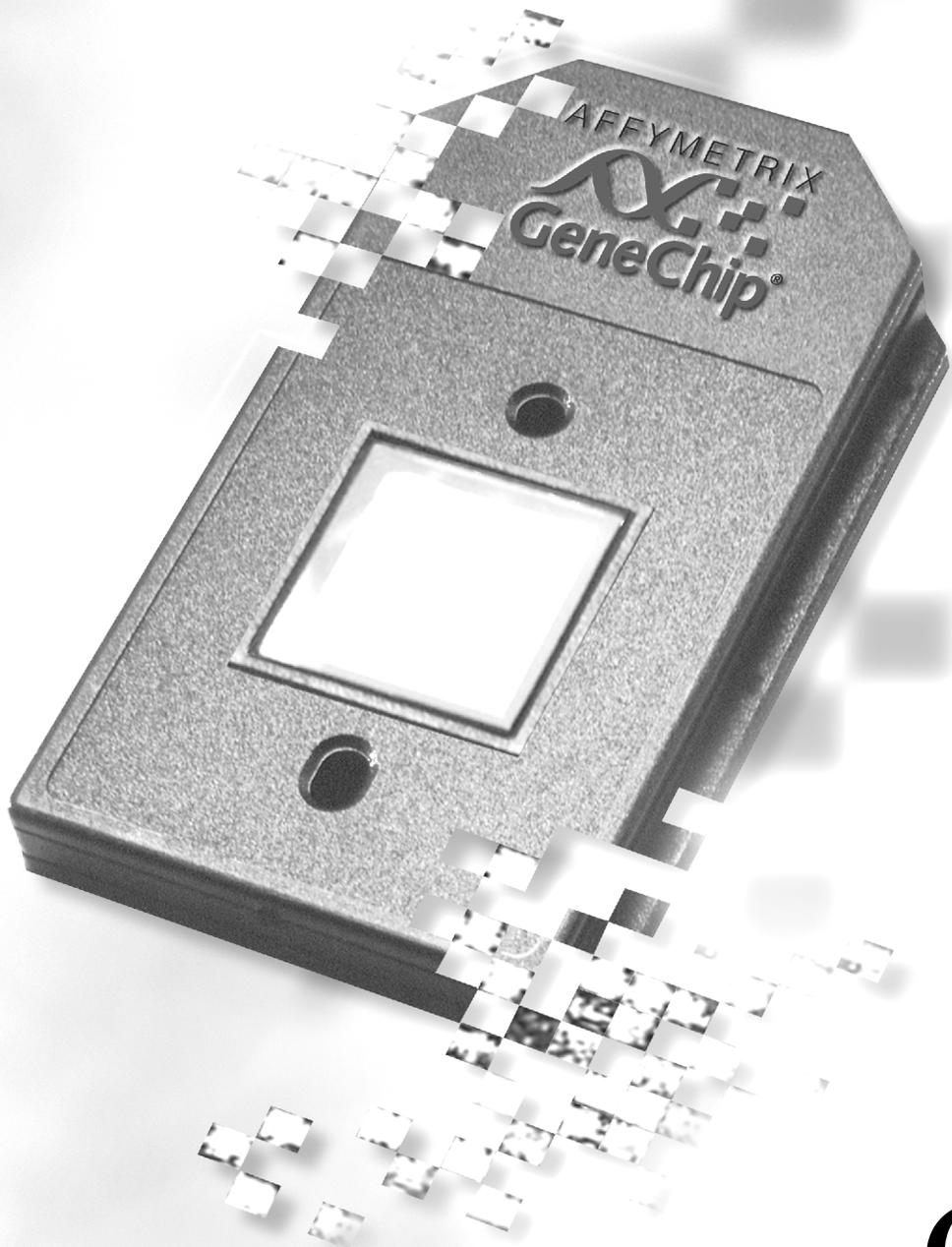
Experiment Data Files

You create the experiment information file (*.exp) during experiment set up. The software generates the other experiment data files types as data analysis proceeds (see [Table B.2](#)).

The default path for *.exp files is C:\GeneChip\TestData. If desired, the files may be written to another directory. The destination path may be set in the Files Locations tab of the Defaults dialog box. (Select **Tools** → **Defaults** from the menu bar and click the Files Locations tab.)

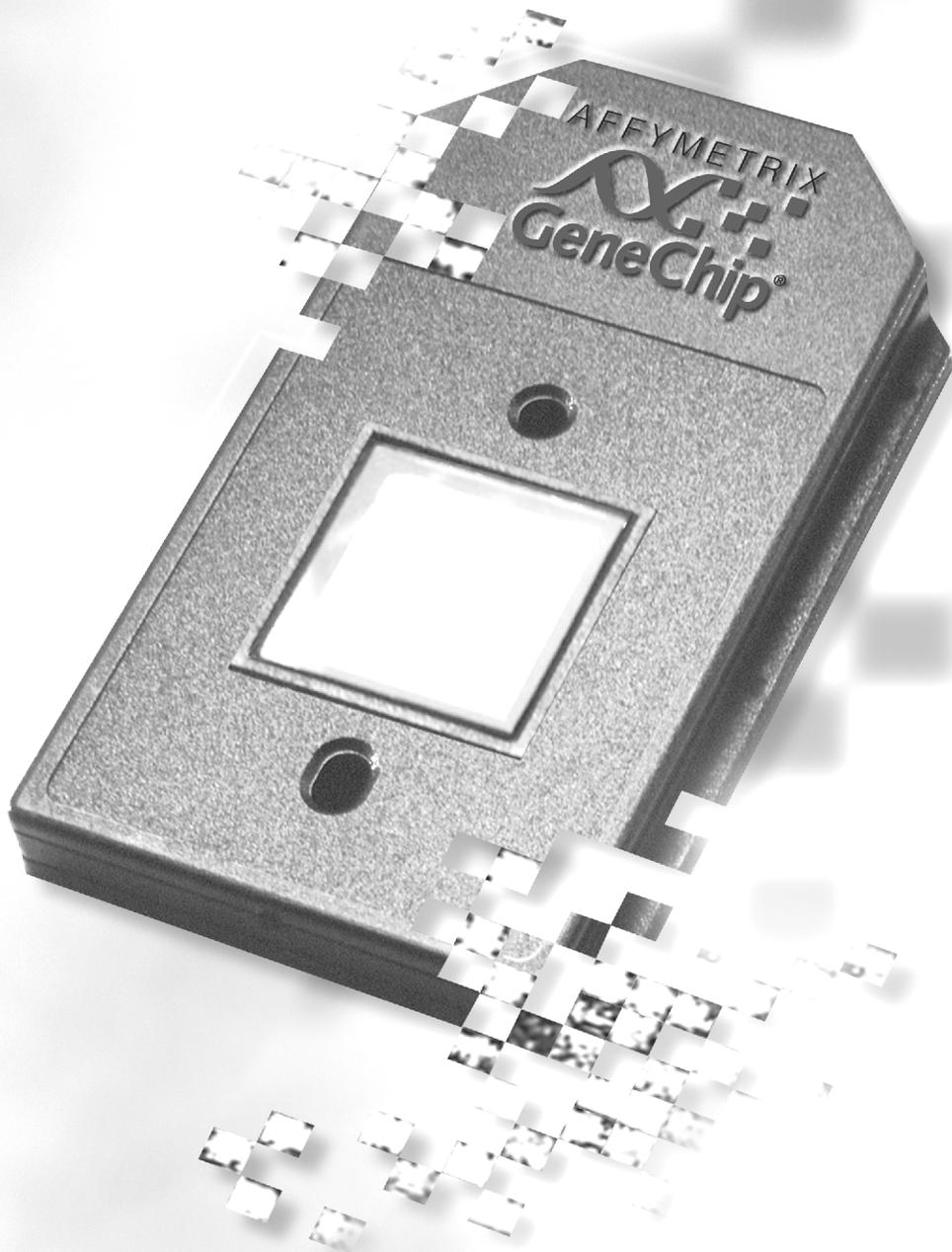
Table B.2
Experiment data files

Experiment Data File Name	File Extension	Description
Experiment Information File	*.exp	Contains information about the experiment name, sample, and probe array type. The experiment name also provides the name for subsequent test data files generated during the analysis of the experiment.
Data File	*.dat	The image of the scanned probe array.
Cell Intensity File	*.cel	The software derives the *.cel file from a *.dat file and automatically creates it upon opening a *.dat file. It contains a single intensity value for each probe cell delineated by the grid (calculated by the Cell Analysis algorithm).
Chip File	*.chp	The output file generated from the analysis of a probe array.
Report File	*.rpt	The report generated from the analysis output file (*.chp).
Experiment Information File	*.tif	A standard file format for graphic images. The Microarray Suite software exports graphic images in this file format.
Data File	*.txt	A standard format for text files. The Microarray Suite software exports text in this file format.
	*.xls	A standard format for Excel files. The Microarray Suite software export text in this file format.



C

Appendix C





Statistical Expression Algorithm

Microarray Suite runs the Statistical expression algorithm. Previous versions of Microarray Suite (lower than 5.0) run the Empirical expression algorithm.

Notation

A GeneChip[®] probe array consists of a number of *cells* (square-shaped areas on the array) and each contains many copies of a unique probe. Probes are tiled in probe pairs consisting of a perfect match (PM) and a mismatch (MM).

The sequence of the PM and MM are the same, except for a base substitution in the middle of the MM probe sequence. A probe set includes a series of probe pairs and represents an expressed transcript ([Figure C.1](#)).

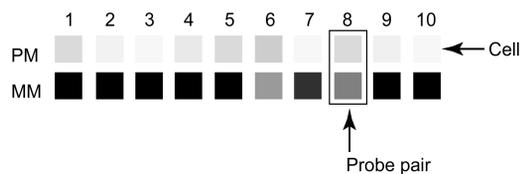


Figure C.1
Probe set that includes 10 probe pairs

Algorithm Output Definitions

Signal	A measure of the abundance of a transcript.
Stat Pairs	The number of probe pairs for a particular probe set on the array.
Stat Pairs Used	$= \text{Pairs} - \text{Masked probe pairs} - \text{Saturated MM probe pairs}$ This is the number of pairs used by the Statistical Expression algorithm to make the detection call in an absolute analysis.
Detection	The call in an absolute analysis that indicates if the transcript was present (P), absent (A), marginal (M), or no call (NC).
Detection p-value	p-value that indicates the significance level of the detection call.
Stat Common Pairs	The intersection of the probe pairs from the baseline and experiment that are used by the Statistical Expression algorithm to make the change call in a comparison analysis.
Change	The call that indicates the change in transcript level between a baseline and an experiment array.
Change p-value	p-value that indicates the significance level of the change call.
Signal log ratio	The change in expression level for a transcript between a baseline and an experiment array. This change is expressed as the \log_2 ratio.
Signal log ratio low	The lower limit of the log ratio within a 95% confidence range.
Signal log ratio high	The upper limit of the signal log ratio within a 95% confidence range.

Background Subtraction

The first step in the analysis is to correct for background across the entire array. The calculated background establishes an intensity *floor* that is subtracted from all intensity values.

Raw cell intensities → Background-adjusted intensities

The algorithm:

- divides the array into equally spaced zones (Figure C.2)
- assigns an average background to the center of each zone
- computes the distance from each cell to the center of every zone
- computes a weighting factor (the reciprocal of the sum of a constant and the square of the distance from the cell to the zone center)
- computes the background of each cell by applying the weighting factor to the zone average (the average background assigned to the center of each zone)

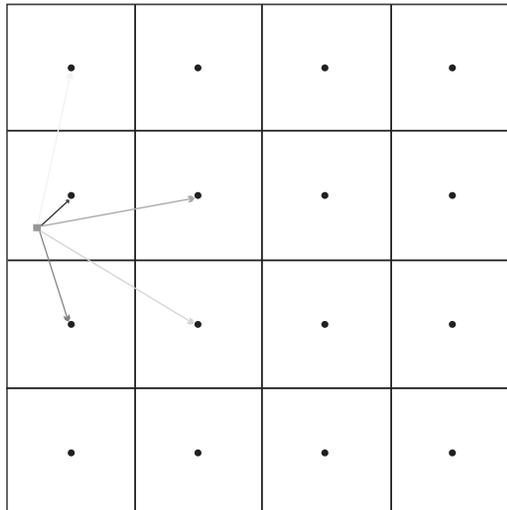


Figure C.2

Array zones for computing background; the arrow color indicates the relative weights

Absolute Expression Analysis

Detection

The detection call answers the question: Is the transcript of a particular probe set reliably detected by the probe array? We want an answer of *absent* or *present*. In this context, absent means the expression level is below the threshold of detection. In the case of uncertainty, we can get a *marginal* call.

The algorithm provides absolute information without reference to numerical values. This makes it easy to filter and interpret results. For example, we may only want to look at genes whose transcripts are present in a particular experiment. An additional advantage is that the detection and the signal values are calculated using independent algorithms that add additional information to the results.

Raw cell intensities → Absent, present, or Marginal detection call plus p-values

✓ NOTE

A No Call detection result occurs if all of the probe pairs of a probe set are excluded from the analysis. A probe pair is excluded if the PM or MM is masked or if the MM is saturated.

Discrimination Value

The algorithm computes a discrimination value that is used as a filter to remove from further consideration all probe sets with insignificant differences between PM and MM.

Discrimination value = $(PM - MM)/(PM + MM)$

The median of the discrimination ratios of all the probe pairs of a probe set is compared to a user-modifiable parameter τ (default = 0.015), and produces an intermediate call.

✓ NOTE

Increasing τ can reduce the number of false present calls, but may also reduce the number of true present calls.

Making the Call

A one-sided Wilcoxon's signed rank test is used to calculate a p-value that reflects the significance of the differences between PM and MM. The p-value or statistical significance of a result is the probability that the observed change in a sample occurred by pure chance. For example, a p-value of 0.05 means there are five chances in 100 that the results are not significant. The lower the p-value, the greater the probability that the results are significant.

To make a call, the p-value for a probe set is examined on an axis with two user-definable thresholds, α_1 and α_2 (defaults $\alpha_1 = 0.04$ and $\alpha_2 = 0.06$). For p-values between zero and 0.5, α_1 and α_2 define the thresholds for the calls (Figure C.3) (see Table C.1).

The result is reported as the *detection call* that is associated with the calculated p-value.

✓ NOTE

Decreasing the significance level α_1 can reduce the number of false detected calls and reduce the number of true detected calls.

Increasing the significance level α_2 can reduce the number of false undetected calls and reduce the number of true undetected calls.

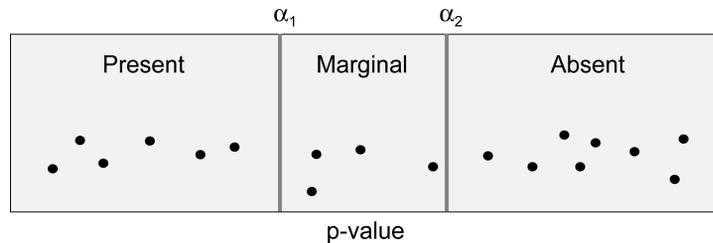


Figure C.3
User-modifiable thresholds, α_1 and α_2

Table C.1
Statistical algorithm detection call rules

Computed Detection p-values	Detection Call
$p < \alpha_1$	Present
$\alpha_1 \leq p < \alpha_2$	Marginal (at the limit of detection)
$p \geq \alpha_2$	Absent

Calculating the Signal

The signal represents the amount of transcript in solution.

Background-adjusted cell intensities → Probe set signal

For each PM intensity, a matching MM probe provides a reference background hybridization intensity. If the MM value is less than the PM value, the algorithm uses the MM value directly. However, if the MM value is larger than the PM value, the algorithm creates an adjusted MM value based on the average difference intensity between \log_2 PM and \log_2 MM, or if that measure is too small, some fraction of PM.

The adjusted MM values are used to calculate the \log_2 PM - log adjusted MM for each probe pair. The signal for a probe set is calculated as the one-step biweight estimate of the combined differences of all the probe pairs in the probe set.

Comparison Expression Analysis

A comparison analysis compares the expression levels of the transcripts on one array to those on another. By directly comparing matching cells on two arrays, any inherent differences in the hybridization efficiency of those cells will cancel out. As a result, this is an accurate and sensitive method of determining changes in expression levels.

Differences between PM and MM and Differences between PM and Background (Experiment and Baseline) → Change call and p-value

The algorithm computes a primary normalization factor and two additional normalization factors that straddle the primary normalization factor. The spread between the normalization factors is determined by the perturbation parameter (d).

To determine the p-values, a signed rank analysis is carried out on the PM and MM differences for each probe pair in a probe set from the two arrays in the comparison. The resulting p-values are used to make the change calls.

To make a call, the p-value for a probe set is examined on an axis with four thresholds (Figure C.4). For p-values between zero and 1.0, γ_1 and γ_2 define the thresholds for the calls.

The result is reported as the *change call* that is associated with the calculated p-value (see Table C.2). The output p-value is the *critical* p-value. The critical p-value is the p-value that determines the change call.

The critical p-value, p , is defined by the following:

- $p = \max(p_1, p_2, p_3)$, if $p_1 < 0.5$, $p_2 < 0.5$ and $p_3 < 0.5$
- $p = \min(p_1, p_2, p_3)$, if $p_1 > 0.5$, $p_2 > 0.5$ and $p_3 > 0.5$
- $p = 0.5$, otherwise

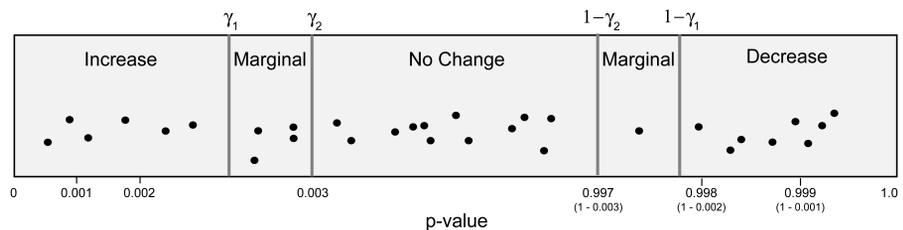


Figure C.4
User-modifiable parameters, γ_1 and γ_2

Table C.2
Statistical algorithm change call rules

Computed Change P-values	Change Call
$p_0 < \gamma_1$ $p_1 < \gamma_1$ $p_2 < \gamma_1$	Increasing
$p_0 < \gamma_2$ $p_1 < \gamma_2$ $p_2 < \gamma_2$	Marginally increasing
Else	No change
$p_0 > 1-\gamma_2$ $p_1 > 1-\gamma_2$ $p_2 > 1-\gamma_2$	Marginally decreasing
$p_0 > 1-\gamma_1$ $p_1 > 1-\gamma_1$ $p_2 > 1-\gamma_1$	Decreasing

You may adjust γ_1 and γ_2 by specifying different values for: γ_{1L} , γ_{1H} , γ_{2L} , γ_{2H} . (See [Appendix D](#) Expression Analysis Settings.)

Decreasing γ_1 can reduce the number of false increase and decrease calls, but can also reduce the number of true increase and decrease calls. Increasing γ_2 can reduce the number of false no-change detected calls, but can also reduce the number of true no-change-detected calls. Increasing the perturbation parameter (d) can increase the number of true no-change-detected calls, but can also increase the number of false no-change-detected calls.

✓ NOTE

A No Call occurs if all of the probe pairs of a probe set are excluded from the analysis. A probe pair is excluded if the PM or MM in the experiment or baseline is masked or saturated.

Signal Log Ratio

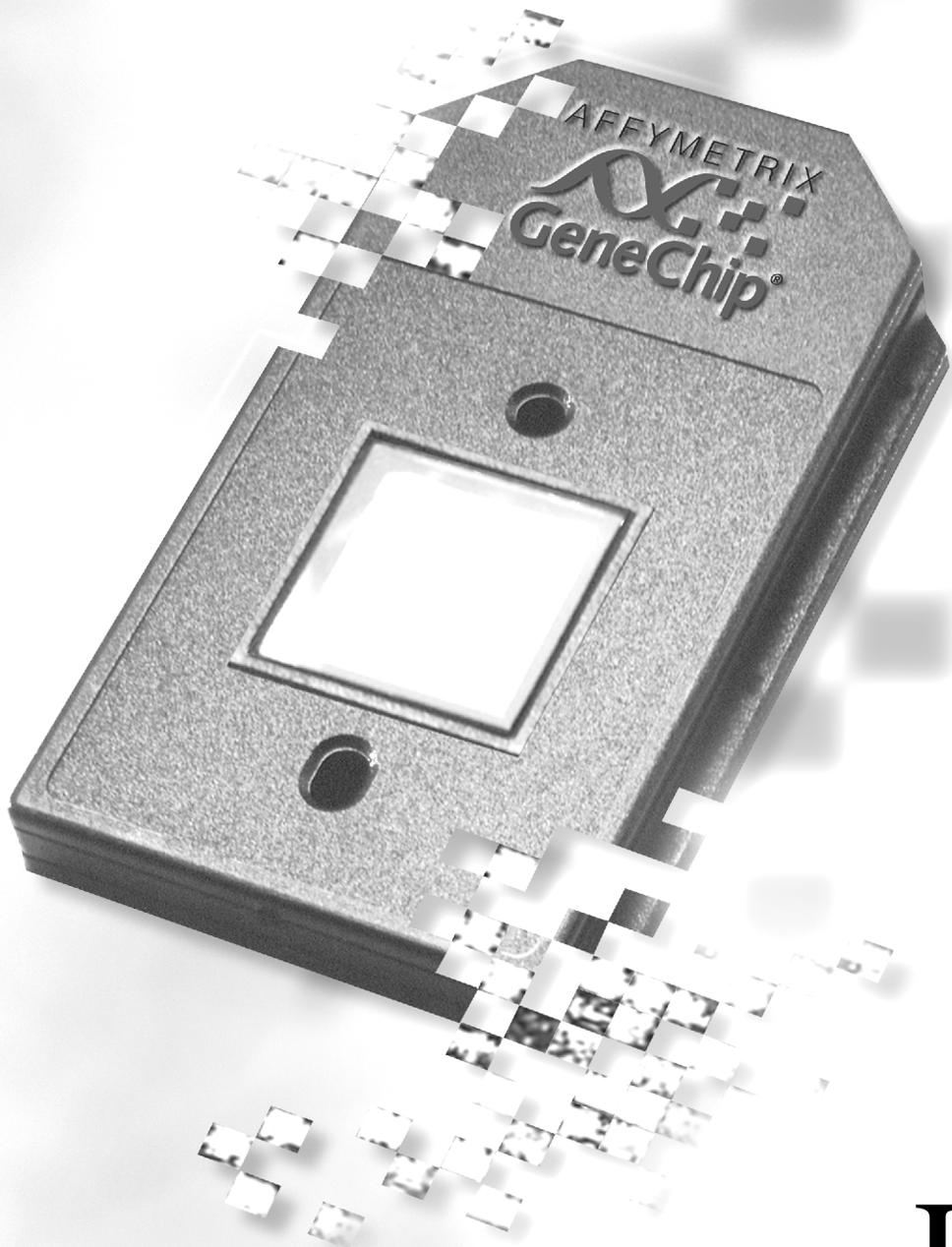
Probes inherently hybridize with slightly different affinities. Relative expression values compare identical probes in the baseline and experiment. As a result, probe specific effects are canceled out.

Adjusted cell intensities (Baseline and Experiment) → Signal log ratio, signal log ratio low, and signal log ratio high

The signal log ratio calculation is an extension of the signal calculation. The discrimination of the log₂ ratio is used to correct for outlier probes. The one-step biweight method is used to compute the average log₂ ratio of the probe set. The upper and lower limits of the 95th confidence interval are reported as signal log ratio high and signal log ratio low.

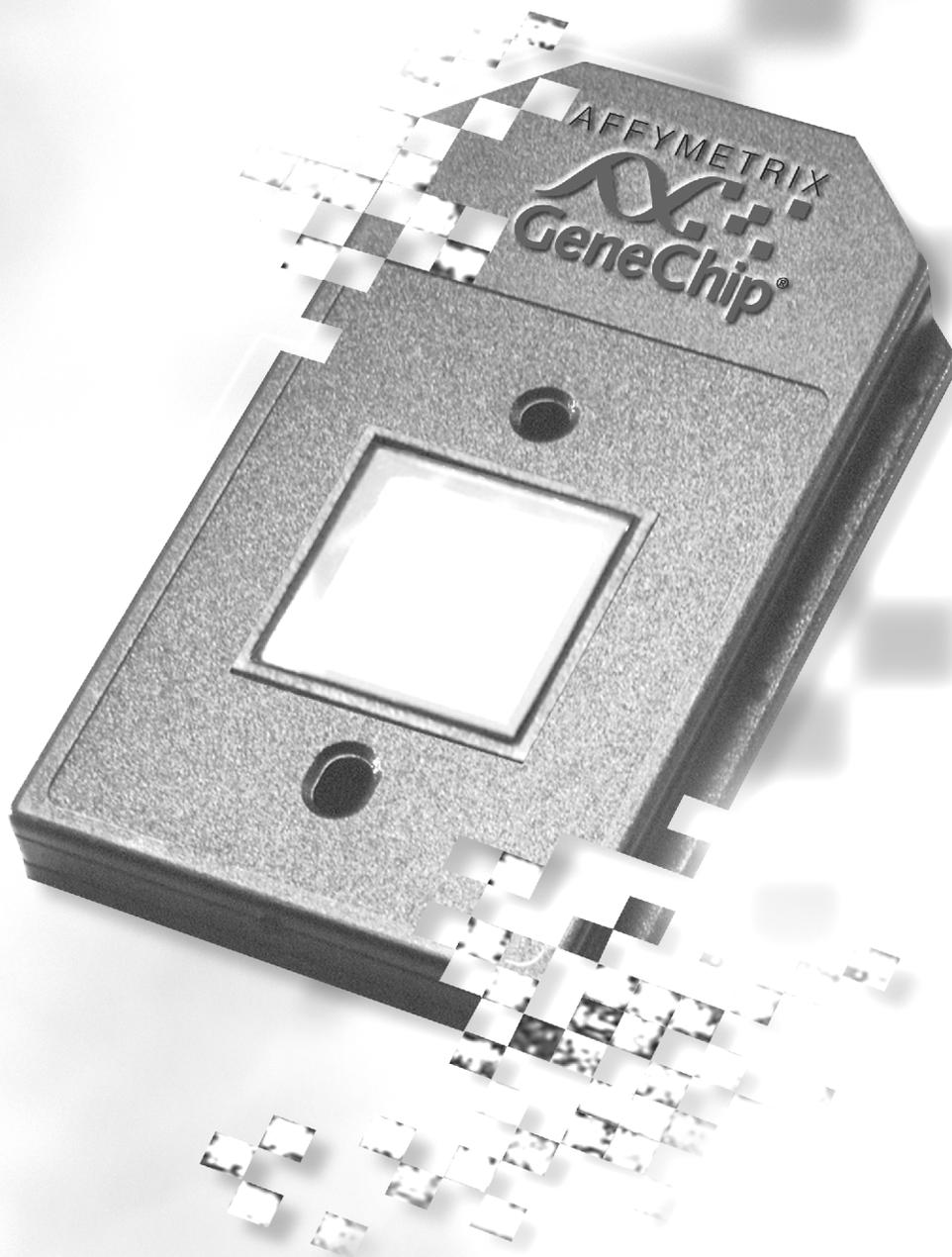
References

1. Hoaglin, D.C., Mosteller, R., Tukey, J.W. 2000. *Understanding Robust and Exploratory Data Analysis*, John Wiley & Sons, New York.
2. Hollander, M., Wolfe, D.A. 1999. *Nonparametric Statistical Methods* (second edition), John Wiley & Sons, New York.
3. Liu, W.M., Mei, R., Bartell, D.M., Di, X., Webster, T.Q., Ryder, T. 2001. Rank-based algorithms for analysis of microarrays, *Proceedings SPIE*, 4266.
4. Wilcoxon, F. 1945. Individual comparisons by ranking methods. *Biometrics*.1:80-83.
5. Roderick, J.A., Little, D., Rubin, B. 1987. *Statistical Analysis With Missing Data*. John Wiley & Sons, New York.



D

Appendix D





Expression Analysis Settings

The expression analysis settings are user-modifiable variables with defaults empirically determined by Affymetrix. They are organized by the tabs of the Expression Analysis Settings dialog box:

- Scaling
- Normalization
- Probe Mask
- Baseline
- Parameters

The Expression algorithm relies on these settings to derive biologically meaningful results from the hybridization intensity data.

Viewing Expression Analysis Settings

1. Click **Analysis Settings** in the shortcut bar, then click **Expression** . Alternatively, select **Tools** → **Analysis Settings** → **Expression** from the menu bar.
⇒ The Expression Analysis Settings dialog box appears ([Figure D.1](#)).

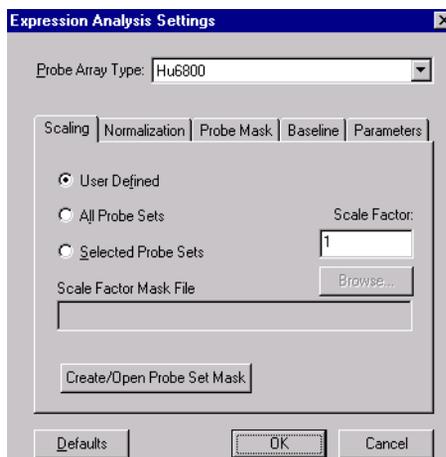


Figure D.1
Expression Analysis Settings, Scaling tab, User Defined scaling option selected

2. Make a selection from the **Probe Array Type** drop-down list.
The settings are specific for the selected probe array type and do not affect the settings for other types of probe arrays.
3. Click a tab to view the different types of expression analysis settings for the selected probe array type.

Scaling Tab

Scaling is a mathematical technique applied to the data from several different probe arrays (of the same type) to minimize discrepancies due to variables such as sample preparation, hybridization conditions, staining, or probe array lot. The Scale Factor is also applied to the noise value.

Microarray Suite offers three types of scaling: User Defined, All Probe Sets, and Selected Probe Sets.

To view the scaling settings:

- Click the Scaling tab in the Expression Analysis Settings dialog box ([Figure D.1](#)).
- Make a selection from the **Probe Array Type** drop-down list.

User Defined Scaling

The **User Defined** scaling option multiplies the signal of each probe set on the array by a user-specified scale factor.

1. In the Scaling tab, choose the **User Defined** option (Figure D.2).

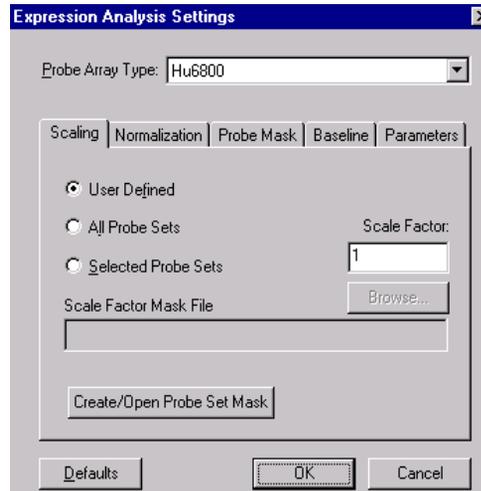


Figure D.2

Expression Analysis Settings, Scaling tab, User Defined scaling option with no scaling specified (scale factor = 1)

2. Enter a number in the **Scale Factor** box.

✓ NOTE

A scale factor of one is equivalent to no scaling.

3. Click **OK** to close the Expression Analysis Settings dialog box.

All Probe Sets Scaling

The **All Probe Sets** scaling option adjusts the trimmed mean signal of a probe array to a user-specified target signal value.

The absolute analysis results (*.chp) from different experiments (probe arrays of the same type) that are scaled to the same target signal using the All Probe Sets scaling option may be directly compared.

1. In the Scaling tab, choose the **All Probe Sets** scaling option (Figure D.3).

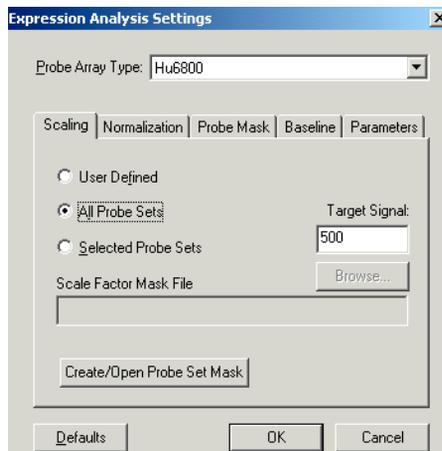


Figure D.3
Expression Analysis Settings, Scaling tab, All Probe Sets scaling option selected

2. Enter a target signal value.
 3. Click **OK** to close the Expression Analysis Settings dialog box.
- Microarray Suite examines all of the probe sets on the array to compute the trimmed mean signal and derive a scale factor for the array so that:
- $$\text{Target Signal} = \text{Scale Factor} \times \text{Trimmed Mean Signal}_{\text{probe array}}$$
- The scale factor standardizes the trimmed mean signal of the array to the target signal (Figure D.4).

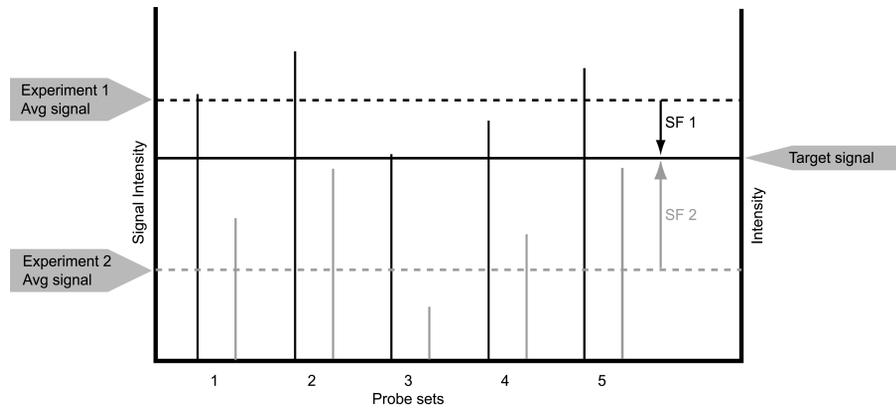


Figure D.4

Scale factors (SF) standardize the average signal of each experiment to a user-specified target signal

Selected Probe Sets Scaling

The **Selected Probe Sets** scaling option adjusts the trimmed mean signal of selected probe sets on a probe array to a user-specified target signal value.

You can directly compare the absolute analysis results (*.chp) of different experiments (probe arrays of the same type) if the same probe sets have been scaled to the same target signal in each experiment using the Selected Probe Sets scaling option.

For this scaling option, Microarray Suite utilizes user-selected probe sets (specified by a Scale Factor mask file) to calculate the trimmed mean signal and derive the scale factor for the probe array so that:

$$\text{Target Signal} = \text{Scale Factor} \times \text{Trimmed Mean Signal}_{\text{selected probe sets}}$$

The scale factor standardizes the trimmed mean signal of an array to the target signal (Figure D.4). Selected Probe Sets scaling does not change the absolute call because the software also multiplies the intensity of each probe set and the probe array noise by the scale factor.

Creating a Scale Factor Mask File

1. In the Scaling tab, choose the **Selected Probe Sets** option (**Figure D.5**).

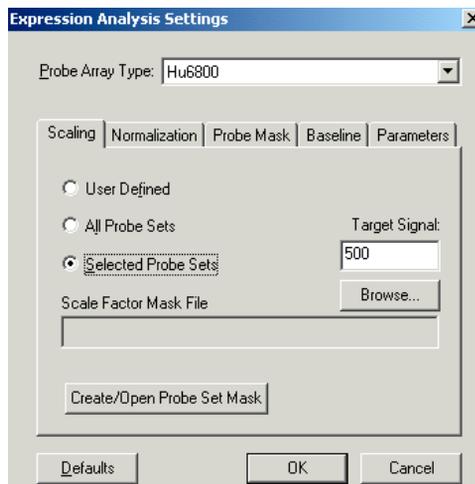


Figure D.5
Expression Analysis Settings, Scaling tab, Selected Probe Sets scaling option

2. Enter a **Target Signal** value, then click **Create/Open Probe Set Mask**.
⇒ The Probe Set Mask File dialog box displays existing mask files (**Figure D.6**).

✓ NOTE

Mask files (.msk) include Scale Factor, Normalization, and Probe mask files. Each type of mask file has a different function and is created in a separate tab of the Expression Analysis Settings dialog box.*



Figure D.6
Probe Set Mask Definition dialog box

3. Enter a name for the scale factor mask file in the **File name** box, then click **Open**.
⇒ The Probe Set Mask Definition dialog box appears (**Figure D.7**).

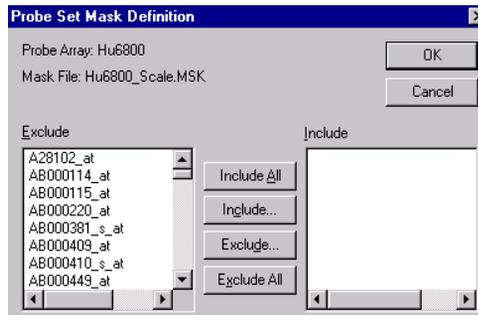


Figure D.7
Probe Set Mask Definition dialog box

4. To select probe sets for the mask file:
 - a. Highlight the desired probe set names in the **Exclude** list.
 - b. Click **Include**.
⇒ This adds the selected probe set names to the **Include** list for the scale factor mask file.

5. To remove a probe set name(s) from the **Include** list:
 - a. Highlight the name(s) in the **Include** list.
 - b. Click **Exclude**.

Include All automatically moves all of the probe set names from the **Exclude** list to the **Include** list.

Exclude All automatically Moves all of the probe set names from the **Include** list to the **Exclude** list.
6. Click **OK** when finished to create the scale factor mask file and close the Probe Set Mask Definition dialog box.

Selecting a Scale Factor Mask File

1. In the Scaling tab, choose the **Selected Probe Sets** option ([Figure D.8](#)).

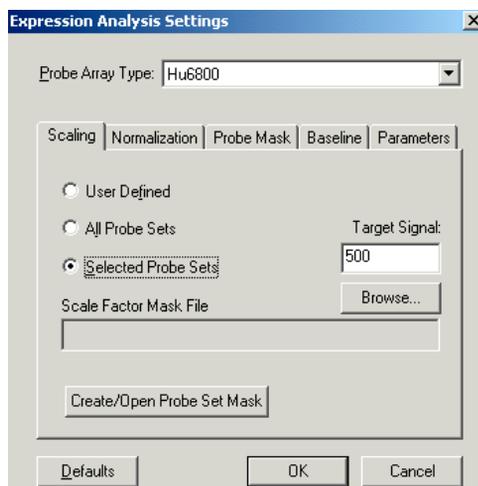


Figure D.8
Expression Analysis Settings, Scaling tab, Selected Probe Sets option

2. Click **Browse** ([Figure D.8](#)).
⇒ The Probe Set Mask File dialog box appears ([Figure D.9](#)).

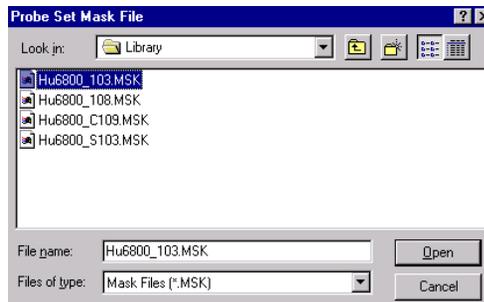


Figure D.9
Probe Set Mask File dialog box

✓ NOTE

Mask files (.msk) are specific for the probe array type. The Microarray Suite software will not open a mask file that is incompatible with the currently selected probe array type.*

3. Double-click the desired *.msk file.
 - ⇒ The Expression Analysis Settings dialog box displays the currently selected *.msk file name in the **Scale Factor Mask File** box (Figure D.10).

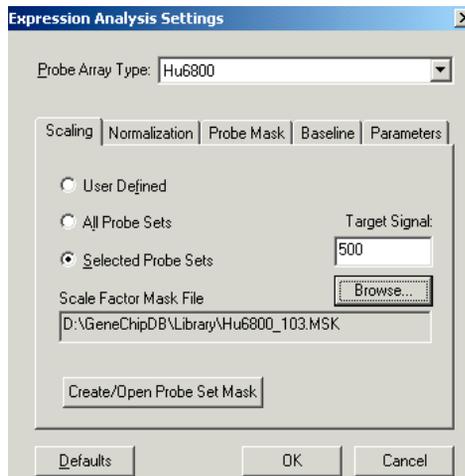


Figure D.10
Expression Analysis Settings, Scaling tab, scale factor mask file selected

Editing a Scale Factor Mask File

1. In the Scaling tab, choose the **Selected Probe Sets** option (Figure D.11).

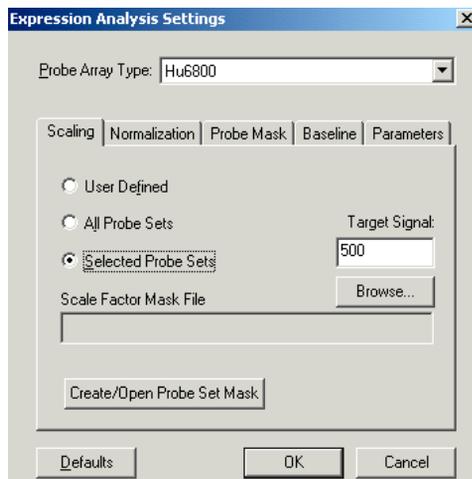


Figure D.11
Expression Analysis Settings, Scaling tab, Selected Probe Sets option

2. Click **Create/Open Probe Set Mask**.
⇒ The Probe Set Mask File dialog box appears (Figure D.12).

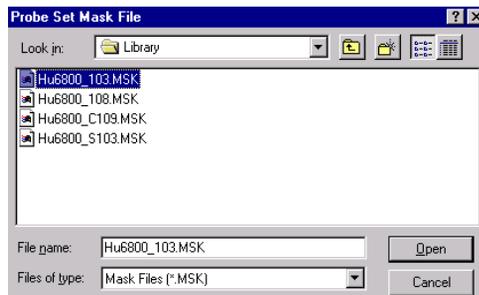


Figure D.12
Probe Set Mask File dialog box

3. Double-click the desired *.msk file.
 - ⇒ The Probe Set Mask Definition dialog box appears (Figure D.13).
 The **Include** list displays the probe sets included in the scale factor mask file. The **Exclude** list displays probe sets not included in the scale factor mask file.

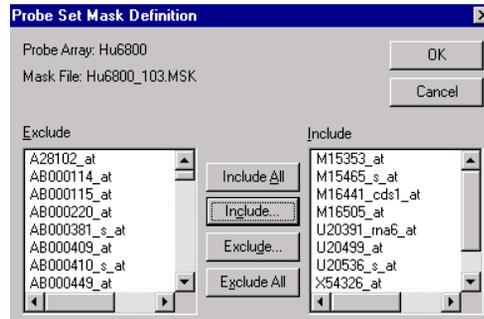


Figure D.13
Probe Set Mask Definition dialog box

4. To remove a probe set(s) from the *.msk file:
 - a. Highlight the probe set name(s) in the **Include** list.
 - b. Click **Exclude**.
5. To add a probe set(s) to the *.msk file, highlight the probe set name(s) in the **Exclude** list, then click **Include**.
 - Include All** automatically moves all of the probe set names from the **Exclude** list to the **Include** list.
 - Exclude All** automatically moves all of the probe set names from the **Include** list to the **Exclude** list.
6. Click **OK** when finished to close the Probe Set Mask Definition dialog box.

Normalization Tab

Normalization is a mathematical technique similar to scaling that enables comparison analysis of an experiment and baseline array.

Microarray Suite offers three types of normalization: User Defined, All Probe Sets, or Selected Probe Sets normalization. All Probe Sets or Selected Probe Sets normalization minimizes discrepancies between an experiment and baseline array due to variables such as sample preparation, hybridization conditions, staining, or probe array lot.

User Defined Normalization

User Defined normalization multiplies the signal of each probe set on an array by a user-specified normalization value.

1. In the Normalization tab, choose the **User Defined** option ([Figure D.14](#)).

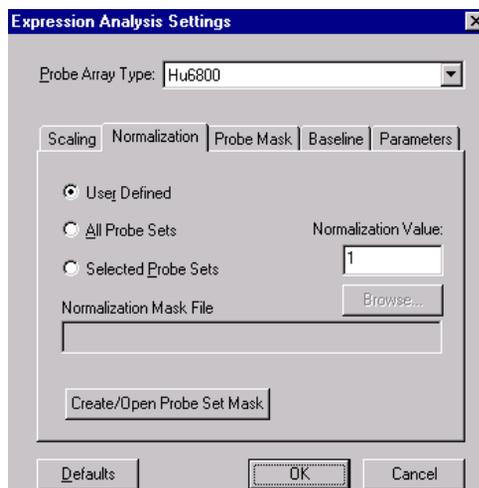


Figure D.14

Expression Analysis Settings, Normalization tab, User Defined normalization option with no normalization specified (Normalization Value = 1)

2. Enter a value in the **Normalization Value** box.

✓ NOTE

A normalization factor of one is equivalent to no normalization.

3. Click **OK** to close the Expression Analysis Settings dialog box.

All Probe Sets Normalization

All Probe Sets normalization adjusts or *normalizes* the trimmed mean signal of the experiment to the trimmed mean signal of the baseline (Figure D.15).

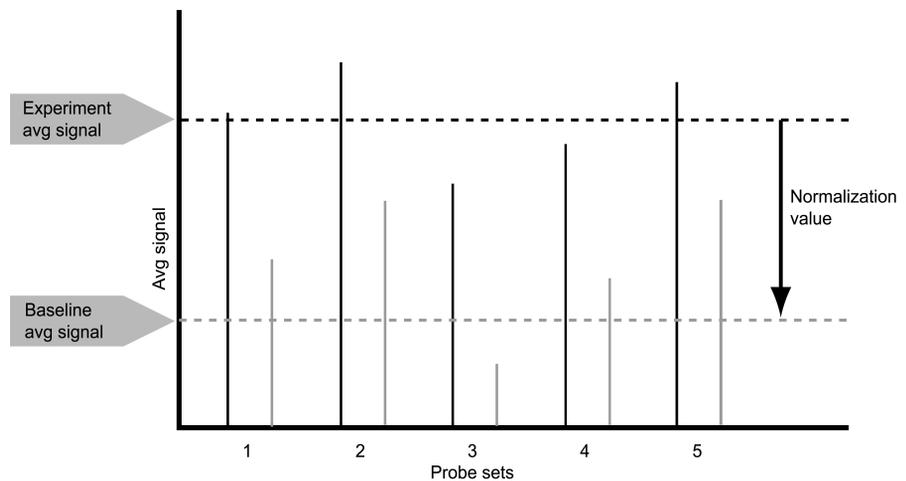


Figure D.15

Normalization Value normalizes the average signal of the experiment to the average signal of the baseline

1. In the Normalization tab, choose the **All Probe Sets** normalization option (Figure D.16).

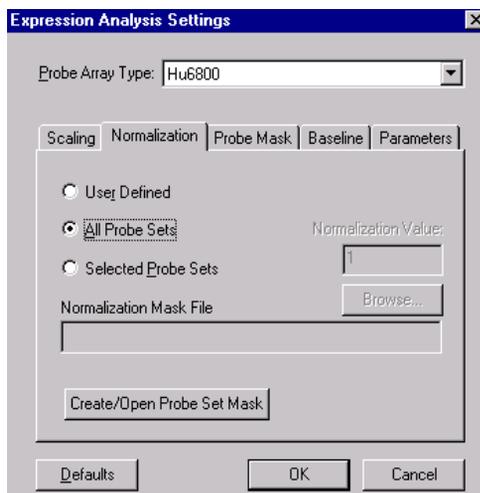


Figure D.16
Expression Analysis Settings, Normalization tab, All Probe Sets normalization option

2. Click **OK** to close the Expression Analysis Settings dialog box.
The software examines all probe sets on the experiment or baseline array to compute a trimmed mean signal for the experiment and a trimmed mean signal for the baseline.

It computes a normalization value so that:

$$\text{Trimmed Mean Signal}_{\text{baseline}} = (\text{Normalization Value}) \times (\text{Trimmed Mean Signal}_{\text{experiment}})$$

Selected Probe Sets Normalization

Selected Probe Sets normalization adjusts or normalizes the trimmed mean signal of the experiment to the trimmed mean signal of the baseline (**Figure D.15**).

For this normalization option, Microarray Suite utilizes user-selected probe sets (specified by a normalization mask file) to compute the trimmed mean signal of the experiment and baseline, and derive a normalization value so that:

$$\text{Trimmed Mean Signal}_{\text{baseline selected probe sets}} = \text{Normalization Value} \times \text{Trimmed Mean Signal}_{\text{experiment selected probe sets}}$$

Creating a Normalization Mask File

1. In the Normalization tab, choose the **Selected Probe Sets** option (Figure D.17).

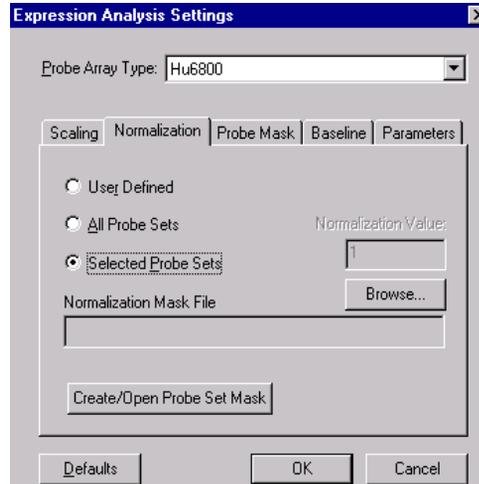


Figure D.17
Expression Analysis Settings, Normalization tab, Selected Probe Sets option

2. Click **Create/Open Probe Set Mask**.

⇒ The Probe Set Mask File dialog box appears (Figure D.18).

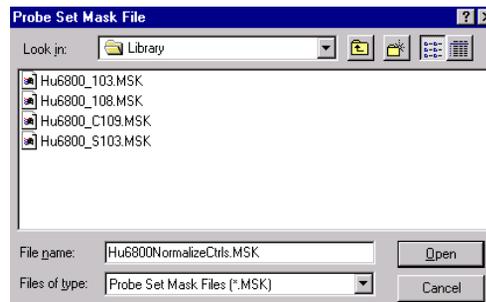


Figure D.18
Probe Set Mask File dialog box

✓ NOTE

The mask files (.msk) include Scale Factor, Normalization, or Probe mask files. Each type of mask file has a different function and is created in a separate tab of the Expression Analysis Settings dialog box.*

3. Enter a name for the new normalization mask file in the **File name** box and click **Open**.
⇒ The Probe Set Mask Definition dialog box appears (**Figure D.19**).

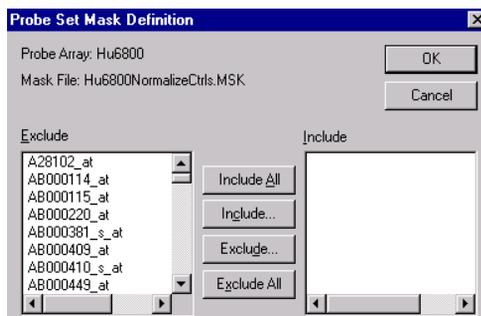


Figure D.19
Probe Set Mask Definition dialog box

4. Highlight the desired probe set names in the **Exclude** list. Click **Include** to add them to the **Include** list for the new *.msk file.
5. To remove a probe set name(s) from the **Include** list, highlight the name(s) in the **Include** list, then click **Exclude**.
Include All automatically moves all of the probe set names from the **Exclude** list to the **Include** list.
Exclude All automatically moves all of the probe set names from the **Include** list to the **Exclude** list.
6. Click **OK** when finished to create the normalization mask file and close the Probe Set Mask Definition dialog box.

Selecting a Normalization Mask File

1. In the Normalization tab, click **Browse** (Figure D.20).
⇒ The Probe Set Mask File dialog box appears (Figure D.21).

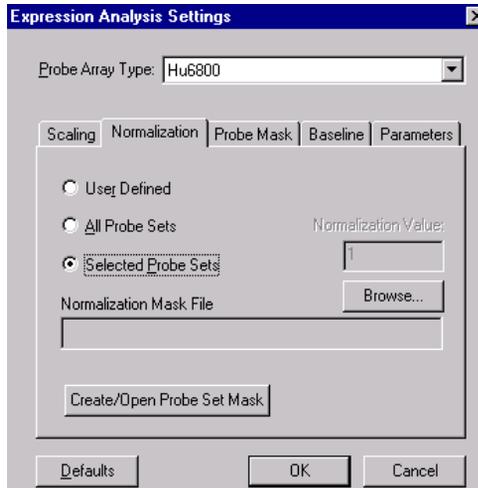


Figure D.20
Expression Analysis Settings, Normalization tab, Selected Probe Sets option

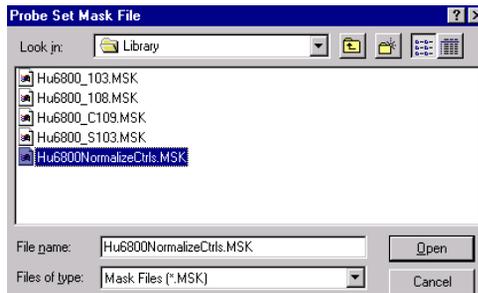


Figure D.21
Probe Set Mask File dialog box

✓ NOTE

Mask files are specific for the probe array type. Microarray Suite will not open a mask file that is incompatible with the currently selected probe array type.

2. Double-click the desired *.msk file.
 - ⇒ The Probe Set Mask File dialog box closes and the Expression Analysis Settings dialog box displays the selected *.msk in the **Normalization Mask File** box (Figure D.22).

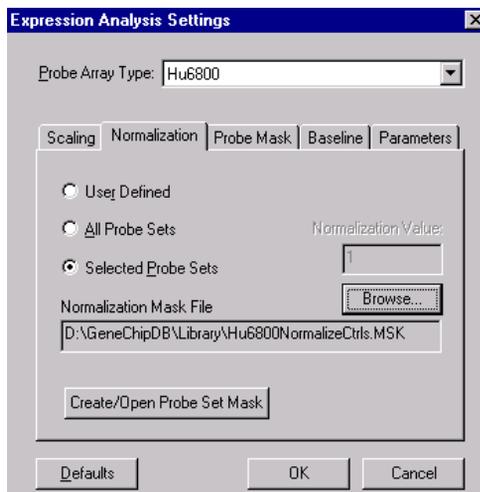


Figure D.22
Expression Analysis Settings, Normalization tab, normalization mask file selected

Editing a Normalization Mask File

1. In the Normalization tab, choose the **Selected Probe Sets** option (Figure D.23).

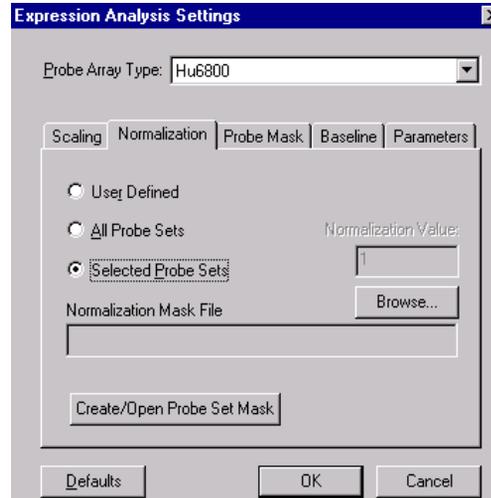


Figure D.23

Expression Analysis Settings, Normalization tab, Selected Probe Sets option

2. Click **Create/Open Probe Set Mask** (Figure D.23).
⇒ The Probe Set Mask File dialog box appears (Figure D.24).



Figure D.24

Probe Set Mask File dialog box

3. Double-click the desired *.msk file.
⇒ The Probe Set Mask Definition dialog box appears (**Figure D.25**).
The **Include** list displays the probe sets included in the *.msk file. The **Exclude** list displays the probe sets not included in the *.msk file.

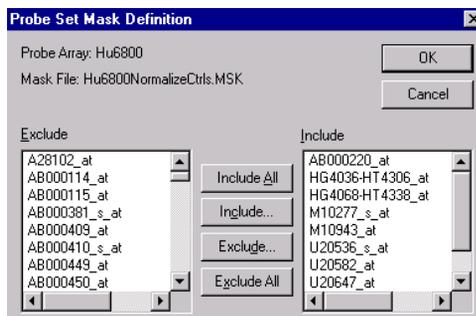


Figure D.25
Probe Set Mask Definition dialog box

4. To remove a probe set(s) from the *.msk file:
 - a. Highlight the probe set name(s) in the **Include** list.
 - b. Click **Exclude**.
5. To add a probe set(s) to the *.msk file:
 - a. Highlight the probe set name(s) in the **Exclude** list
 - b. Click **Include**.

Include All automatically moves all of the probe set names from the **Exclude** list to the **Include** list.

Exclude All automatically moves all of the probe set names from the **Include** list to the **Exclude** list.
6. Click **OK** when finished to close the Probe Set Mask Definition dialog box.

Probe Mask Tab

User-selected probe pairs may be excluded or *masked* from an expression analysis. In the Probe Mask tab ([Figure D.26](#)), you can create a probe mask file that specifies the probe pairs to exclude from an analysis.

Microarray Suite also offers advanced features that automatically generate three types of probe masks: cross hybridization, hybridization, or spike mask. (See [Appendix F](#) for more information about probe masks.)

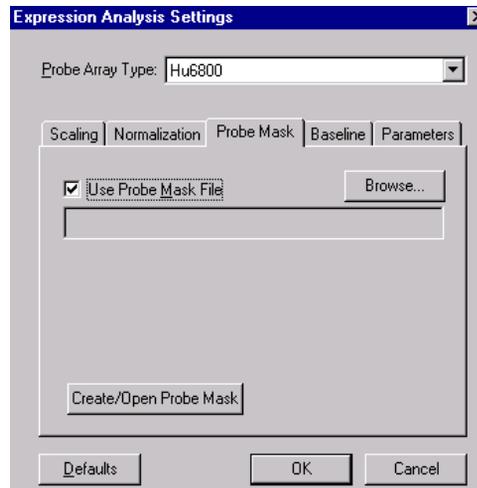


Figure D.26
Expression Analysis Settings, Probe Mask tab

Creating a Probe Mask

1. In the Probe Mask tab, click **Create/Open Probe Mask** (Figure D.26).
⇒ The Probe Mask File dialog box appears (Figure D.27).

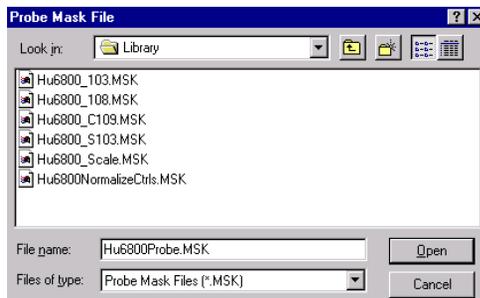


Figure D.27
Probe Mask File dialog box

✓ NOTE

*The *.msk files include Scale Factor, Normalization, or Probe mask files. Each type of mask file has a different function and is created in a separate tab of the Expression Analysis Settings dialog box.*

2. Enter a name for the new probe mask file in the **File name** box and click **Open**.
⇒ The Probe Mask Definition dialog box appears (Figure D.28) and displays the probe array type and the probe mask name.

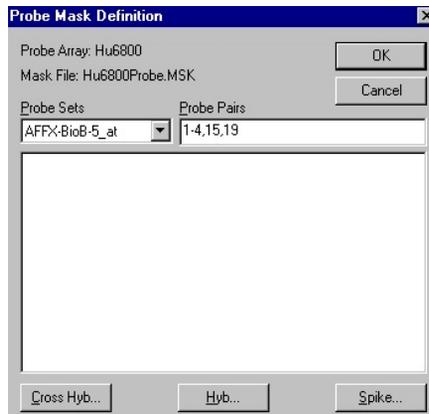


Figure D.28
Probe Mask Definition dialog box

3. Select the desired probe set name from the **Probe Sets** drop-down list and enter the desired probe pair numbers in the **Probe Pairs** field (following the example format: 1-4,7,15,19).
These probe pairs will be omitted from the analysis.
4. Click the box below the **Probe Sets** drop-down list to add the probe pairs to the probe mask (**Figure D.29**).

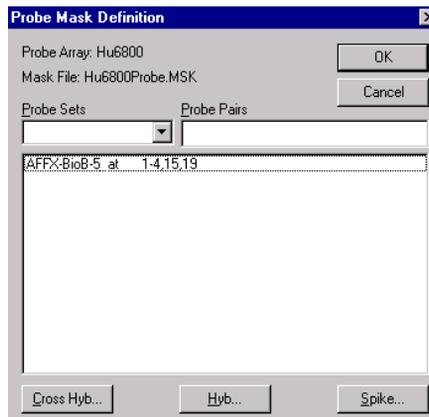


Figure D.29
Probe pairs selected for the probe mask file

5. Repeat [step 3](#) and [step 4](#) to specify additional probe pairs for the probe mask.
6. Click **OK** when finished adding probe pairs to the probe mask.

Selecting a Probe Mask

1. In the Probe Mask tab, choose the **Use Probe Mask File** option ([Figure D.30](#)).

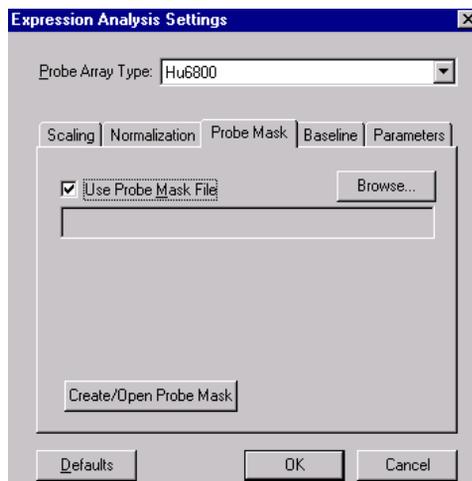


Figure D.30
Expression Analysis Settings, Probe Mask tab

2. Click **Browse**.
⇒ The Probe Mask File dialog box appears ([Figure D.31](#)).

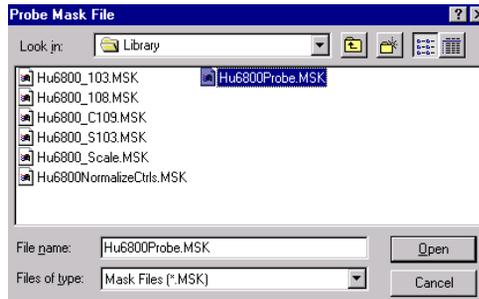


Figure D.31
Probe Mask File dialog box

3. Double-click the desired probe mask.
⇒ The Probe Mask tab displays the selected probe mask (Figure D.32).

✓ NOTE

Mask files are specific for the probe array type. Microarray Suite will not open a mask file that is incompatible with the currently selected probe array type.

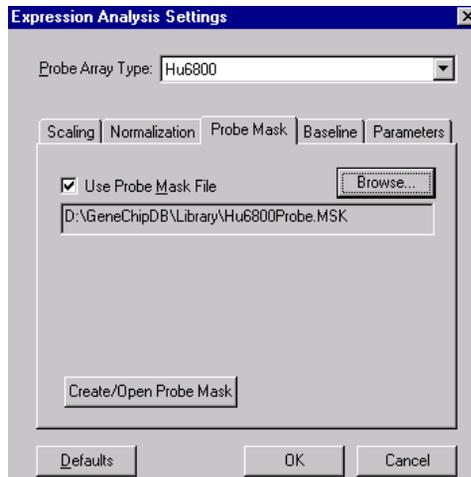


Figure D.32
Expression Analysis Settings, Probe Mask tab

Editing a Probe Mask

1. In the Probe Mask tab, click **Create/Open Probe Mask** (Figure D.33).
⇒ The Probe Mask File dialog box appears (Figure D.34).

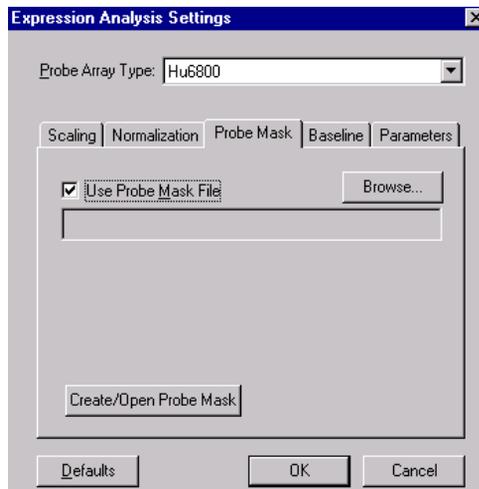


Figure D.33
Expression Analysis Settings

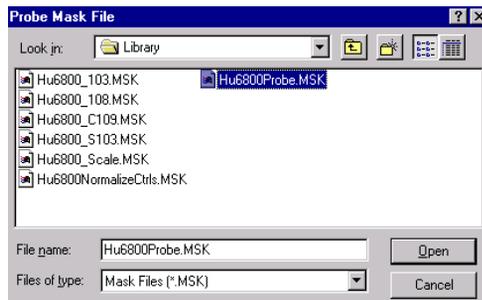


Figure D.34
Probe Mask File dialog box

2. Click the desired probe mask file and click **OK** (or double-click the file name).
 - ⇒ The Probe Mask Definition dialog box displays the probe pairs in the selected probe mask (**Figure D.35**).

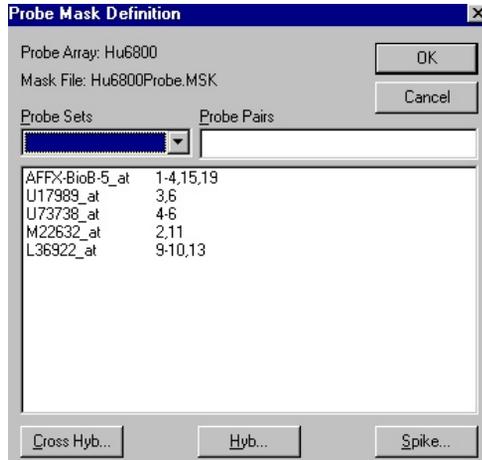
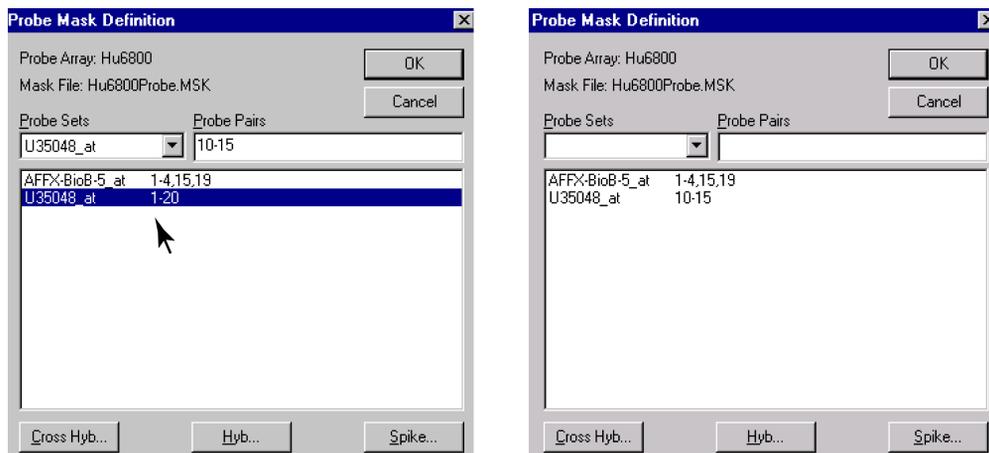


Figure D.35
Probe Mask Definition dialog box

3. Click the probe set/probe pair entry you wish to edit.
 - ⇒ The Probe Sets and Probe Pairs box automatically display the selection (**Figure D.36**).

**Figure D.36**

Edit entries in the Probe Pairs box and click the lower box (left) to update the probe pair list (right)

4. Edit or delete the entry in the Probe Pairs field, then click the lower field to update the list of probe pairs (Figure D.36).
5. Click **OK** when finished to close the Probe Mask Definition dialog box.

Baseline Tab

In the baseline tab, the user selects the analysis output file (*.chp) that will serve as the baseline in a comparison expression analysis (Figure D.37).

✓ NOTE

Do not choose the **Use Baseline Comparison File** option for an absolute analysis.

1. In the Baseline tab, choose the **Use Baseline Comparison File** option (Figure D.37).

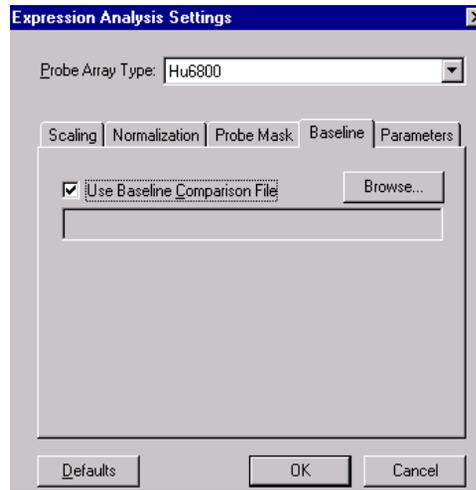


Figure D.37
Expression Analysis Settings, Baseline tab

2. Click **Browse**.
⇒ The Baseline Comparison File dialog box appears (Figure D.38).

✓ NOTE

The baseline file must be derived from the same type of probe array selected in the **Probe Array Type** drop-down list in the Expression Analysis Settings dialog box (Figure D.37).

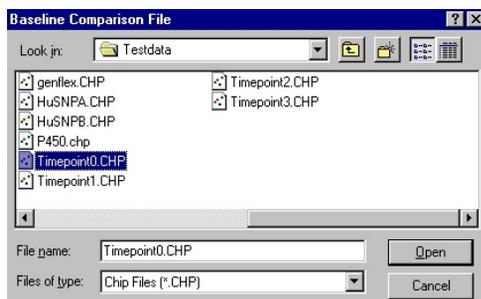


Figure D.38
Baseline Comparison File dialog box

3. Double-click the baseline *.chp and click **Open**.
The Baseline tab displays the selected baseline file for comparison analyses (**Figure D.39**).

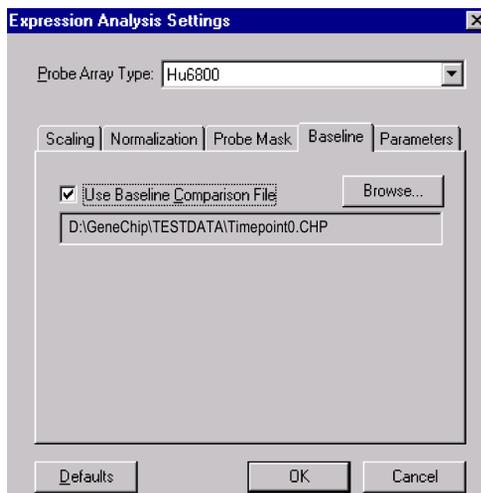


Figure D.39
Expression Analysis Settings, Baseline tab

Parameters Tab

The Parameters tab displays user-modifiable Statistical Expression algorithm parameters ([Figure D.40](#)). [Table D.1](#) describes the parameters.

✓ NOTE

Versions of Microarray Suite lower than 5.0 run the Empirical Expression algorithm. (See [Appendix E](#) for more information about the Empirical Expression algorithm.)

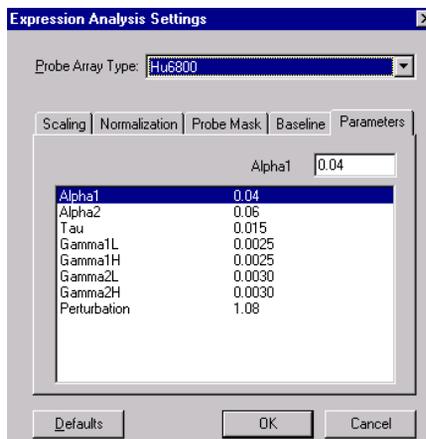


Figure D.40
Expression Analysis Settings, Parameters tab

The parameter default values were determined through extensive empirical testing at Affymetrix. Changing a parameter value affects the:

- algorithm output (see [Appendix E](#) for more information about the Statistical and Empirical Expression algorithm outputs)
- assay sensitivity (ability to make true detection or change calls)
- assay specificity (ability to minimize false detection or change calls)

Table D.1
User-modifiable Statistical Expression algorithm parameters

Parameter	Description	Effect of Changing Parameter
Alpha1 (α_1)	The p-value threshold that defines present and marginal detection calls. If: $p < \alpha_1$, the call is present $\alpha_1 \leq p < \alpha_2$, the call is marginal	Decreasing α_1 may reduce the number of false present calls (increases specificity) and the number of true present calls (decreases sensitivity).
Alpha2 (α_2)	The p-value threshold that defines marginal and absent detection calls. If: $\alpha_1 \leq p < \alpha_2$, the call is marginal $p \geq \alpha_2$, the call is absent	Increasing α_2 may reduce the number of false absent calls (increases sensitivity) and the number of true absent calls (decreases specificity).
Tau (τ)	Threshold for the probe set discrimination score. If the discrimination score $< \tau$, the difference between the PM and MM cells in the probe set is considered insignificant. The algorithm does not make a call for the probe set.	Increasing τ may reduce the number of false present calls (increases specificity) and the number of true present calls (decreases sensitivity).
Gamma1L (γ_{1L})	The lower limit of γ_1 which is the p-value threshold that defines increase and marginal increase change calls.	Decreasing γ_{1L} and γ_{1H} decreases γ_1 . Decreasing γ_1 may reduce the number of false increase and decrease calls (increases specificity) and may also reduce the number of true increase and decrease calls (decreases sensitivity).
Gamma1H (γ_{1H})	The upper limit of γ_1 which is the p-value threshold that defines increase and marginal increase change calls.	
Gamma2L (γ_{2L})	The lower limit of γ_2 which is the p-value threshold that defines marginal increase and no change calls.	Increasing γ_{2L} and γ_{2H} increases γ_2 . Increasing γ_2 may reduce the number of false no change calls (increases specificity) and may also reduce the number of true no change calls (decreases sensitivity).
Gamma2H (γ_{2H})	The upper limit of γ_2 which is the p-value threshold that defines marginal and no change change calls.	
Perturbation (d)	Determines the spread between the three normalization factors the algorithm computes for an experiment and baseline in a comparison analysis.	Increasing d may increase the number of true no change calls (increases sensitivity) and may also increase the number of false no change calls (decreases specificity).

Changing a Parameter Value

1. In the Parameters tab and make a selection from the **Probe Array Type** drop-down list.
2. Click the parameter you wish to change.
⇒ The parameter box displays the current value (Figure D.41).

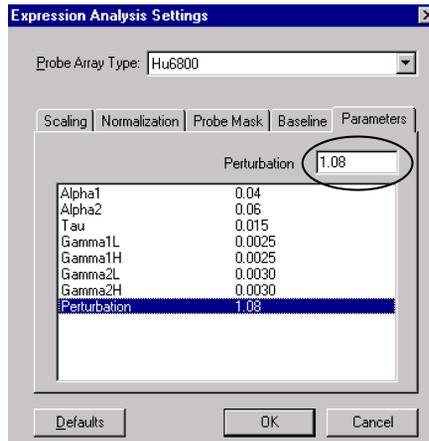
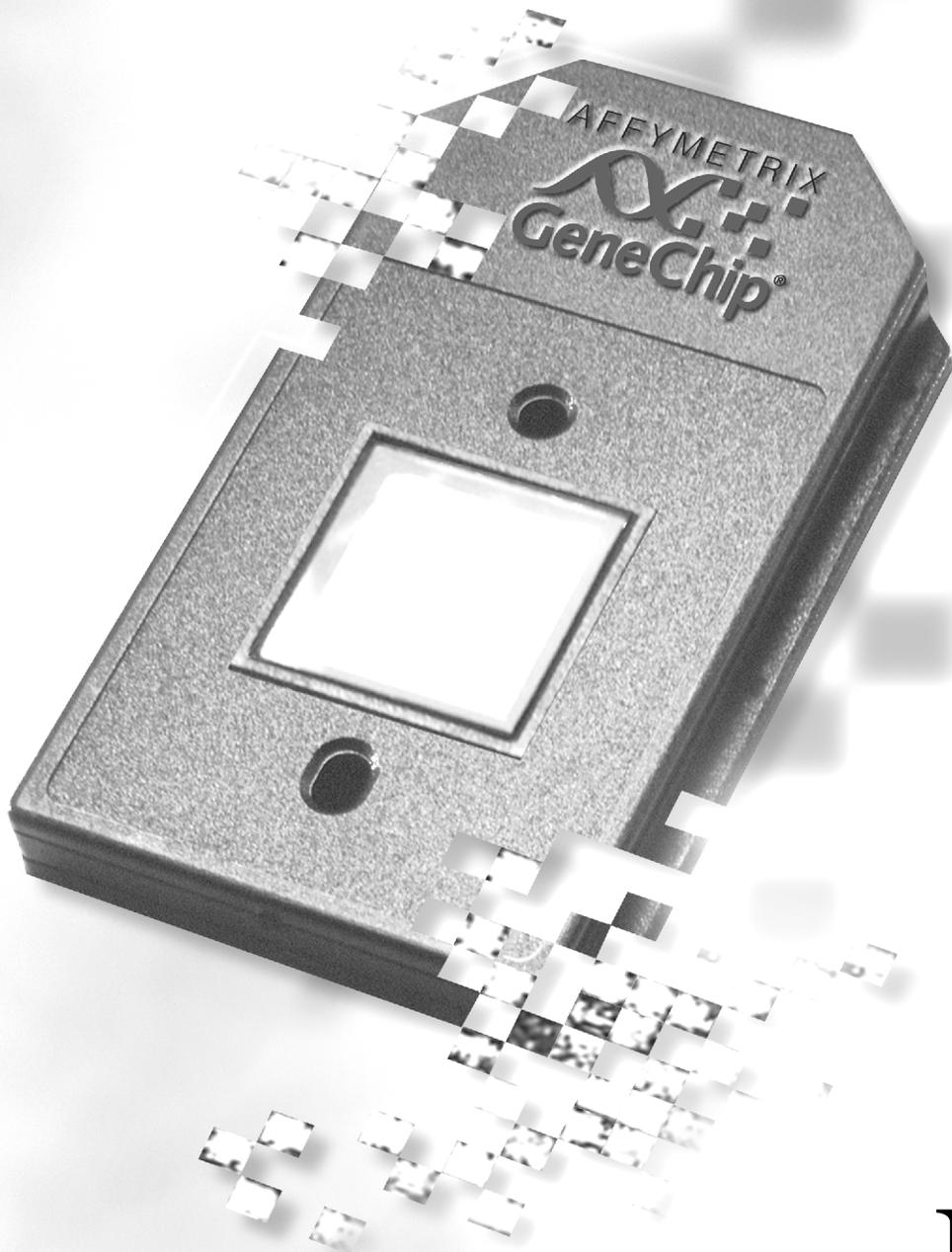


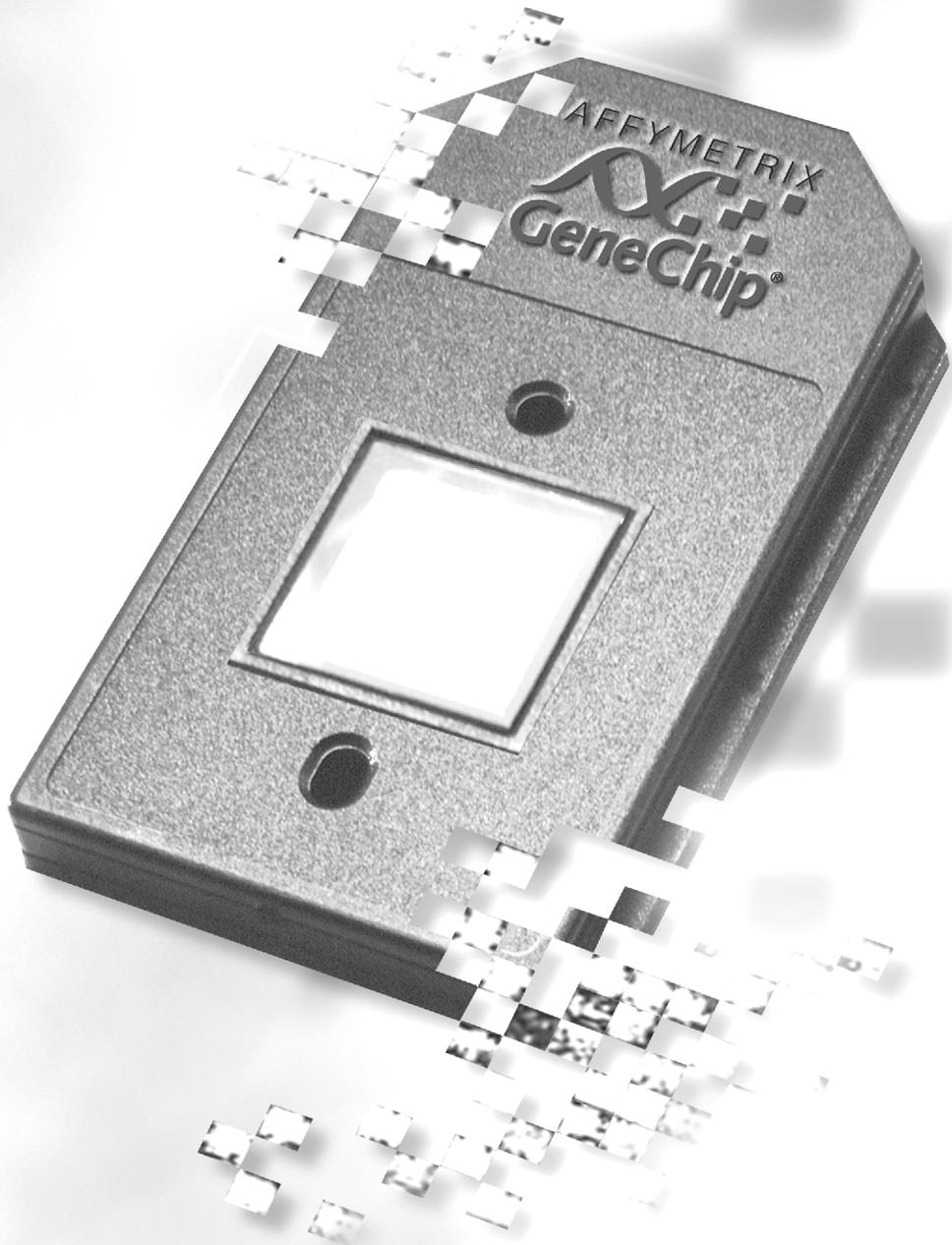
Figure D.41
Expression Analysis Settings, Parameters tab

3. Enter a new value for the parameter in the parameter box.
4. Click **OK** to close the Expression Analysis Settings dialog box.



E

Appendix E





Expression Analysis Metrics

Microarray Suite runs the Statistical expression algorithm. Previous versions of Microarray Suite (lower than 5.0) run the Empirical expression algorithm.

Statistical Expression Algorithm

Absolute Results

Signal

A measure of the abundance of a transcript.

Detection

The call that indicates whether the transcript is detected (P, present), undetected (A, absent), or at the limit of detection (M, marginal).

Detection p-value

p-value that indicates the significance of the detection call.

Stat Pairs

The number of probe pairs for a particular probe set on an array.

Stat Pairs Used

The number of probe pairs per probe set used in the analysis. This may be the total number of probes per probe set on the probe array or the number of probe pairs in a pre-designated subset (for example, probe pairs specified by a probe mask file and/or a masked image). Stat Pairs Used = total probe

pairs per probe set – (probe pairs masked in a mask file) – (probe pairs masked in the image) – (saturated MM probe pairs).

Comparison Results

Signal Log Ratio

The change in the expression level of a transcript between a baseline and an experiment array. This change is expressed as the \log_2 ratio. A \log_2 ratio of 1 is equal to a fold change of 2.

Signal Log Ratio Low

The lower limit of the log ratio within a 95% confidence interval.

Signal Log Ratio High

The upper limit of the log ratio within a 95% confidence interval.

Change

The call that indicates the change in the transcript level between a baseline and experiment (increase (I), marginal increase (MI), no change (NC), marginal decrease (MD), decrease (D))

Change p-value

p-value that indicates the significance of the change call.

Stat Common Pairs

The intersection of the probe pairs from the baseline and experiment that are used by the Expression algorithm to make the change call in a comparison analysis.

Empirical Expression Algorithm

Previous versions of Microarray Suite (lower than 5.0) run the Empirical Expression algorithm.

Absolute Analysis Results

Positive & Negative

The number of probe pairs scored positive or negative. A probe pair is called positive if the intensity of the PM probe cell is significantly greater than that of the corresponding MM probe cell. A probe pair is called negative if the intensity of the MM probe cell is significantly greater than that of the corresponding PM probe cell. To evaluate the intensity, the algorithm calculates the ratio and difference associated with each probe pair and compares these values to the Statistical Difference Threshold (SDT) and the Statistical Ratio Threshold (SRT).

A probe pair is Positive if: $PM - MM \geq SDT$ and $PM/MM \geq SRT$

A probe pair is Negative if: $MM - PM \geq SDT$ and $MM/PM \geq SRT$

The SDT is a function of the noise (Q) and is calculated by the software: $SDT = Q * SDT_{mult}$. The SDT_{mult} and the SRT are user-modifiable parameters. The SDT_{mult} is set at 2.0 for the standard staining protocol or 4.0 for the antibody amplification protocol. (Refer to the *Expression Analysis Technical Manual* or the *HuSNP Mapping Assay User Manual*). The default SRT value is 1.5.

✓ NOTE

Increasing the SDT_{mult} and SRT increases analysis stringency, reducing these thresholds decreases analysis stringency.

The number of positive and negative probe pairs is determined for every probe set and are used to derive parameters that describe probe set performance.

Pairs

The number of probe pairs for a particular probe set on an array.

Pair Used

The number of probe pairs per probe set used in the analysis. This may be the total number of probes per probe set on the probe array or the number of probe pairs in a pre-designated subset (for example, probe pairs specified by a probe mask file and/or a masked image).

Pairs Used = total probe pairs per probe set – (probe pairs masked in a mask file) – (probe pairs masked in the image).

Pairs in Avg

A *trimmed* probe set that excludes probes with extremely intense or weak signal from the analysis. If 8 or fewer probe pairs are used, Pairs in Avg = Pairs Used (or the number of probe pairs per probe set minus any that are masked).

Super scoring is performed if more than 8 probe pairs are used. Superscoring is a process that excludes probe pairs from calculation of the Avg Diff and Log Avg Ratio if they are outside a given intensity range. Microarray Suite calculates the mean and standard deviation of the intensity differences (PM – MM) for an entire probe set (excluding the highest and lowest values). Those values within a set number of standard deviations (STP) are included in the calculation of the Avg Diff or Log Avg Ratio. The STP is a user-modifiable parameter with a default value = 3.

Pos Fraction

positive probe pairs/# probe pairs used

Log Avg

Describes the hybridization performance of a probe set and is determined by calculating the ratio of the PM/MM intensities for each probe pair in a probe set, taking the logs of the resulting values, and averaging them for the probe set:

$\text{Log Avg} = 10 \times \{ [\sum \log (\text{PM/MM})] / \text{Pairs in Avg} \}$

✓ NOTE

Log Avg = 0 indicates random cross hybridization. The higher the Log Avg, the more confidence the transcript is present.

Pos/Neg

The ratio of Positive probe pairs to Negative probe pairs in a probe set (# Positive probe pairs / # Negative probe pairs).

Avg Diff

$$\text{Avg Diff} = \Sigma (\text{PM} - \text{MM}) / \text{Pairs in Avg}$$

This parameter serves as a relative indicator of the level of expression of a transcript. It is used to determine the change in the hybridization intensity of a given probe set between two different experiments.

The Avg Diff is calculated by taking the difference between the PM and MM of every probe pair (excluding the probe pairs where PM – MM is outside the STP standard deviation of the mean of PM-MM) in a probe set and averaging the differences for the entire probe set.

✓ NOTE

The Avg Diff cannot be used to compare the hybridization intensity levels of two different probe sets on the same array.

Absolute Call

Each transcript in an absolute analysis has three possible Absolute Call outcomes: *Present* (P), *Absent* (A), or *Marginal* (M). The absolute call is derived from the Pos/Neg, Positive Fraction, and Log Avg absolute call metrics. Each absolute call metric is weighted and entered into a decision matrix to determine the status of the transcript.

Comparison Analysis Results

Increase

A probe pair is considered to increase if the intensity difference between the PM and MM probe cells in the experimental sample is significantly higher than in the baseline sample. Two criteria must be met for a probe pair to show a significant increase:

$(\text{PM} - \text{MM})_{\text{exp}} - (\text{PM} - \text{MM})_{\text{base}} \geq \text{Change Threshold (CT)}, \text{ and}$

$[(\text{PM} - \text{MM})_{\text{exp}} - (\text{PM} - \text{MM})_{\text{base}}] / \max [Q/2, \min(|\text{PM} - \text{MM}|_{\text{exp}}, |\text{PM} - \text{MM}|_{\text{base}})] \geq \text{Percent Change Threshold}/100$

Decrease

A probe pair is considered to decrease if the intensity difference between the PM and MM probe cells in the experimental sample is significantly lower than in the baseline sample. Two criteria must be met for a probe pair to show a significant decrease:

$(PM - MM)_{base} - (PM - MM)_{exp} \geq \text{Change Threshold (CT)}$, and

$[(PM - MM)_{base} - (PM - MM)_{exp}] / \max [Q/2, \min(|PM - MM|_{exp}, |PM - MM|_{base})] \geq \text{Percent Change Threshold}/100$

The software calculates the Change Threshold (CT) using the SDT (Statistical Difference Threshold) of both the experimental and baseline data. Alternatively, the user may define the CT by entering a value for the CT Multiplier (in the Parameters tab of the Expression Analysis Settings dialog box), which is multiplied by the noise (Q) of the baseline or experimental data, whichever is greater. The Percent Change Threshold is a user-specified value (also set in the Parameters tab of the Expression Analysis Settings dialog box).

Inc Ratio

For each transcript: # Increased probed pairs / # probe Pairs Used

Dec Ratio

For each transcript: # Decreased Probe pairs / # probe Pairs Used

Pos Change

Positive probe pairs_{exp} - # Positive probe pairs_{baseline}

Neg Change

Negative probe pairs_{exp} - # Negative probe pairs_{baseline}

Inc/Dec

For each transcript: the # increased probe pairs / # decreased probe pairs

DPos-DNeg Ratio

(Positive Change – Negative Change)/# probe Pairs Used

The DPos – DNeg Ratio and Log Avg Ratio Change are usually positive when a transcript changes from a very low to a relatively high expression level and are typically negative when the expression level changes from a high to a very low or undetectable level. Both metrics may have values close to zero if the transcript is present in both the baseline and experimental samples despite an increase or decrease in the level of the transcript.

Log Avg Ratio Change

$\text{Log Avg}_{\text{exp}} - \text{Log Avg}_{\text{base}}$

The difference between the Log Avg Ratio of the baseline and experimental probe array data (in a comparison analysis) for each transcript. The Log Avg Ratios are recomputed for each for each probe set based on probe pairs used in both the baseline and experimental probe arrays (the recomputed values are not displayed by the software).

Difference Call

Each transcript in a comparison analysis has five possible Difference Call outcomes: (1) *Increase (I)*, (2) *Marginally Increase (MI)*, or (3) *Decrease (D)*, (4) *Marginally Decrease (MD)*, and (5) *No Change (NC)*. The difference call is derived from the comparison metrics: Max [Increase/Total, Decrease/Total], Increase/Decrease Ratio, Log Average Ratio Change, and Dpos – Dneg Ratio. Each comparison metric is weighted and entered into a decision matrix to determine the status of the transcript.

Avg Diff Change

The Avg Diff values are recomputed for each probe set based on probe pairs used in both the baseline and experimental probe arrays (the recomputed values are not displayed by the software).

$\text{Avg Diff Change} = \text{Avg Diff}_{\text{exp}} - \text{Avg Diff}_{\text{baseline}}$

B=A

An asterisk (*) in this column indicates the transcript is called absent (A) in the baseline.

Fold Change

The Fold Change indicates the relative change in the expression levels between the experiment and baseline targets. The Fold Change for a transcript is a positive number when the expression level in the experiment increases compared to the baseline and is a negative number when the expression level in the experiment declines. The Fold Change (FC) is calculated as:

$$FC = \left(\frac{AvgDiffChange}{\max[\min(AvgDiff_{base}, AvgDiff_{exp}), Q_M \times Q_C]} \right)^+ \begin{cases} +1 & \text{if } (AvgDiff_{exp} \geq AvgDiff_{base}) \\ -1 & \text{if } (AvgDiff_{exp} < AvgDiff_{base}) \end{cases}$$

$$AvgDiffChange = AvgDiff_{exp} - AvgDiff_{base}$$

$$Q_C = \max(Q_{exp}, Q_{base})$$

$$Q_M = 2.1 \text{ for a } 50\mu\text{m feature or } Q_M = 2.8 \text{ for a } 24\mu\text{m feature}$$

The normalized or scaled Avg Diff values are recomputed in both the experimental and baseline data sets to include only probe pairs used in both the baseline and experiment arrays. Then the Avg Diff Change is calculated as:

$$Avg\ Diff\ Change = Avg\ Diff_{exp} - Avg\ Diff_{base}$$

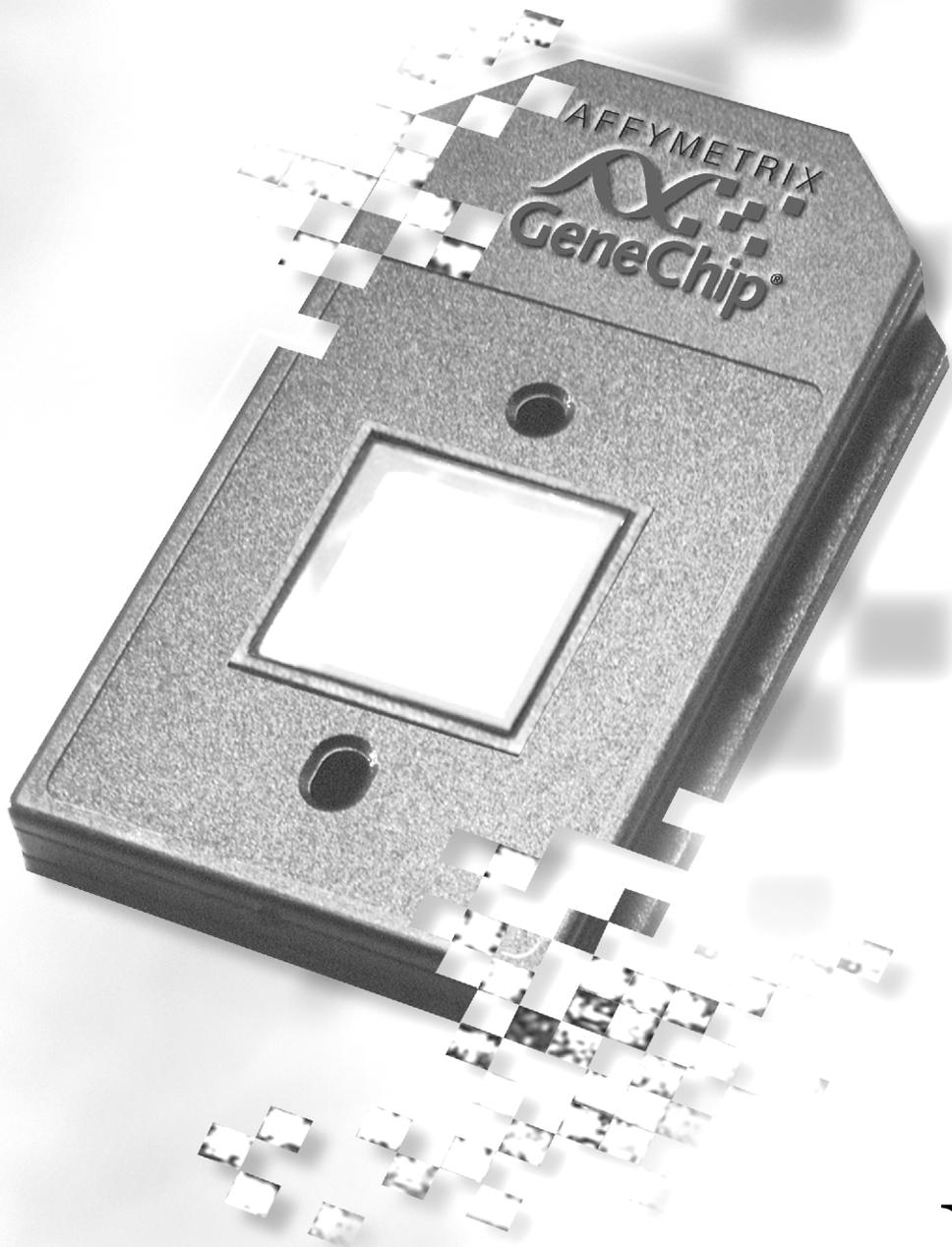
$$\text{and } Q_C = \max(Q_{exp}, Q_{base})$$

Q multiplier: $Q_M = 2.1$ for a $50\ \mu\text{m}$ feature array or $Q_M = 2.8$ for a $24\ \mu\text{m}$ feature array

If the noise (Q) of the experiment or baseline array is greater than the Avg Diff of the transcript (the baseline or experimental data), the Fold Change is calculated over the noise and is an approximation [a tilde character (~) precedes the approximated Fold Change value in the *.chp file].

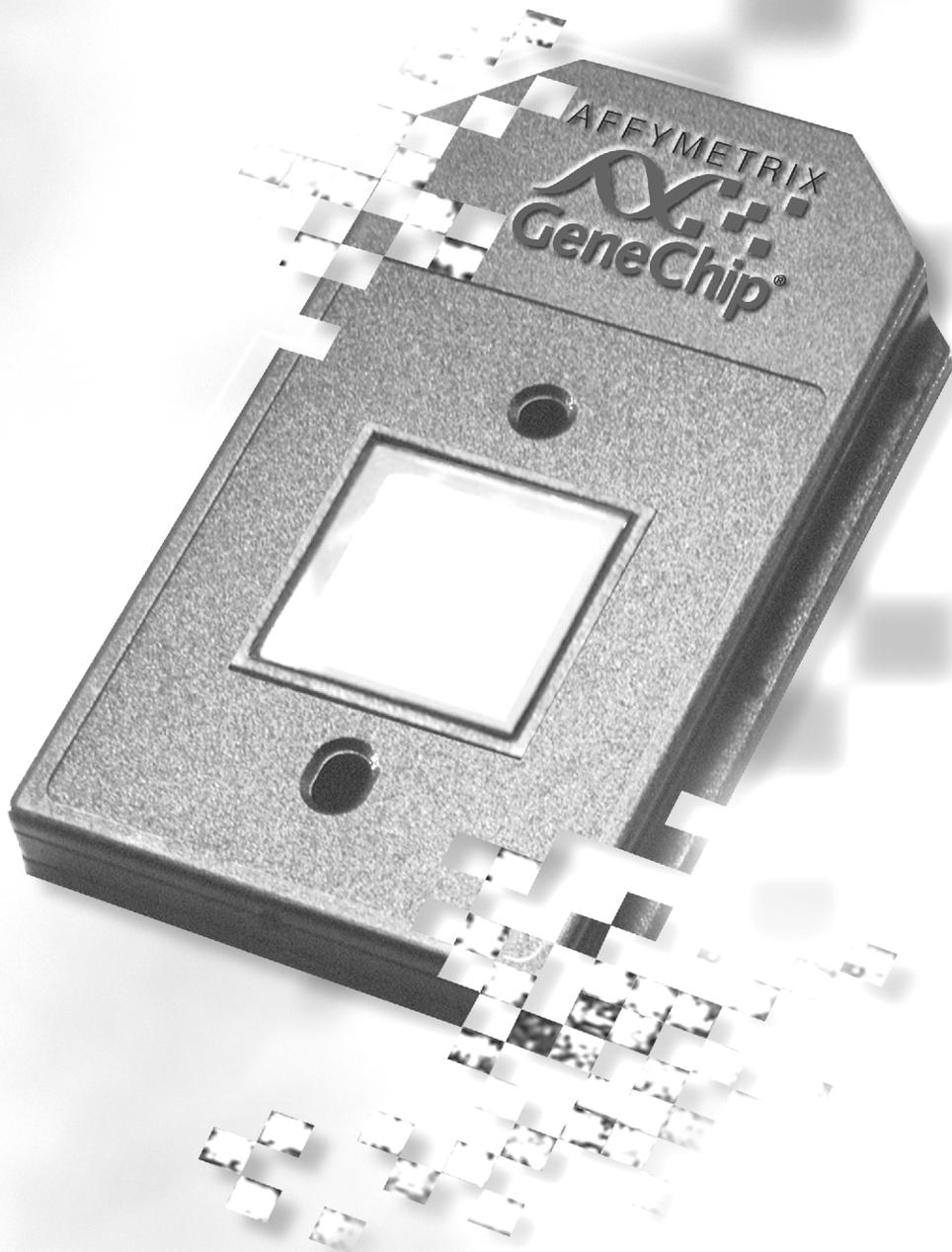
Sort Score

The Sort Score is a ranking based on the Fold Change and the Avg Diff Change. The higher the Fold Change and the Avg Diff Change, the higher the Sort Score.



F

Appendix F





Cross Hybridization, Hybridization, & Spike Probe Masks

A probe mask specifies probe pairs that are to be excluded from an expression analysis. There are three types of probe masks:

- Cross hybridization probe mask specifies probe pairs that include a perfect match (PM) or mismatch (MM) probe cell intensity greater than a user-specified limit.
- Hybridization probe mask includes probe pairs that meet either of the following criteria:

$$\text{PM} - \text{MM} \leq \text{Difference Threshold}$$

or

$$\text{PM}/\text{MM} \leq \text{Ratio Threshold}$$

- Spike probe mask includes probe pairs that meet either of the following criteria:

$$(\text{PM} - \text{MM})_{\text{spike}} - (\text{PM} - \text{MM})_{\text{unspike}} < \text{Difference Threshold},$$

or

$$(\text{PM} - \text{MM})_{\text{spike}} / (\text{PM} - \text{MM})_{\text{unspike}} < 1 + \text{Ratio Threshold}$$

Viewing the Probe Mask Tab

You can create, view, or edit probe masks in the Probe Mask tab of the Expression Analysis Settings dialog box.

To view the probe mask tab:

1. Click **Analysis Settings** in the shortcut bar, then click **Expression** . Alternatively, select **Tools** → **Analysis Settings** → **Expression** from the menu bar.
⇒ The Expression Analysis Settings dialog box appears ([Figure F.1](#)).
2. Click the Probe Mask tab.

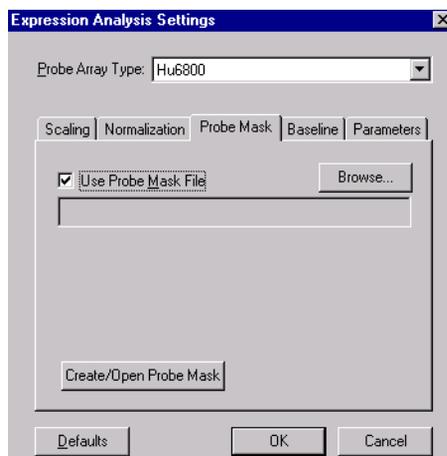


Figure F.1
Expression Analysis Settings, Scaling tab, User Defined scaling option selected

3. Make a selection from the **Probe Array Type** drop-down list.
The settings are specific for the selected probe array type and do not affect the settings for other types of probe arrays.

Cross Hybridization Probe Mask

Microarray Suite analyzes one or more user-specified *.cel files to generate a cross hybridization probe mask. The cross hybridization probe mask specifies probe pairs that have a perfect match (PM) or mismatch (MM) probe cell with an intensity that exceeds a user-specified limit.

A series of *.cel files may be analyzed and the results from each analysis may be combined in an *And* (intersection) or *Or* (union) fashion to create a composite cross hybridization probe mask. The probe pairs specified by the cross hybridization probe mask are excluded or *masked* from an expression analysis.

Creating a Cross Hybridization Probe Mask

1. In the shortcut bar, click **Analysis Settings**, then click **Expression** . Alternatively, select **Tools** → **Analysis Settings** → **Expression** from the menu bar.
⇒ The Expression Analysis Settings dialog box appears (Figure F.2).

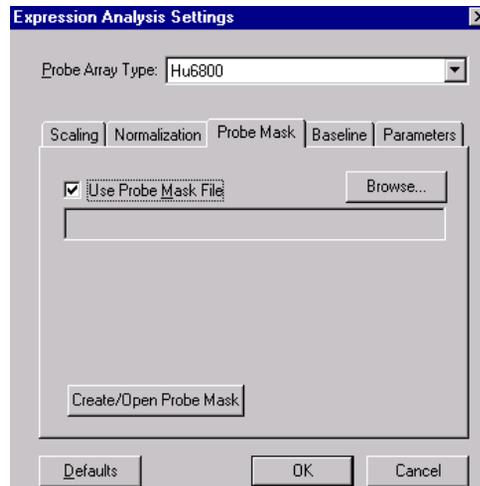


Figure F.2
Expression Analysis Settings, Probe Mask tab

2. Click the Probe Mask tab
3. Click **Create/Open Probe Mask**.
⇒ The Probe Mask File dialog box appears ([Figure F3](#)).

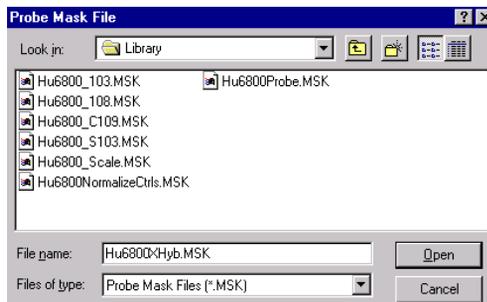


Figure F3
Probe Mask File dialog box

4. Enter a name for the new mask in the **File Name** box and click **Open**.
⇒ The Probe Mask Definition dialog box appears ([Figure F4](#)).

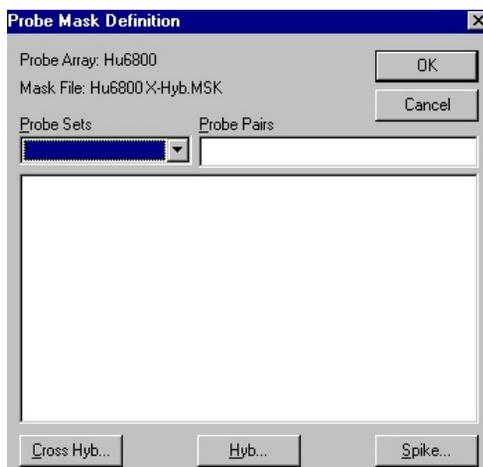


Figure F4
Probe Mask Definition dialog box

5. Click Cross Hyb.

⇒ The Cross Hybridization dialog box appears (Figure F5).

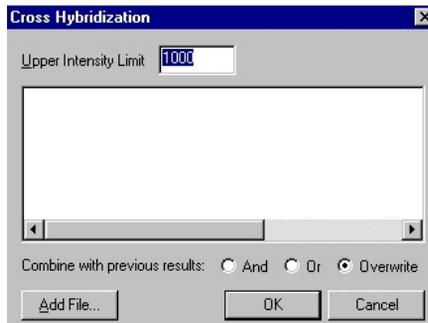


Figure F5
Cross Hybridization dialog box

6. Enter a new upper intensity limit (if necessary).

7. Click Add File.

⇒ This displays the Open dialog box (Figure F6).

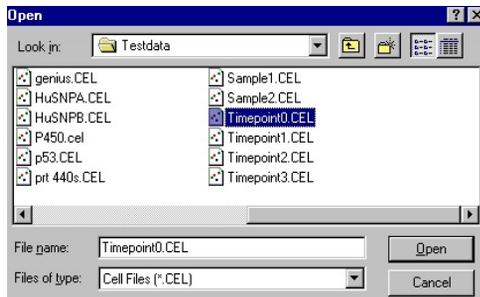


Figure F6
Open dialog box displays *.cel files available for the cross hybridization probe mask

8. Double-click a *.cel file in the Open dialog box.

⇒ This adds the selected *.cel to the Cross Hybridization dialog box (Figure F7).

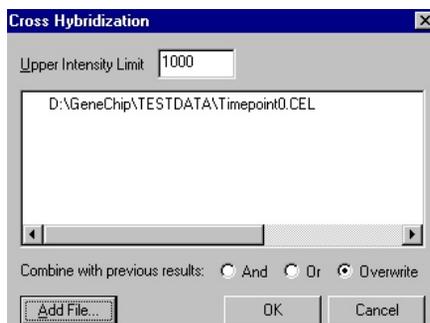


Figure F.7
Cross Hybridization dialog box displays *.cel files selected for the cross hybridization probe mask

9. Repeat and 8 [step 7](#) and [step 8](#) until all desired *.cel files have been added to the Cross Hybridization dialog box.

Microarray Suite identifies the probe pairs in each *.cel file that include a PM or MM probe cell with an intensity greater than the user-specified Upper Intensity Limit (the default is 1000).

The results from each *.cel file may be combined in an *And* (intersection) or *Or* (union) fashion to create the cross hybridization probe mask.

10. Double-click the second (and any subsequent) *.cel file name to toggle between the **And** and **Or** option ([Figure F.8](#)).

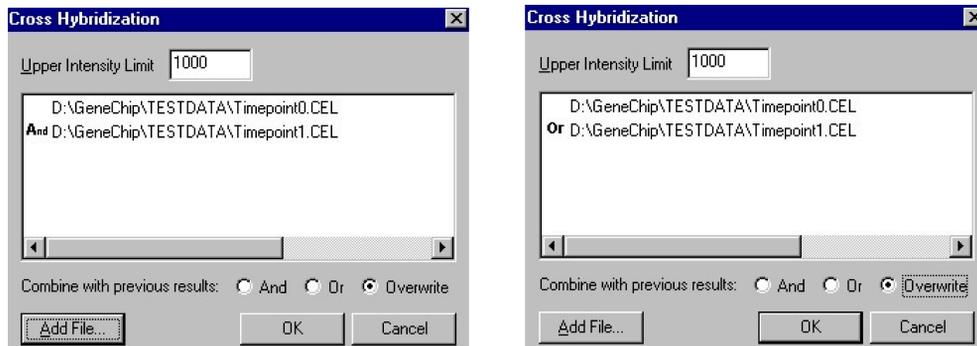


Figure F.8
And combined *.cel files (left), *Or* combined *.cel files (right)

11. Click **OK** to generate the cross hybridization probe mask.
 - ⇒ The Probe Mask Definition dialog box displays a list of the probe pairs in the cross hybridization probe mask (**Figure F.9**).

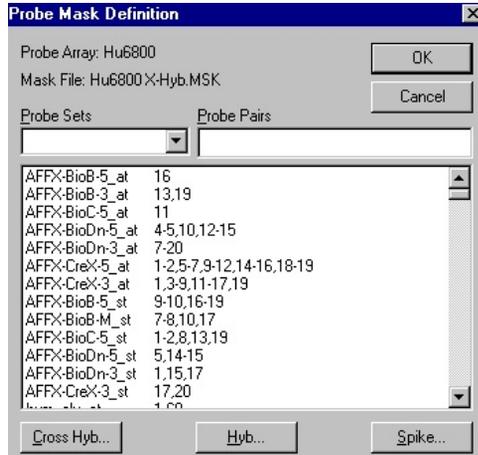


Figure F.9
Probe pairs in the Cross Hybridization Probe mask

12. Click **OK** to close the Probe Mask Definition dialog box.

Editing a Cross Hybridization Probe Mask

1. In the Probe Mask tab, click **Create/Open Probe Mask** (Figure F.10).
⇒ The Probe Mask File dialog box appears (Figure F.11).

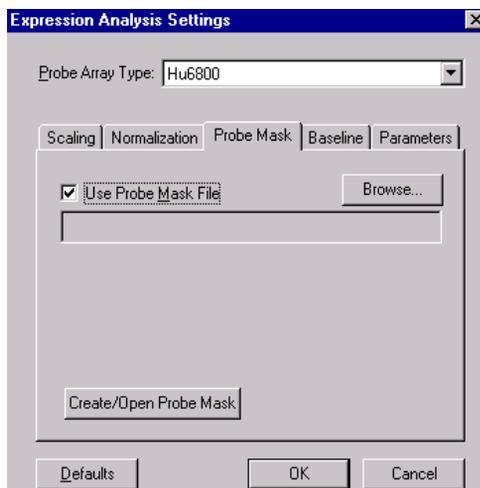


Figure F.10
Expression Analysis Settings dialog box, Probe Mask tab

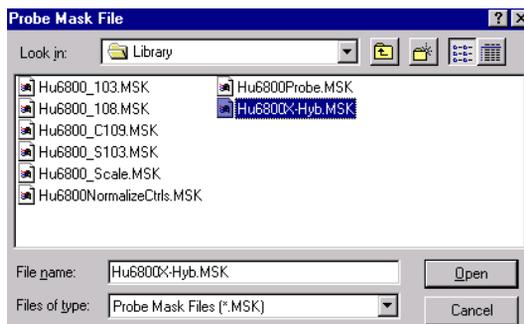


Figure F.11
Probe Mask File dialog box

2. Double-click the cross hybridization probe mask file (*.msk) you wish to edit (Figure F.11).
 - ⇒ The Probe Mask Definition dialog box displays the probe pairs in the currently selected probe mask (Figure F.12).

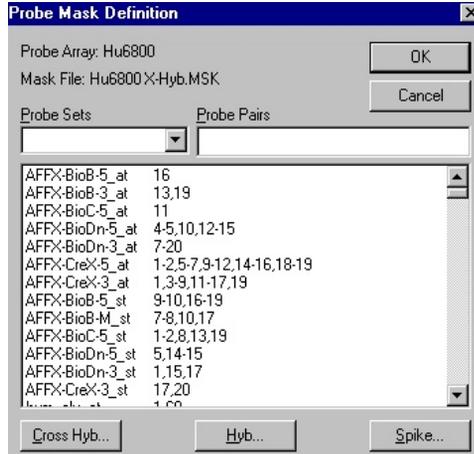
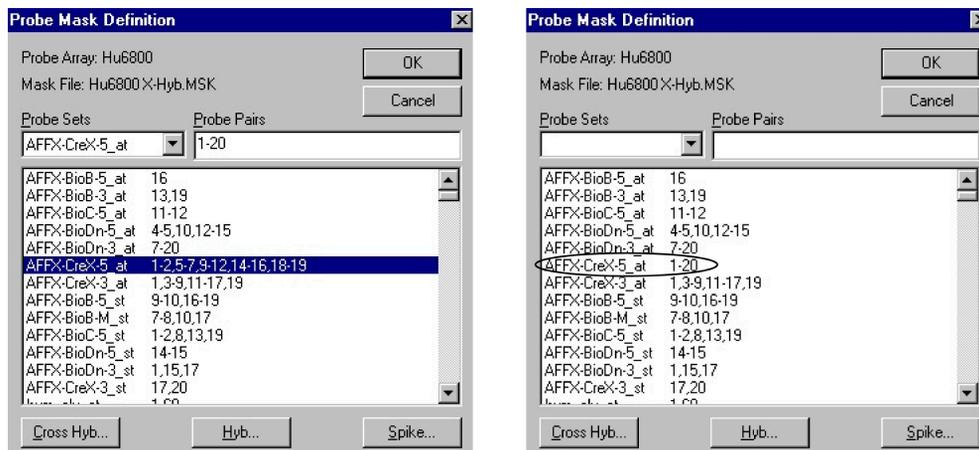


Figure F.12
Probe Mask Definition dialog box

3. In the lower box, click the Probe Set/Probe Pair entry you wish to edit.
 - ⇒ The **Probe Sets** box and **Probe Pairs** box automatically display the selection (Figure F.13).

**Figure F.13**

Edit the Probe Pairs box and then click the lower box (left) to update the probe pair list (right)

4. Edit or delete the entry in the **Probe Pairs** field.
5. Click the lower field to update the list of probe pairs (**Figure F.13**).
6. Click **OK** when finished to close the Probe Mask Definition dialog box.

Combining Cross Hybridization Probe Masks

A cross hybridization mask may be combined with (*And* or *Or* fashion) or overwrite an existing cross hybridization mask.

1. In the Probe Mask tab, click **Create/Open Probe Mask** (**Figure F.14**).
⇒ The Probe Mask File dialog box appears (**Figure F.15**).

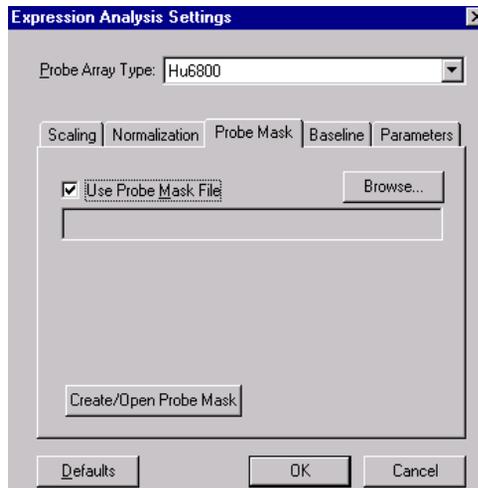


Figure F.14
Expression Analysis Settings dialog box, Probe Mask tab

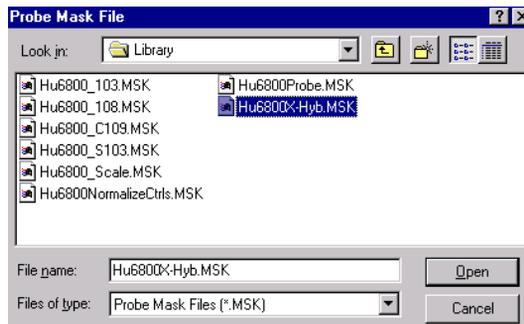


Figure F.15
Probe Mask File dialog box

2. In the Probe Mask File dialog box, double-click the cross hybridization probe mask you wish to combine with another cross hybridization probe mask(s) (Figure F.15).
⇒ The Probe Mask Definition dialog box appears (Figure F.16).

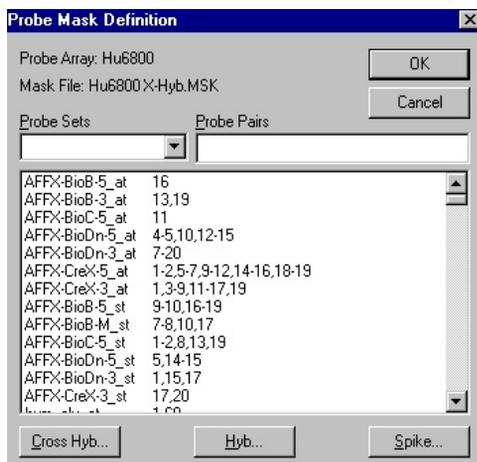


Figure F.16
Probe Mask Definition dialog box

3. Click **Cross Hyb.**

⇒ The Cross Hybridization dialog box appears (**Figure F.17**).

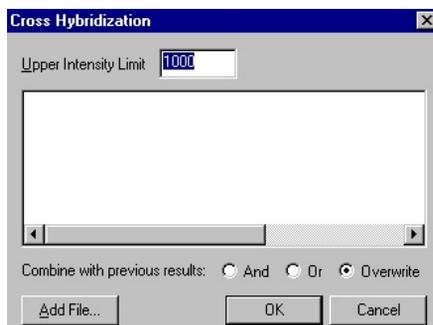


Figure F.17
Cross Hybridization dialog box

4. Enter a new **Upper Intensity Limit** if desired.

5. Click **Add File**.

⇒ This displays the Open dialog box (**Figure F.18**).

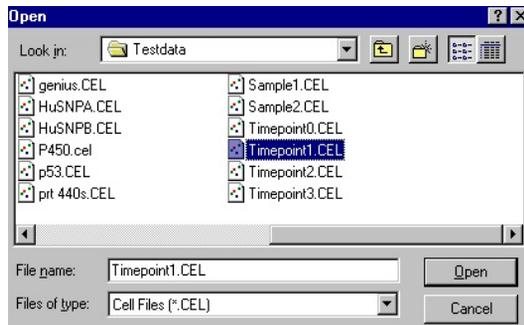


Figure F.18
Open dialog box

6. Double-click the *.cel file you wish to combine with the cross hybridization mask selected in [step 2](#).
⇒ The Cross Hybridization dialog box displays the selected *.cel file ([Figure F.19](#)).

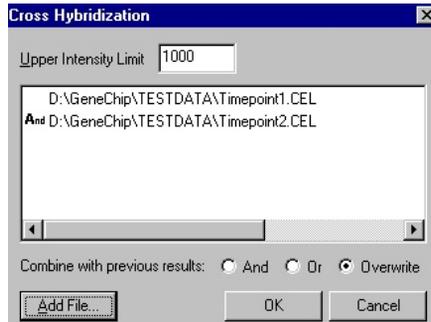


Figure F.19
Cross Hybridization dialog box

7. Click **OK** and repeat [step 5](#) and [step 6](#) to add additional *.cel files to the Cross Hybridization dialog box.
8. Double-click the second (and any subsequent) *.cel file name to toggle between *And* and *Or* ([Figure F.19](#)).

9. Choose a **Combine with previous results** option (And, Or, or Overwrite), then click **OK**.

Microarray Suite analyzes the selected *.cel files and identifies all probe pairs with a PM or MM probe cell that has an intensity greater than the user-specified **Upper Intensity Limit**. It combines the results from the *.cel files (And or Or or Overwrite) to create a composite cross hybridization probe mask.

This composite probe mask is combined with the probe mask selected in [step 2](#) (for example, Hu6800X-hyb.MSK in [Figure F.15](#)) as specified in [step 9](#) (*And* or *Or* fashion, or overwrites the selected cross hybridization probe mask).

Hybridization Probe Mask

Microarray Suite analyzes one or more user-specified cell intensity files (*.cel) to generate a hybridization probe mask comprised of probe pairs that meet either of the following criteria:

$PM - MM \leq \text{Difference Threshold}$

or

$PM/MM \leq \text{Ratio Threshold}$

The Difference and Ratio Thresholds are user-modifiable with defaults of 30 and 1.5 respectively. (See Appendix [Appendix E](#) for more information about the Difference Threshold and Ratio Threshold.)

You can analyze a series of *.cel files and combine the results from each analysis in an *And* (intersection) or *Or* (union) fashion to create a composite hybridization probe mask.

Creating a Hybridization Probe Mask

1. In the Probe Mask tab, click **Create/Open Probe Mask** (Figure F.20).
⇒ The Probe Mask File dialog box appears (Figure F.21).

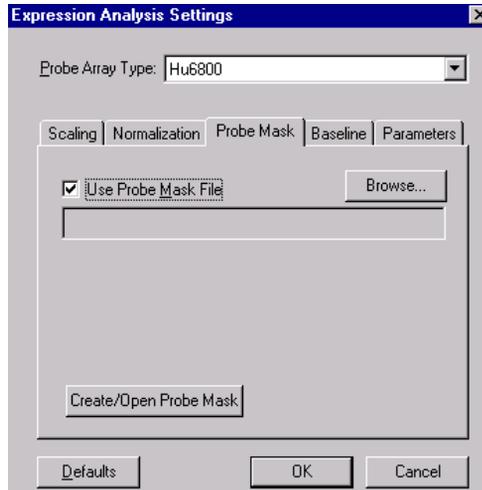


Figure F.20
Expression Analysis Settings dialog box, Probe Mask tab

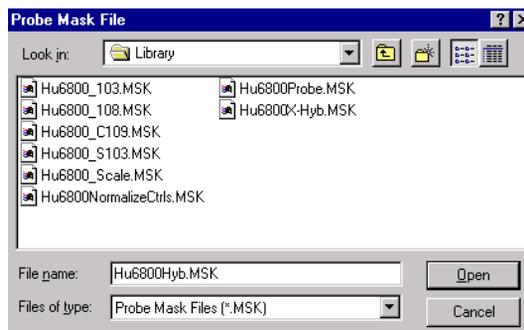


Figure F.21
Probe Mask File dialog box

2. Enter a name for the new hybridization probe mask in the **File Name** box, then click **Open**.

⇒ The Probe Mask Definition dialog box appears (Figure F.22).

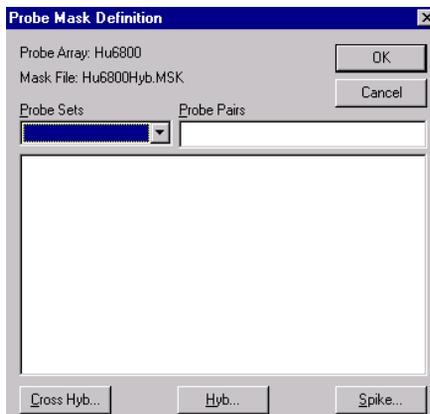


Figure F.22
Probe Mask Definition dialog box

3. Click **Hyb**.

⇒ The Hybridization Behavior dialog box appears (Figure F.23).

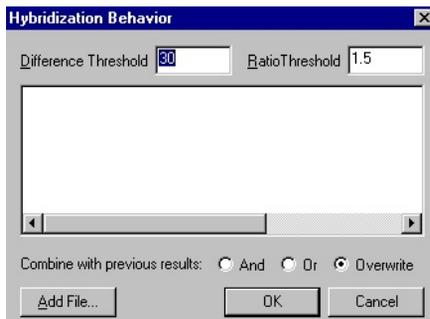


Figure F.23
Hybridization Behavior dialog box

4. Click Add File.

⇒ This displays the Open dialog box (Figure F.24).

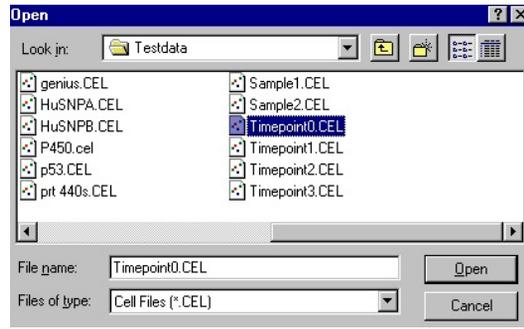


Figure F.24
Open dialog box

5. Double-click the *.cel file of interest.

⇒ This adds the selected *.cel file to the Hybridization Behavior dialog box (Figure F.25).

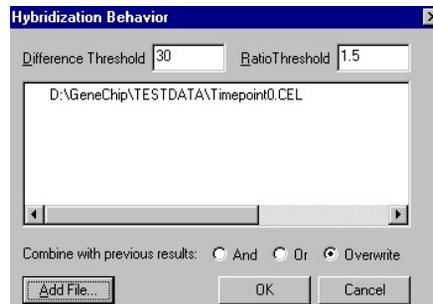


Figure F.25
Hybridization Behavior dialog box

- 6. If desired, click Add File again to select and add another *.cel file to the Hybridization Behavior dialog box. Repeat step 4 and step 5 until all desired *.cel files have been added.**

You can combine the results from the analysis of each *.cel file in an *And* (intersection) or *Or* (union) fashion to create the hybridization probe mask.

7. Double-click the second (and each subsequent) *.cel file name to toggle between the *And* and *Or* option (Figure F.26).

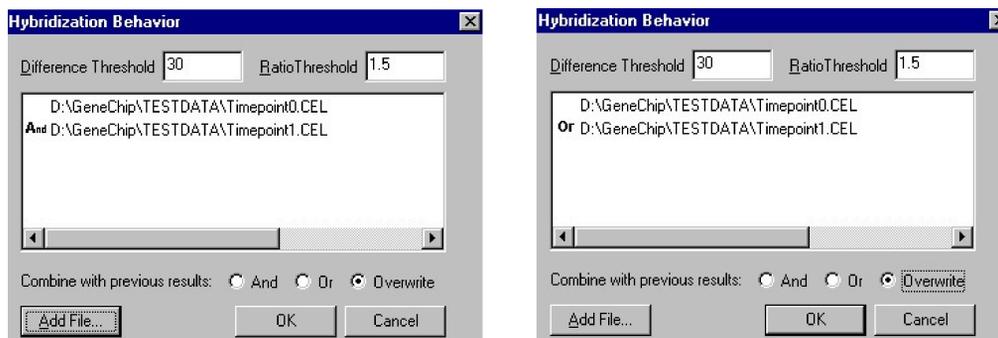


Figure F.26

And combined *.cel files analysis results (left), *Or* combined *.cel file analysis results (right)

8. Click **OK** to generate the hybridization probe mask.
⇒ The Probe Mask Definition dialog box displays the probe pairs of the Hybridization probe mask (Figure F.27).

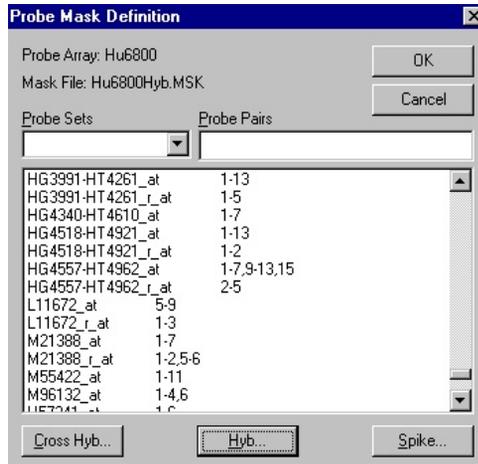


Figure F27
Probe pairs in the hybridization probe mask

Editing a Hybridization Probe Mask

1. In the Probe Mask tab, click **Create/Open Probe Mask** (Figure F28).
⇒ The Probe Mask File dialog box appears (Figure F29).

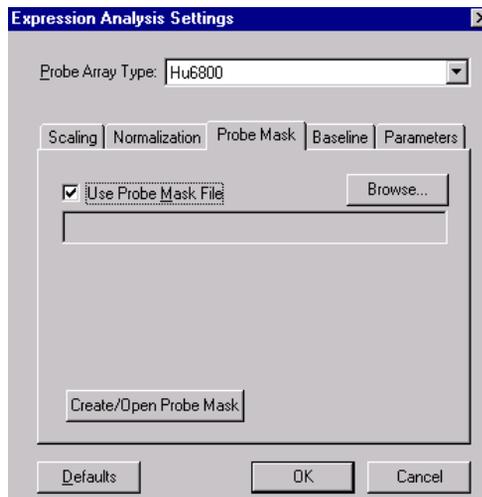


Figure F28
Expression Analysis Settings dialog box, Probe Mask tab



Figure F29
Probe Mask File dialog box

2. Double-click the hybridization probe mask (*.msk) of interest (Figure F29).
⇒ The Probe Mask Definition dialog box displays the probe pairs in the selected hybridization probe mask (Figure F30).

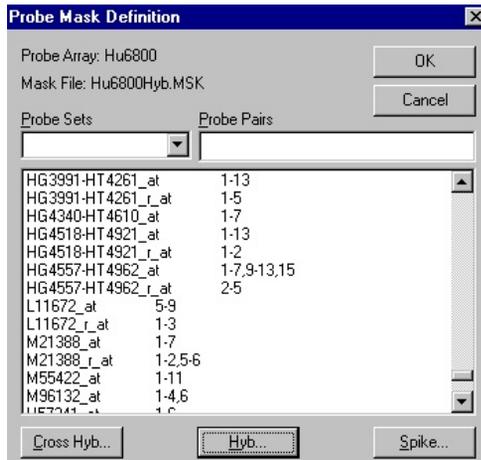


Figure F30
Probe Mask Definition dialog box

3. In the lower box, click the desired Probe Set/Probe Pair entry you wish to edit.
⇒ The **Probe Sets** box and **Probe Pairs** box automatically display the selection (Figure F31).

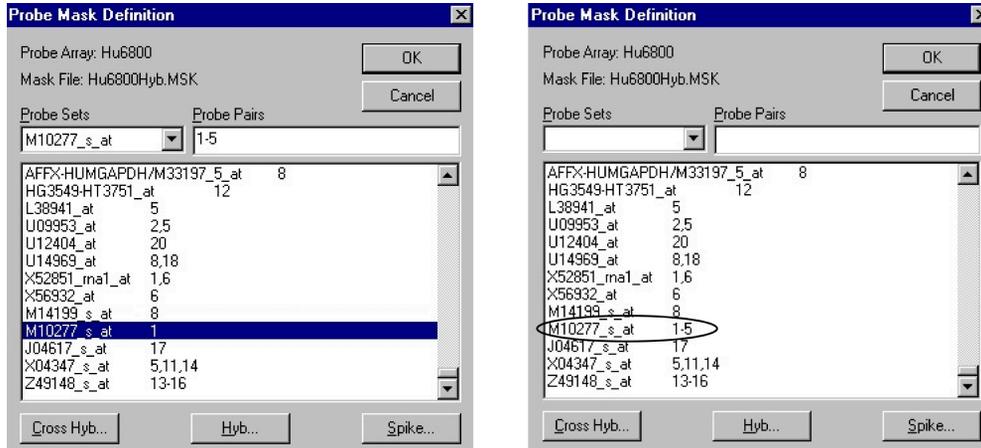


Figure F31

Edit the entry in the Probe Pairs box (left) and click the lower box to update the probe pair list (right)

4. Edit or delete the entry in the **Probe Pairs** box, then click the lower box to update the list of probe pairs (Figure F31).
5. Click **OK** when finished to close the Probe Mask Definition dialog box.

Combining Hybridization Probe Masks

A hybridization probe mask may be combined with (*And* or *Or* fashion) or overwrite an existing hybridization probe mask.

1. In the Probe Mask tab, click **Create/Open Probe Mask** (Figure F32).
⇒ The Probe Mask File dialog box appears (Figure F33).

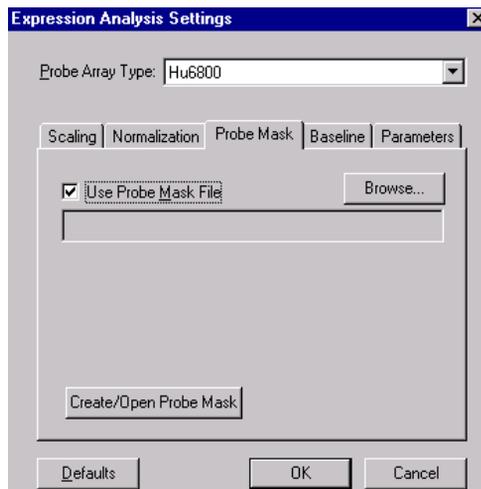


Figure F32
Expression Analysis Settings dialog box, Probe Mask tab

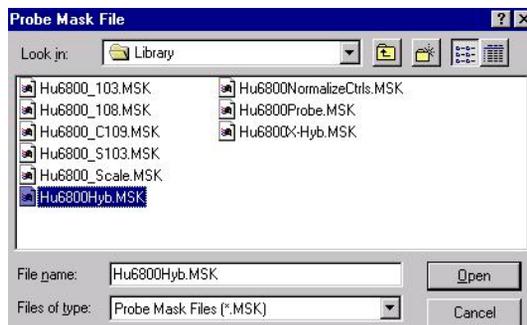


Figure F33
Probe Mask File dialog box

2. Double-click the hybridization probe mask (*.msk) of interest.
⇒ The Probe Mask Definition dialog box appears (Figure F34).

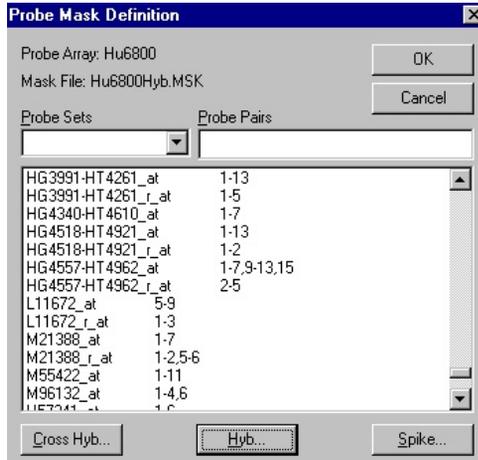


Figure F34
Probe Mask Definition dialog box

3. Click **Hyb.**
⇒ The Hybridization Behavior dialog box appears (Figure F35).

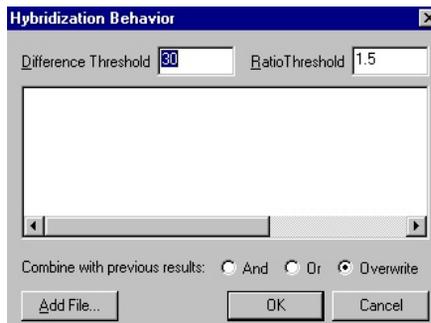


Figure F35
Hybridization Behavior dialog box

4. Enter new Difference Threshold or Ratio Threshold values if desired.
5. Click **Add File** in the Hybridization Behavior dialog box.
⇒ This displays the Open dialog box (Figure F.36).

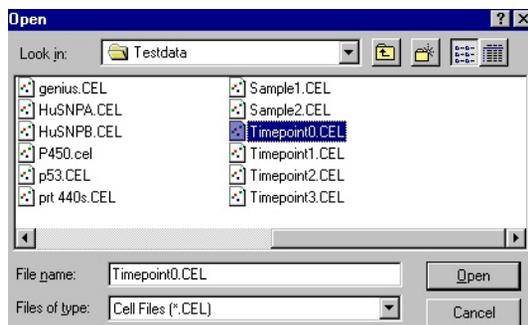


Figure F.36
Open dialog box

6. Double-click the *.cel file you wish to combine with the hybridization mask selected in step 2.
⇒ The Hybridization Behavior dialog box displays the selected *.cel file (Figure F.37).
7. Click **OK**. Repeat step 5 and step 6 to add additional *.cel files to the Hybridization Behavior dialog box (Figure F.37).

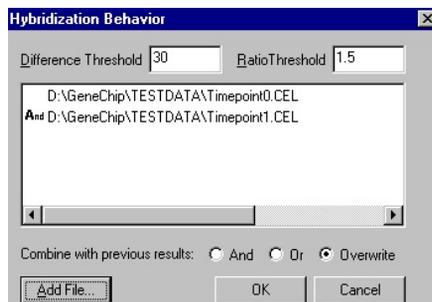


Figure F.37
Hybridization Behavior dialog box

8. Double-click the second (and any subsequent) *.cel file name to combine the *.cel files in an *And* and *Or* fashion.
9. Choose a **Combine with previous results** option (*And*, *Or*, or *Overwrite*), then click **OK**.

The Microarray Suite software analyzes the selected *.cel files to identify the probe pairs where $PM - MM \leq \text{Difference Threshold}$ or $PM / MM \leq \text{Ratio Threshold}$. The software combines the results from each *.cel file in an *And* or *Or* fashion to create a new composite Hybridization probe mask.

This combines the composite hybridization mask with the mask selected in [step 2](#) (for example, Hu6800Hyb.msk in [Figure F33](#)) as specified in [step 9](#) (*And* or *Or* fashion, or overwrites the selected hybridization mask).

Spike Probe Mask

Microarray Suite compares two *.cel files, one derived from a *spiked* target that contained known amounts of a control transcript(s) and one derived from an unspiked target.

The software identifies any probe pairs where:

$$(PM - MM)_{\text{spike}} - (PM - MM)_{\text{unspike}} < \text{Difference Threshold},$$

or

$$(PM - MM)_{\text{spike}} / (PM - MM)_{\text{unspike}} < 1 + \text{Ratio Threshold}$$

and automatically generates a spike probe mask (*.msk) comprised of all probe pairs that meet either of the above criteria. The Difference and Ratio Thresholds are user modifiable with defaults of 30 and 1 respectively.

Creating a Spike Probe Mask

1. In the Probe Mask tab, click **Create/Open Probe Mask** (Figure F.38).
⇒ The Probe Mask File dialog box appears (Figure F.39).

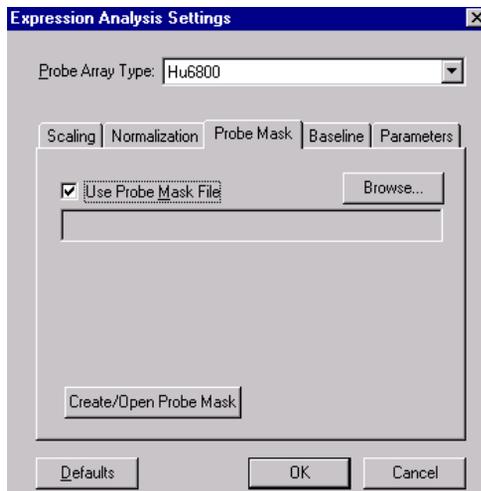


Figure F.38
Expression Analysis Settings dialog box, Probe Mask tab

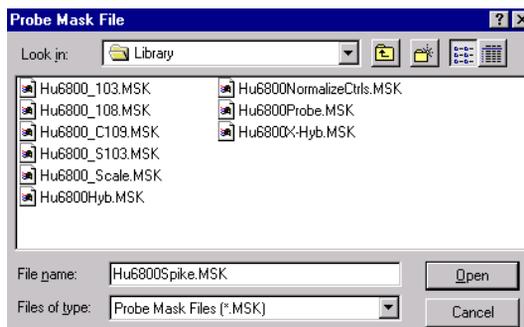


Figure F.39
Probe Mask File dialog box

2. Enter a name in the **File Name** box for the new spike probe mask, then click **Open** (Figure F.39).
⇒ The Probe Mask Definition dialog box appears (Figure F.40).

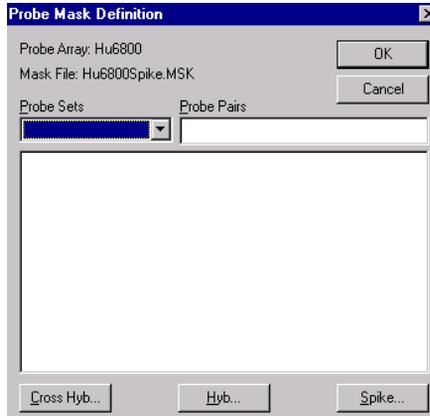


Figure F.40
Probe Mask Definition dialog box

3. Click **Spike**.
⇒ The Spike Behavior dialog box appears (Figure F.41).

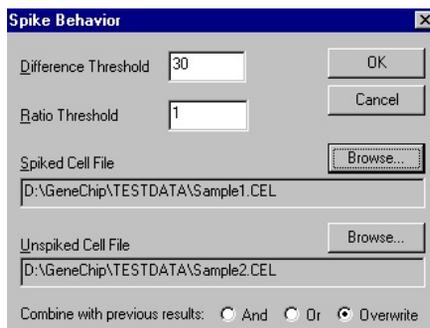


Figure F.41
Spike Behavior dialog box

4. To select a spiked *.cel file, click **Browse** (upper button).
⇒ This displays the Open dialog box (Figure F.42).

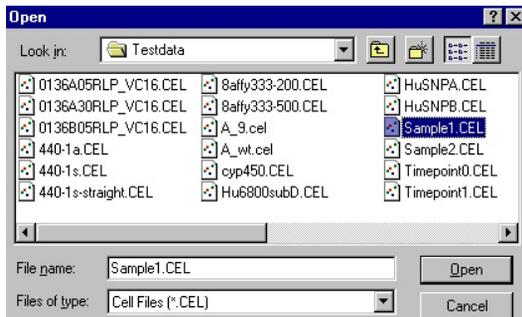


Figure F.42
Open dialog box

5. Double-click the *.cel file derived from the spiked target.
⇒ This adds the selected *.cel to the **Spiked Cell File** box (Figure F.41).
6. To select an unspiked *.cel, click **Browse** (lower button) in the Spike Behavior dialog box (Figure F.41).
⇒ This displays the Open dialog box (Figure F.42).
7. Double-click the *.cel file derived from the unspiked target.
⇒ This adds the selected *.cel to the **Unspiked Cell File** box (Figure F.41).
8. Click OK to generate the spike probe mask.
⇒ The Probe Mask Definition dialog box displays the probe pairs of the spike probe mask (Figure F.43).

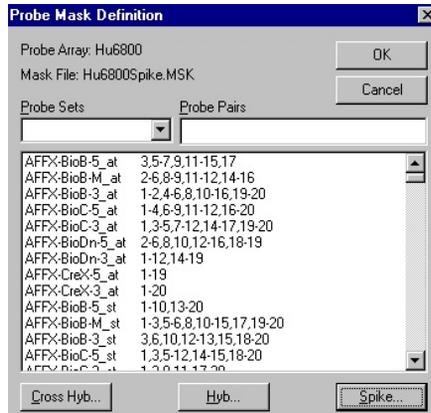


Figure F43
Probe Mask Definition dialog box

Editing a Spike Probe Mask

1. In the Probe Mask tab, click **Create/Open Probe Mask** (Figure F44).
⇒ The Probe Mask File dialog box appears (Figure F45).

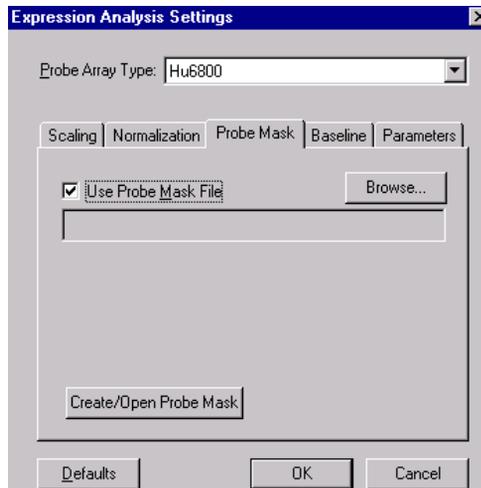


Figure F44
Expression Analysis Settings, Probe Mask tab

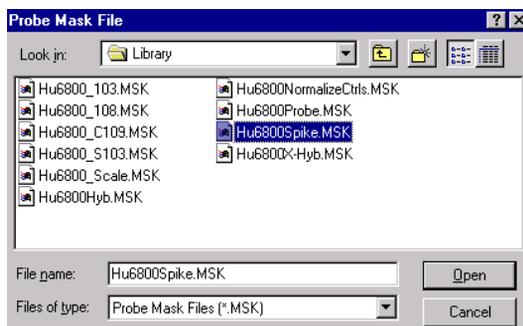


Figure F.45
Probe Mask File dialog box

2. Double-click the spike probe mask you wish to edit (Figure F.45).
⇒ The Probe Mask Definition dialog box displays the probe pairs in the selected spike probe mask (Figure F.46).

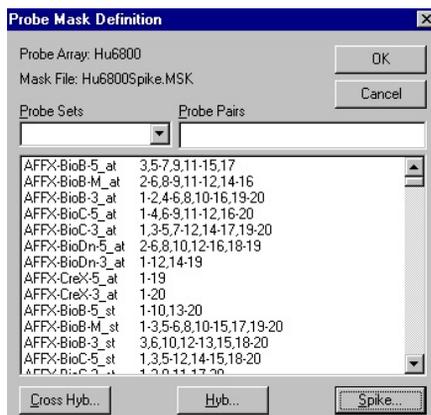


Figure F.46
Probe Mask Definition dialog box

3. In the lower box, click the probe set/probe pair entry you wish to edit.
⇒ The Probe Sets box and Probe Pairs box automatically display the selection (Figure F.47).

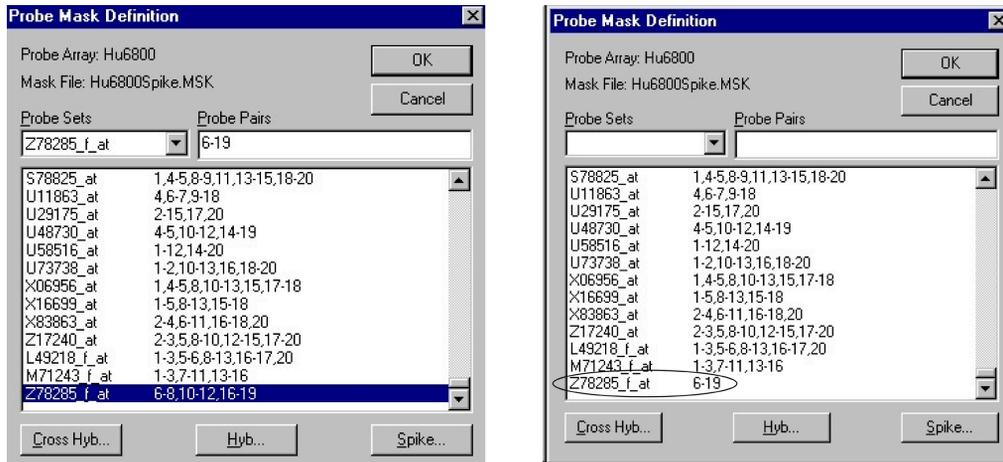


Figure F.47

Edit the Probe Pairs box and click the lower box (left) to update the probe pair list (right)

4. Edit or delete the entry in the **Probe Pairs** box, then click the lower field to update the list of probe pairs (Figure F.47).
5. Click **OK** when finished to close the Probe Mask Definition dialog box.

Combining Spike Probe Masks

A spike probe mask may be combined with (*And* or *Or* fashion) or overwrite an existing spike probe mask.

1. In the Probe Mask tab, click **Create/Open Probe Mask** (Figure F.48).
⇒ The Probe Mask File dialog box appears (Figure F.49).

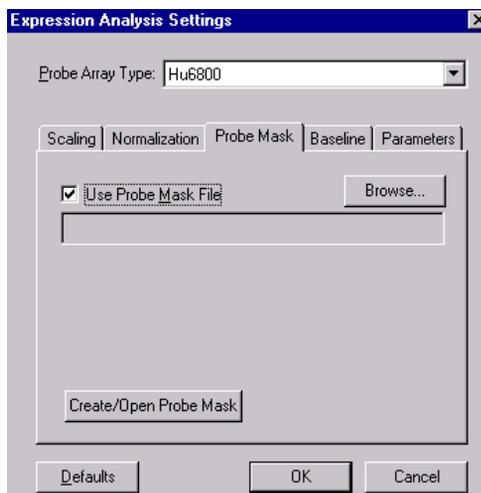


Figure F.48
Expression Analysis Settings, Probe Mask tab

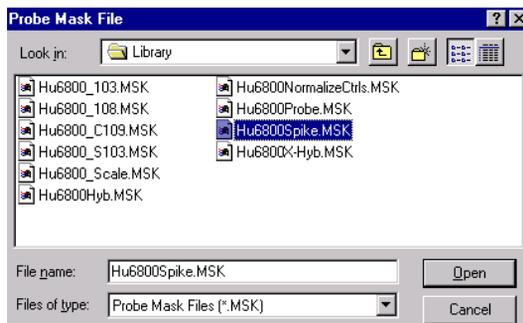


Figure F.49
Probe Mask File dialog box

2. Double-click the spike probe mask of interest (Figure F.49).
⇒ The Probe Mask Definition dialog box displays the probe pairs in the selected spike probe mask (Figure F.50).

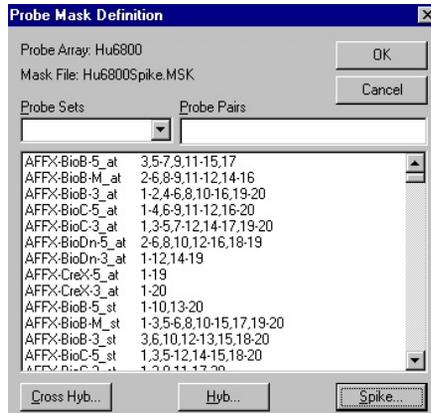


Figure F50
Probe Mask Definition dialog box

3. Click Spike.

⇒ The Spike Behavior dialog box appears (**Figure F51**).

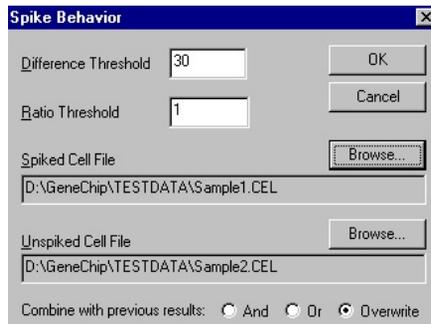


Figure F51
Spike Behavior dialog box

4. To select a spiked *.cel file, click **Browse (upper button).**

⇒ This displays the Open dialog box (**Figure F52**).

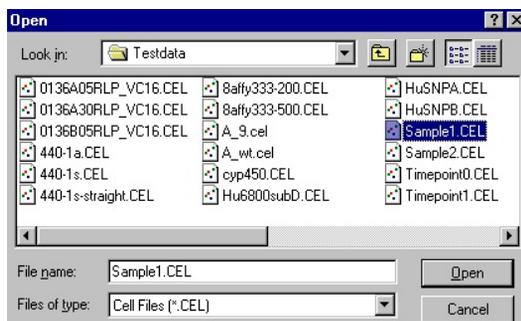


Figure F.52
Open dialog box

5. Double-click the *.cel file derived from the spiked target.
⇒ This adds the selected *.cel to the **Spiked Cell File** box (Figure F.51).
6. To select an unspiked *.cel, click **Browse** (lower button) in the Spike Behavior dialog box (Figure F.51).
⇒ This displays the Open dialog box (Figure F.52).
7. Double-click the *.cel file derived from the unspiked target.
⇒ This adds the selected *.cel to the **Unspiked Cell File** box (Figure F.51).
8. Select a **Combine with previous results** option (*And*, *Or*, or *Overwrite*) and click **OK**.

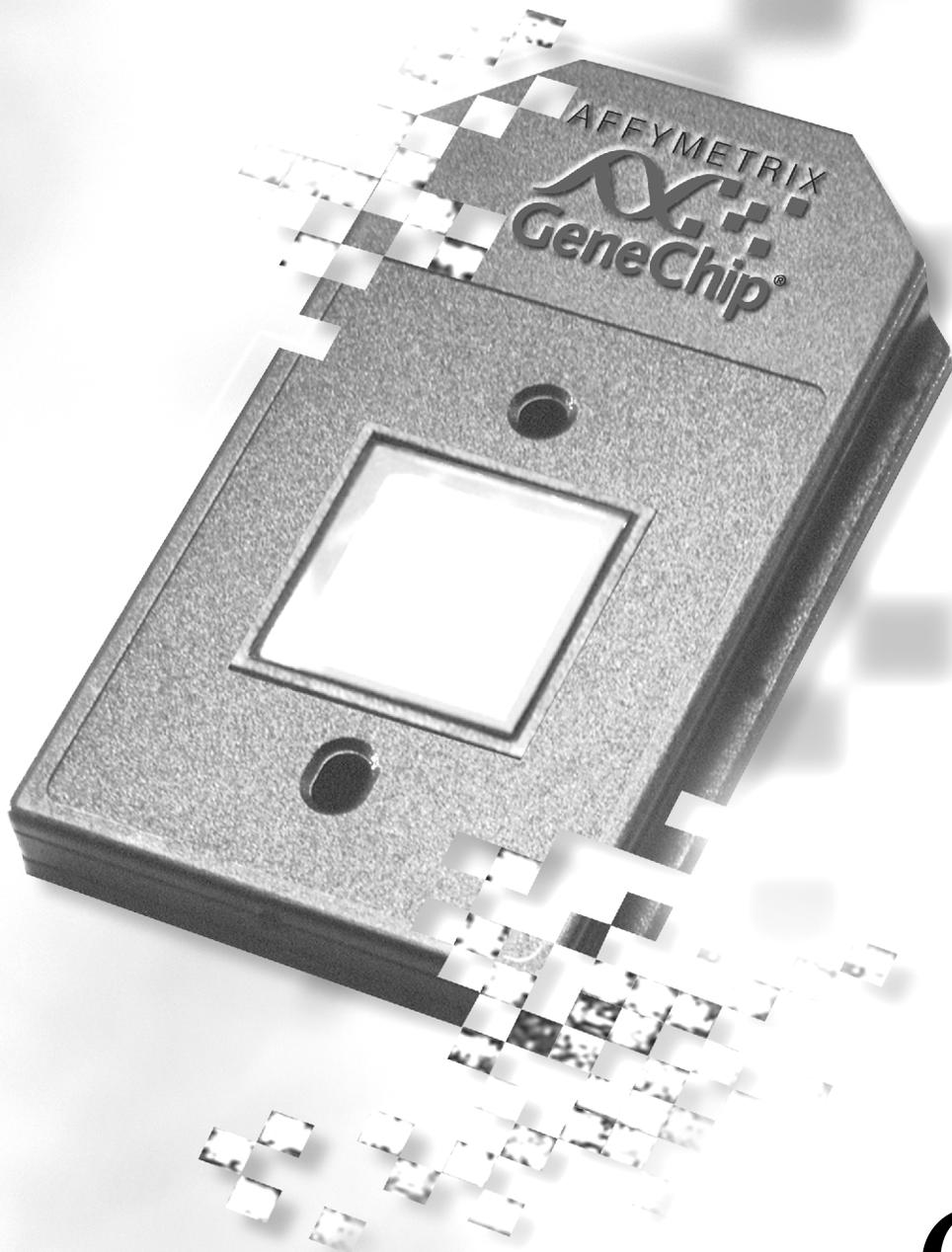
The Microarray Suite software analyzes the two selected *.cel files and identifies all probe pairs that meet either of the following criteria:

$$(PM - MM)_{\text{spike}} - (PM - MM)_{\text{unspike}} < \text{Difference Threshold}$$

or

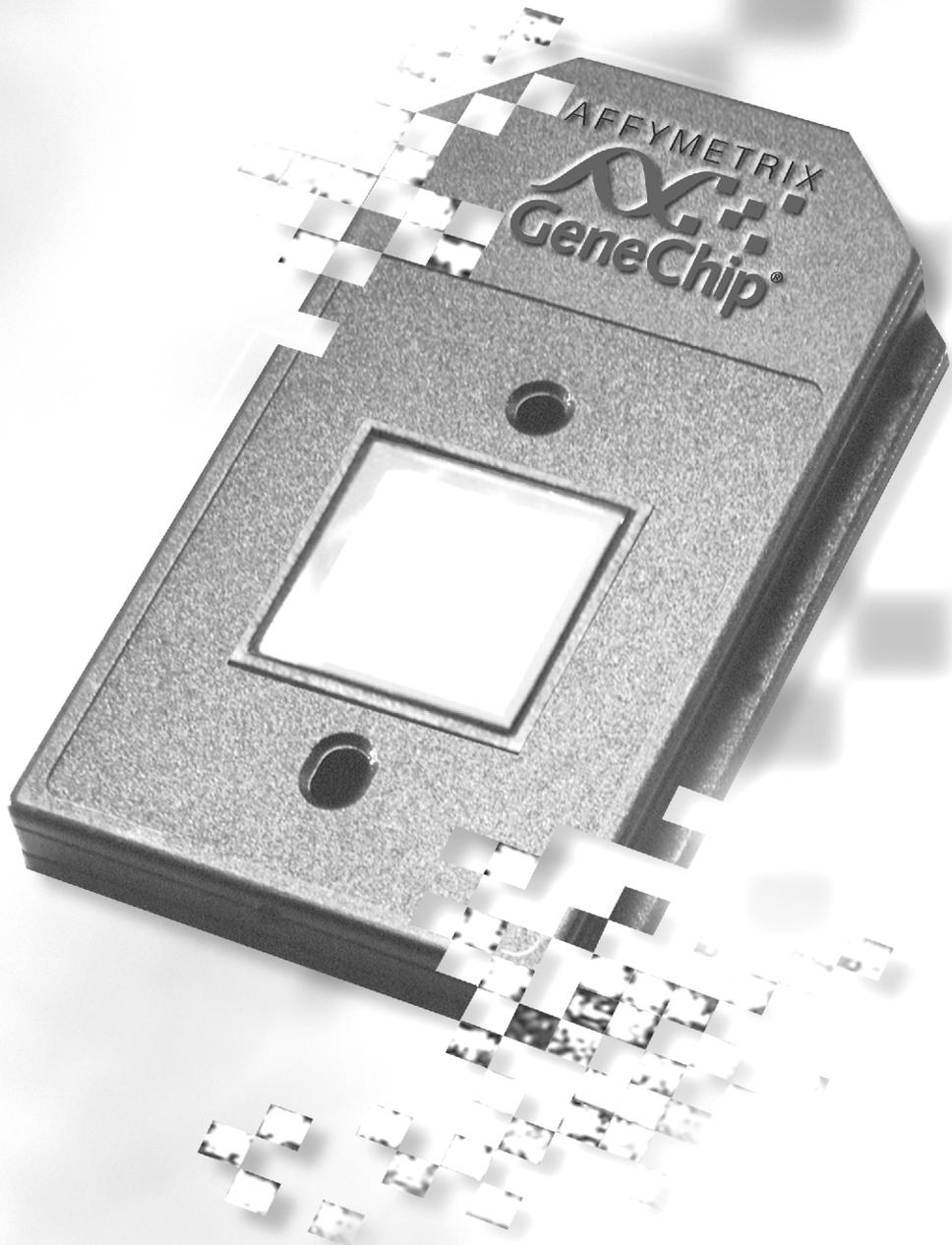
$$(PM - MM)_{\text{spike}} / (PM - MM)_{\text{unspike}} < 1 + \text{Ratio Threshold}$$

The software combines the results from the *.cel files to create a new composite spike probe mask. This combines the composite spike probe mask with the spike probe mask selected in step 2 (for example, Hu6800Spike.MSK in Figure F.45) as specified in step 8 (*And* or *Or* fashion, or overwrites the selected spike probe mask).



G

Appendix G



Mixture Detection Analysis Settings

Some input parameters for the Mixture Detection algorithm are user modifiable. The default values were empirically optimized through extensive testing at Affymetrix and should only be modified by expert users.

The Mixture Detection algorithm relies on the analysis settings to derive biologically meaningful results from GeneChip® p53 probe array hybridization intensity data.

✓ NOTE

The Genotyping algorithm is only available when Microarray Suite is run in disk files mode.

Viewing & Editing Analysis Settings

1. In the shortcut bar, click **Analysis Settings**, then click **Mixture Detection** . Alternatively, select **Tools** → **Analysis Settings** → **Mixture Detection** from the menu bar.
⇒ The Mixture Detection Analysis Settings dialog box appears ([Figure G.1](#)).

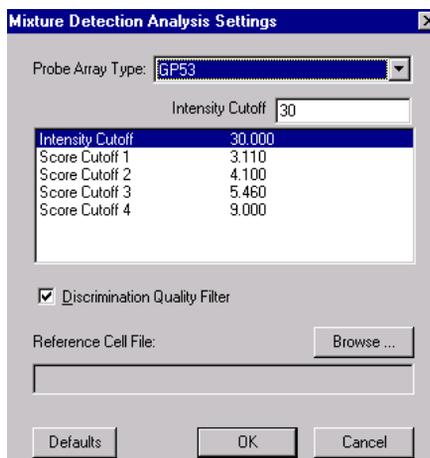


Figure G.1
Mixture Detection Analysis Settings dialog box

2. Select **GP53** from the **Probe Array Type** drop-down list.
The settings are specific for the selected probe array type and do not affect the settings for other types of probe arrays.
3. To edit a parameter value (**Figure G.2**):
 - a. Highlight the parameter in the lower box.
 - b. Enter the new value in the upper box.
 - c. Click the lower box to display the new parameter value.

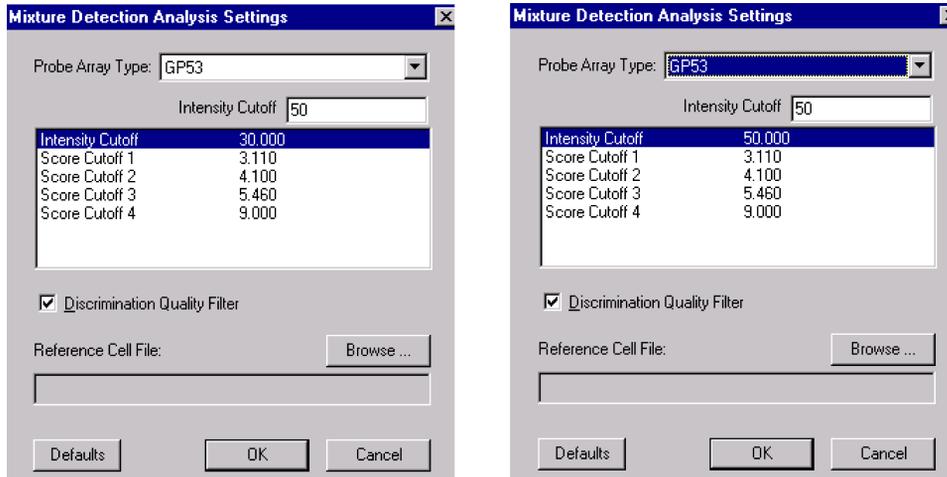


Figure G.2 Mixture Detection Analysis Settings; enter a new parameter value in the upper box and click the lower field (left) to display the new parameter value (right)

Intensity Cutoff

If all of the background-subtracted intensities of the probe cells in a probe set are less than the Intensity Cutoff, the probe set is removed from the analysis.

Score Cutoff

The Score Cutoffs should only be modified by expert users. Increasing the Score Cutoff increases the stringency of the analysis.

The Mixture Detection algorithm computes a score for each probe set used in the analysis. The score value increases with increasing fractions of mutant base in the target.

Microarray Suite calls a putative mutant for the position if the probe sets that interrogate a position pass the quality control filters and the sum of their scores is greater than an empirically determined Score Cutoff.

The Score Cutoff is a function of the number of probe sets that interrogate a position (standard tiles only or standard and alternate tiles). For example, if

two probe sets analyze a position (the sense and anti-sense standard tiles) and pass the algorithm quality control filters, then the sum of the scores for the two probe sets must be greater than Score Cutoff 2 to call a putative mutant at the position ([Table G.1](#)).

Table G.1
Score cutoffs

Number of Probe Sets Used to Analyze a Position	Score Cutoff
1	Score Cutoff 1
2	Score Cutoff 2
3-10	Score Cutoff 3
>10	Score Cutoff 4

The Score Cutoffs are user-modifiable. The default values were determined through extensive empirical testing at Affymetrix ([Figure G.1](#)).

Discrimination Quality Filter

The Discrimination Quality Filter measures the difference in the cell intensities of the wildtype and the three non-wildtype probe cells in a probe set. It is the ratio of the background-subtracted intensities of the *wildtype probe/average intensity of the non-wildtype probes* for each probe set.

The average value for all of the positions in an exon is computed. If this value is less than a threshold value, all positions in the exon are called ambiguous (n).

$$\text{Discrimination Quality Filter} = \left(\sum_{i=1}^{i=\text{NumberPS}} \text{IntDiscPS}_i \right) / \text{NumberPS}$$

NumberPS = The number of standard and alternate probe sets for a given exon minus outlier probe sets.

Outlier probe set = a probe set where the *mean intensity – background* < 0 for the A, C, G, or T-substituted cells in the probe set. The intensity of each probe cell in the probe set must be greater than the background in order to include the probe set in the computation of the Discrimination Quality Filter.

$\text{IntDiscPS}_i =$

$$\text{MeanIntensityWTProbe} - \text{Background} / \left(\frac{1}{3} \sum_{i=1}^{i=3\text{substitutionprobes}} (\text{MeanIntensitynonWTProbe}_i - \text{Background}) \right)$$

Reference Cell File

The Mixture Detection algorithm compares the hybridization intensity data of the target and a p53 wildtype reference DNA sample (see the GeneChip® p53 probe array package insert). Prior to an analysis, you must specify a *.cel file for the reference wildtype p53 DNA sample in the Mixture Detection Analysis Settings dialog box, otherwise consensus base calls are not generated.

Selecting a Reference Cell Intensity File

1. In the Mixture Detection Analysis Settings dialog box ([Figure G.3](#)), click **Browse**.
⇒ This displays the Open dialog box ([Figure G.4](#)).

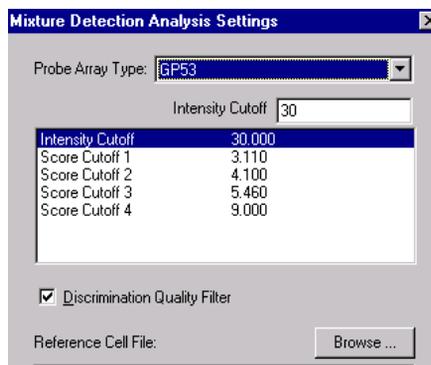


Figure G.3
Mixture Detection Analysis Settings dialog box

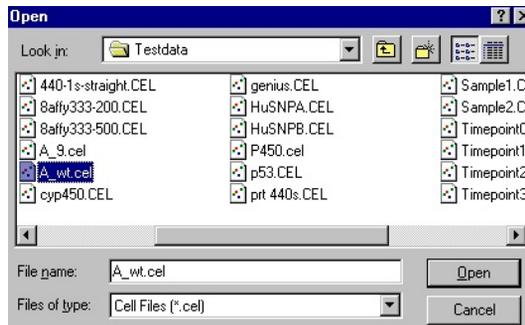


Figure G.4
Open dialog box

2. Select the desired *.cel file and click the **Open** button (or double-click the *.cel file name).
 - ⇒ The Mixture Detection Analysis Settings dialog box displays the directory location and name of the reference p53 wildtype reference DNA file (**Figure G.5**).

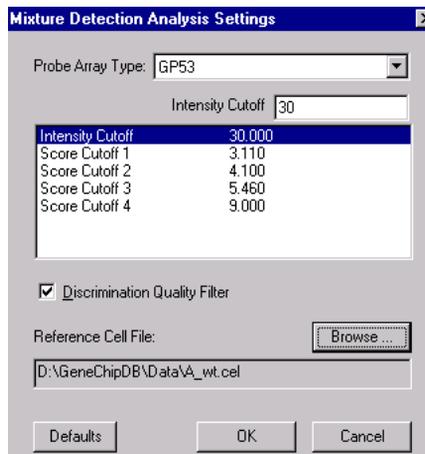
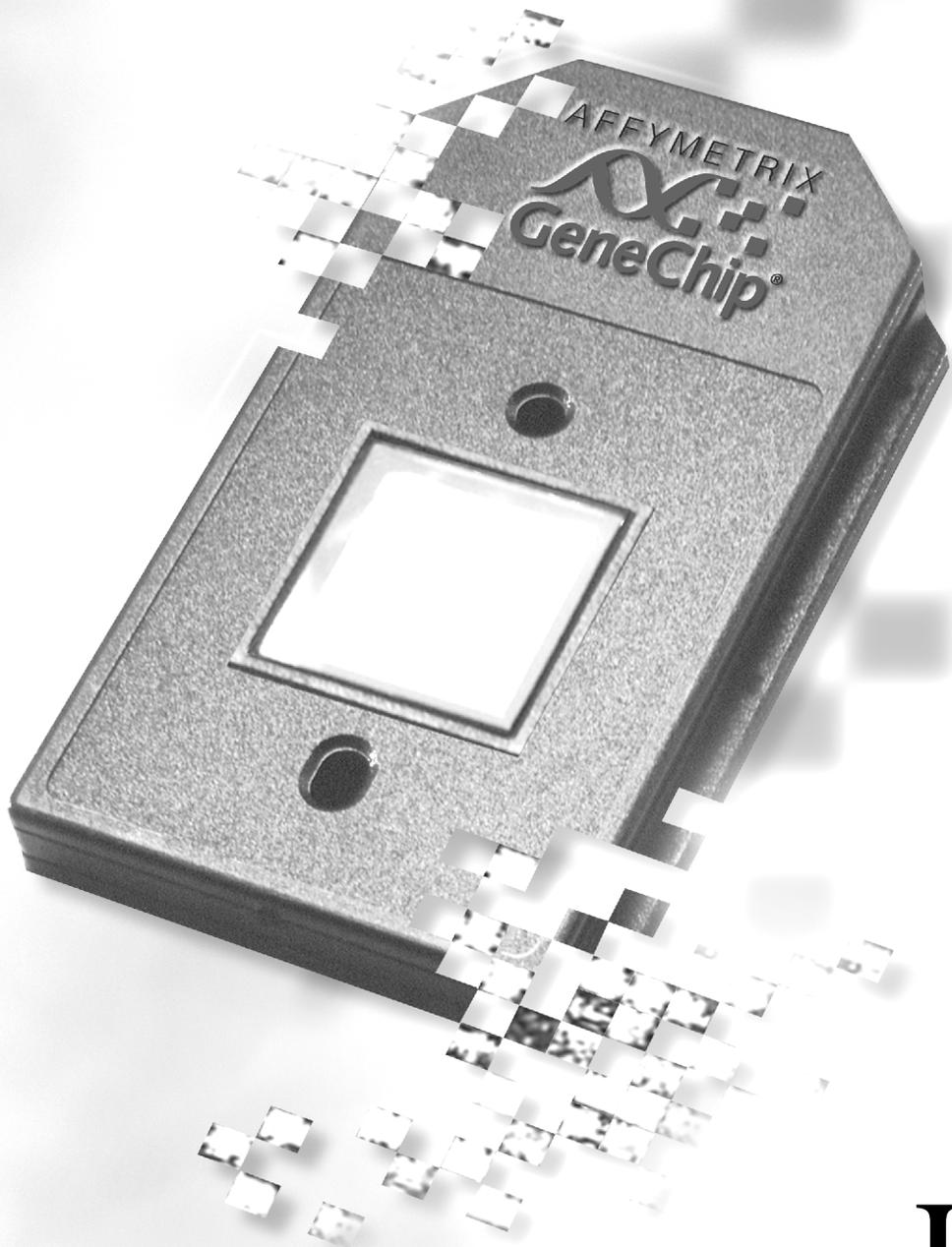
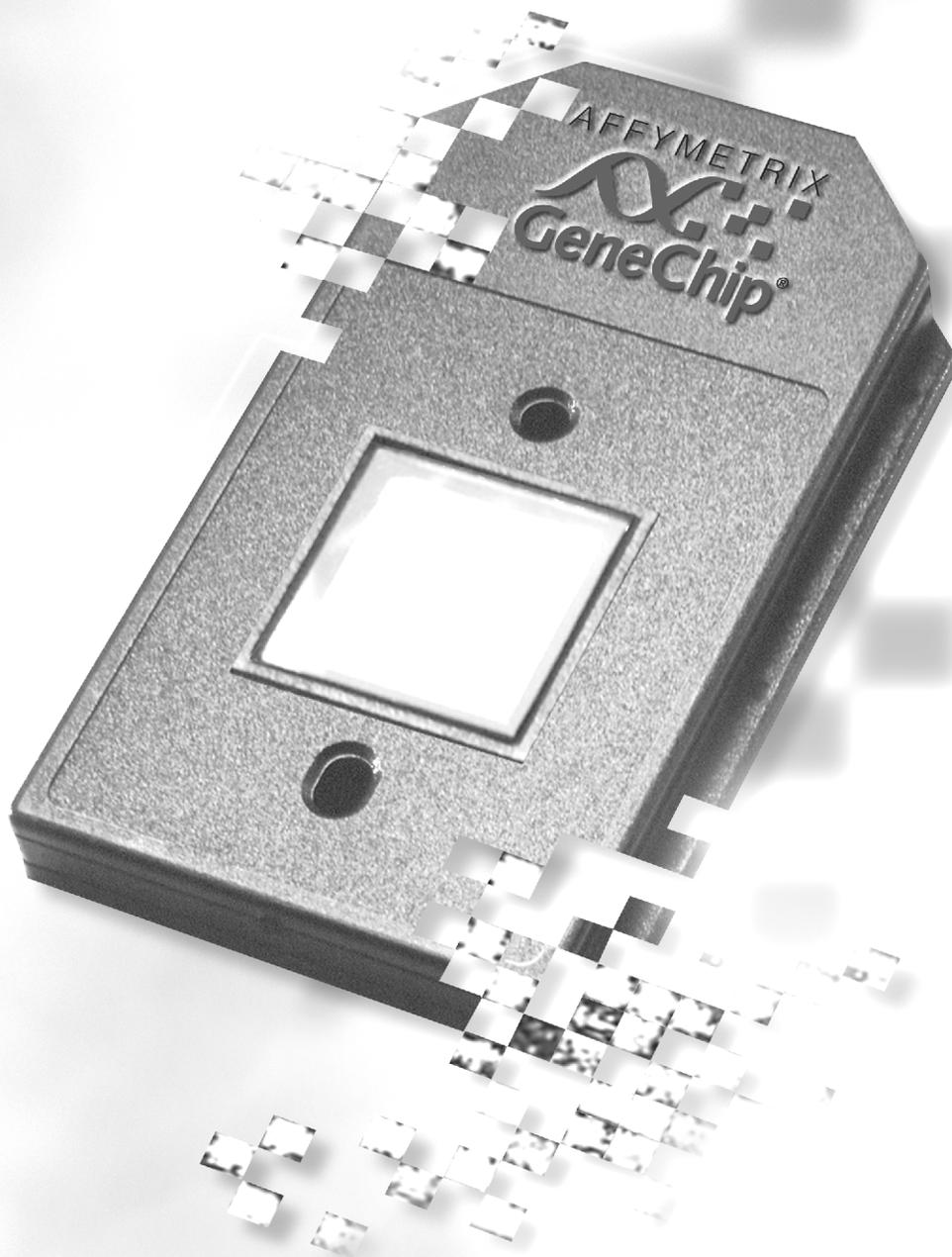


Figure G.5
Mixture Detection Analysis Settings dialog box



H

Appendix H





Genotyping Analysis Settings

The Genotyping algorithm evaluates the quality of the hybridization intensity data from each miniblock (a set of four probes: PM_A , MM_A , PM_B , MM_B), which must pass quality controls tests; otherwise, the data from a block (which includes five miniblocks) is not used in the analysis.

The Genotyping algorithm relies on the analysis settings to derive biologically meaningful results from GeneChip® HuSNP™ Mapping probe array intensity data. Some of the analysis settings are user-modifiable. The default values were empirically optimized through extensive testing at Affymetrix and should only be modified by expert users.

✓ NOTE

The Genotyping algorithm is only available when Microarray Suite is run in disk files mode.

Viewing & Editing Analysis Settings

1. In the shortcut bar, click **Analysis Settings**, then click **Genotyping** . Alternatively, select **Tools** → **Analysis Settings** → **Genotyping** from the main menu.
⇒ The Genotyping Analysis Settings dialog box appears ([Figure H.1](#)).

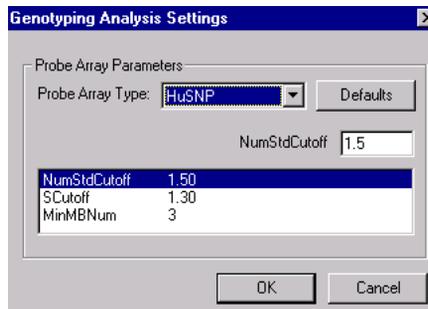


Figure H.1
Genotyping Analysis Settings dialog box

2. Select **HuSNP** from the **Probe Array Type** drop-down list.
The settings are specific for the selected probe array type and do not affect the settings for other types of probe arrays.
3. To edit a parameter value (**Figure H.2**):
 - a. Highlight the parameter in the lower box.
 - b. Enter the new value in the upper box.
 - c. Click the lower box to display the new parameter value.

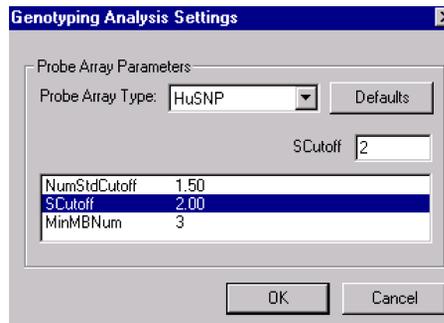
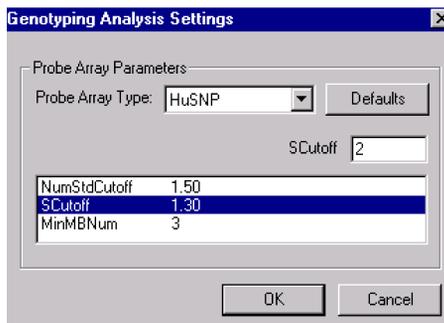


Figure H.2
Genotype Analysis Settings; enter a new parameter value in the upper box and click the lower box (left) to save the new parameter value (right)

Number of Standard Deviations Cutoff

The number of standard deviations cutoff (NumStdCutoff) is a number of standard deviation units. The intensity data from a miniblock of four probe cells passes one of the quality controls tests if:

$$(PM_A + PM_B - \text{Mean})/\text{Std} > \text{NumStdCutoff}$$

where:

PM_A = probe designed to be a perfect match to allele A, PM_B = probe designed to be a perfect match to allele B

$$\text{Mean} = \frac{\left(\sum_{i=1}^{Nmm} MMi \right)}{Nmm}$$

Nmm = number of mismatch probes (both A and B) for the block

MMi = the intensity of mismatch probe i

$$\text{Std}^2 = \left(\sum_{i=1}^{Nmm} (MMi - \text{Mean})^2 \right) / (Nmm - 1)$$

$$\text{Std} = \sqrt{\text{Std}^2}$$

Increasing the NumStdCutoff threshold increases the stringency of the analysis.

S Cutoff

The intensity data from a miniblock of four probe cells passes one of the quality control tests if:

$$PM_A/MM_A \geq S \text{ Cutoff OR } PM_B/MM_B \geq S \text{ Cutoff}$$

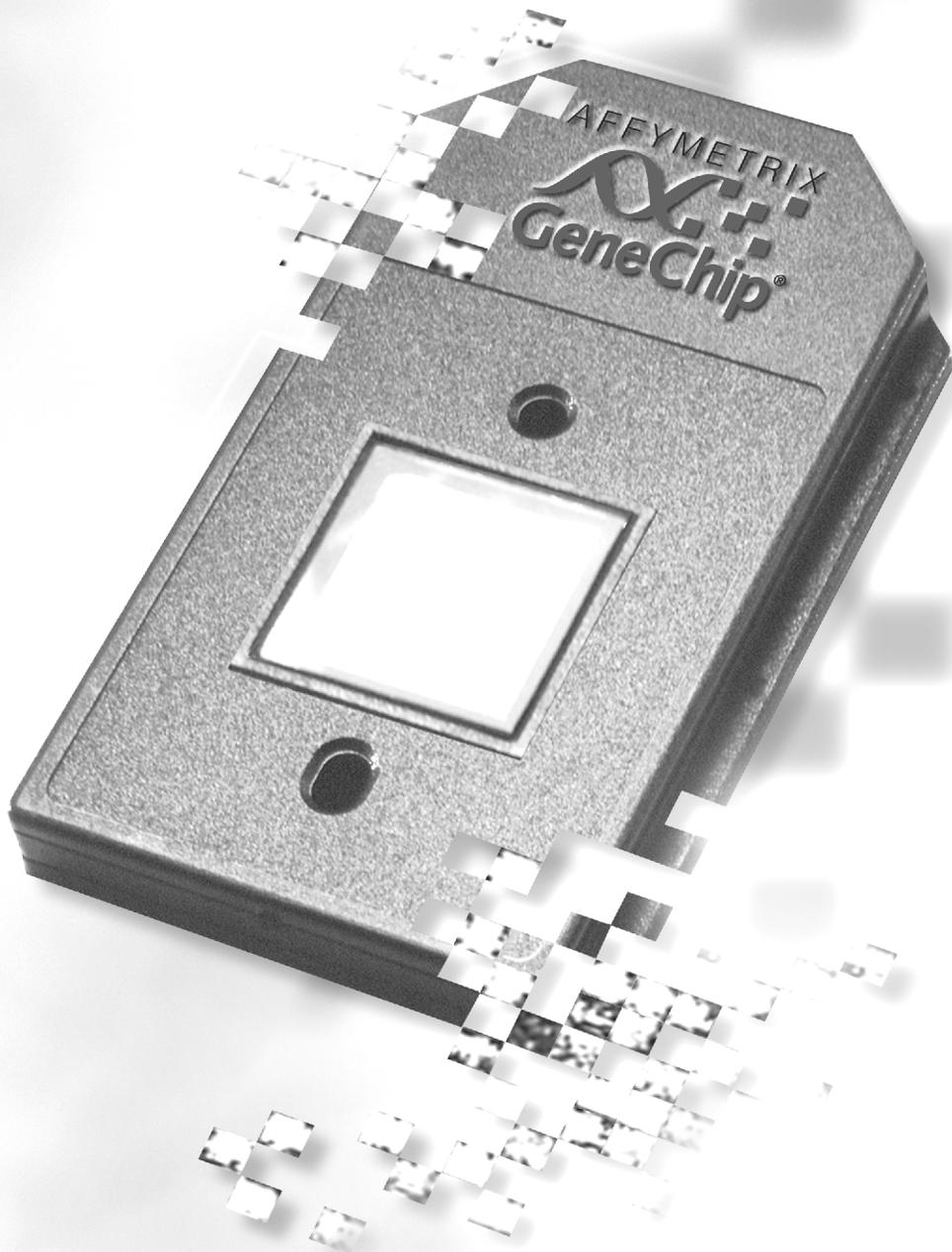
where PM_A = perfect match probe designed to be perfectly complementary to allele A, PM_B = perfect match probe designed to be perfectly complementary to allele B, MM_A = mismatch probe to allele A, MM_B = mismatch probe to allele B.

Increasing the S Cutoff threshold increases the stringency of the analysis.

Minimum Miniblock Number

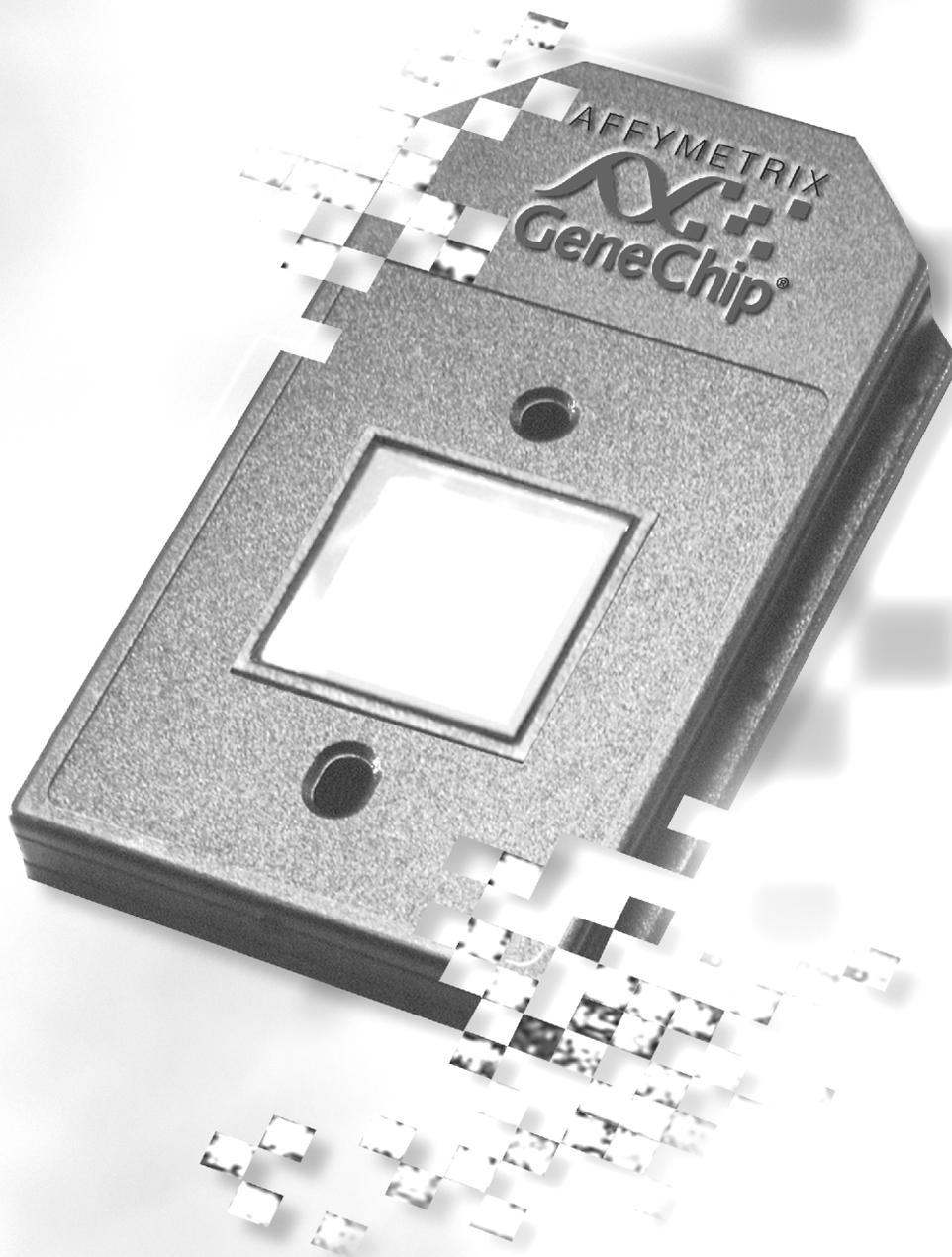
The minimum miniblock number (MinMBNum) specifies the minimum number of miniblocks per block that must pass all of the quality control tests so that the data from that block may be used in the analysis.

Increasing the minimum miniblock number increases the stringency of the analysis.



I

Appendix I





Genotype Analysis Viewer Report

Overview

The Genotyping (GT) Analysis Viewer report facilitates in-depth analysis of Affymetrix® HuSNP™ probe array data. It provides the relative allele signal (RAS) for each marker used in the Affymetrix® HuSNP™ assay. RAS is a quantitative representation of the presence of the two possible alleles A and B. If only one sample *.cel (*A.cel and *B.cel) is loaded into the GT Viewer, the report simply displays the RAS value for the sense and anti-sense block of each marker under the headings RAS1 and RAS2 respectively.

The GT Viewer algorithm detects variations between the genotype calls of a control and an experimental sample. When you first load a control and then an experimental sample into the GT Viewer, the algorithm computes the difference in RAS (Delta RAS) between the two samples. Delta RAS indicates the shift in signal from the control genotype call to the experimental genotype call and reports this value in the output for each marker.

The RAS distribution of a particular SNP is dependent on many factors that may be unique to that SNP, such as surrounding sequences. To address the widely distributed RAS values for different SNPs, the algorithm uses the information in the HuSNP.tab table file. The table contains the standard deviation of RAS (Std RAS) for all heterozygote (AB) calls for each marker. The data were obtained from an Affymetrix® reference set of 82 unrelated CEPH DNA samples that were analyzed without the pattern recognition steps of the Affymetrix® HuSNP™ algorithm. The GT Viewer Analysis report uses information from this table to compute and display a normalized version of the DeltaRAS values in the StdUnits column.

The StdUnits is an approximate normalization of Delta RAS in the presence of SNP-to-SNP variations. It is important to avoid over-interpreting the StdUnit data. There are two reasons to be cautious in deriving the

significance of DeltaRAS from the StdUnit value. First we assume a normal distribution when computing a standard deviation value. This may not be a safe assumption since some RAS values have been shown to deviate from normality to various uncharacterized degrees.

Secondly, it may be reasonable to expect that the standard deviations supplied in the default HuSNP.tab may only approximately represent what is observed under different experimental conditions.

Due to these factors you are strongly encouraged to create a customized table of RAS standard deviations from their control samples. In addition, you should consider both the raw DeltaRAS and the StdUnit values when attempting to detect patterns of allelic shifts

HuSNP.tab File

The HuSNP.tab table enables the GT Analysis Viewer report to incorporate empirical data on the likely RAS distribution of Affymetrix® HuSNP™ markers. Ideally, users analyzing a large number of control and experimental pairs can create a customized version of the HuSNP.tab table in order to capture the RAS standard deviation of the AB genotype call from the control set of samples. We recommend that the customized table be based on data from a minimum of ten control samples. (See the following section on how to customize the HuSNP.tab heterozygote standard deviation table.)

Table 1.1
HuSNPtab table

Marker Name	HuSNP Unit #	Std_Blkl	Std_Blkl2	N1	N2
Waif-A	1	0.01	-1000	33	0
Waif-B	2	0.1	-1000	21	0
Waif-C	3	0.15	0.44	5	5
Waif-D	4	0.2	0.1	10	10
Waif-E	5	0.36	-1000	15	14
Waif-F	6	0.1	0.4	7	8

Marker Name	The Affymetrix® HuSNP™ SNP name.
HuSNP Unit #	The consecutive number for the SNP on the Affymetrix® HuSNP™ array.
Std_Blkl	The standard deviation of the heterozygote (<i>AB</i>) distribution of RAS values for the first used block of the SNP.
Std_Blkl2	The standard deviation of the heterozygote (<i>AB</i>) distribution of RAS values for the second used block of the SNP. Note: the majority of HuSNP markers do not have a second used block.
N1	The number of individuals used to derive the standard deviation values for the first used block of the SNP.
N2	The number of individuals used to derive the standard deviation values for the second used block of the SNP.

The HuSNP.tab table has one row per SNP and displays the standard deviation for RAS values (Std RAS) for each marker in Std_Blkl. If a second block RAS does not exist on the probe array or was not used, the table displays a value of -1000. If available, the table displays a second Std RAS value in Std_Blkl2. If less than three data points were available, the Std is set to 0.039 which is the median value of the distribution.

Columns N1 and N2 are not used by the GT Viewer report and are included as readable information only. N is the number of individuals from which the std values were derived.

The value in the StdUnit column is computed automatically by the report function as shown below.

If only one block is available:

$$StdUnit = (DeltaRAS)/(StdBlkl)$$

If two block are available:

$$StdUnit = (DeltaRAS)/(\sqrt{StdBlk1^2 + StdBlk2^2})$$

Making a Customized HuSNPtab Table

The following instructions explain how to make a customized HuSNP.tab heterozygote standard deviation table from a user-defined data set.

- 1.** Analyze the reference sample set in Microarray Suite using the Genotyping algorithm (HuSNP™ analysis).
- 2.** Run the Genotyping Viewer report for all control samples, either individually or in pairs.
- 3.** Save the RAS values for Block 1 and Block 2 for each sample to a user-defined database.
- 4.** For as many control samples as possible, collect the heterozygote calls for each marker.
- 5.** Compute the standard deviation for heretozygote RAS values. When both are available, compute a separate value for RAS1 and RAS2 for each marker.
- 6.** Create a custom version of the HuSNP.tab file with user-defined standard deviation values for RAS.
- 7.** Replace the HuSNP.tab file in the library directory of Microarray Suite with the custom *.tab file created in [step 6](#).

✓ NOTE

The custom table must be named HuSNP.tab.

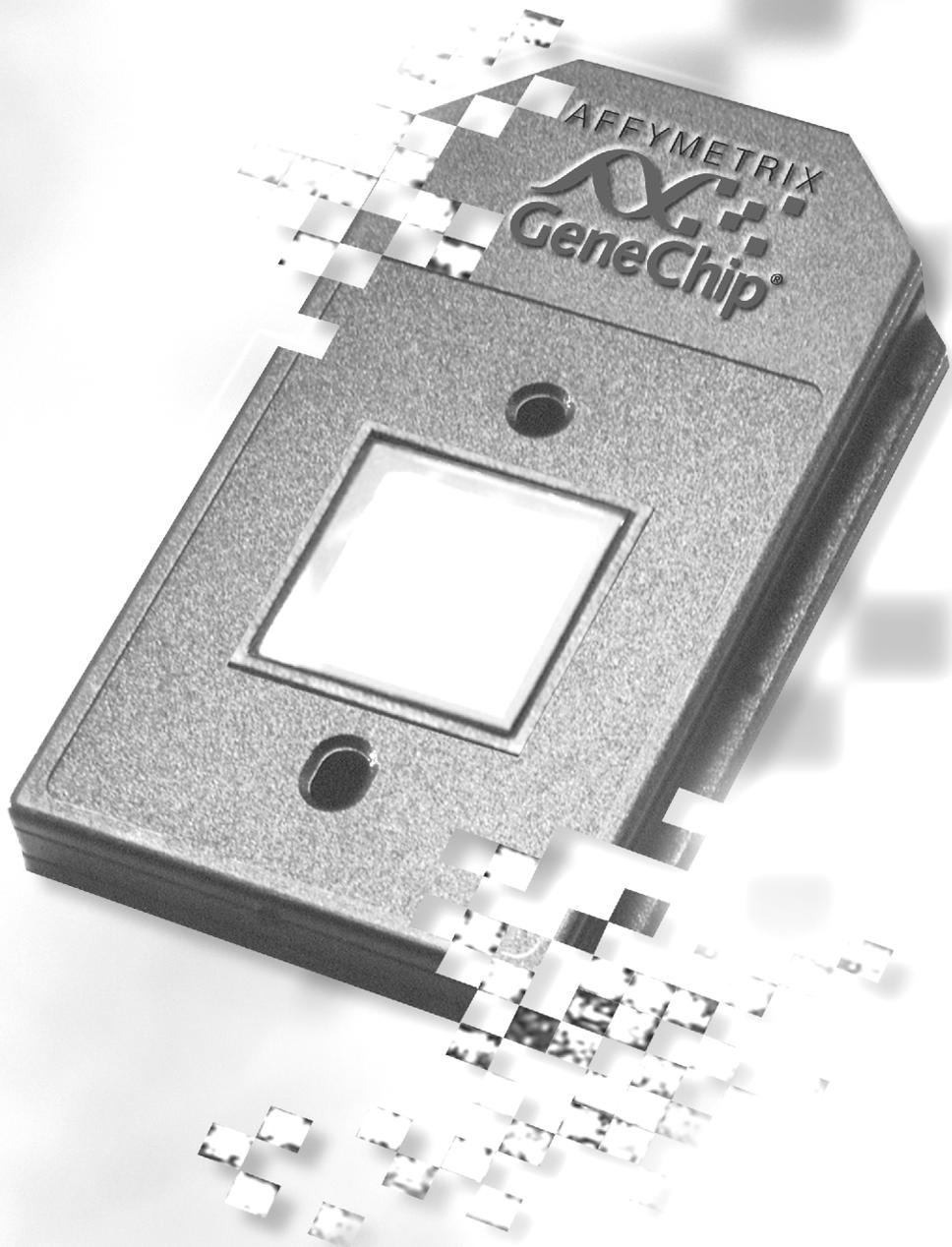
The original HuSNP.tab default file can be renamed or archived in a different location.

- 8.** If the standard deviation is available for both blocks of RAS, both should be listed in the table, under columns Std_Blkl and Std_Blk2. If std RAS for a marker is available for only one block, regardless if it is from sense or anti-sense, it should be listed in the Std_Blkl column. In the absence of a second std RAS value, enter *-1000* in the Std_Blk2 column.

9. If you cannot determine the standard deviation for a particular marker, the standard deviation value provided in the original HuSNP.tab from Affymetrix can be used for that marker. Alternatively, enter 0.039, the median value of the std distribution. If there is no value listed in the HuSNP.tab for a marker's Std RAS in either column, the report output automatically displays the -9 code in the StdUnit column.

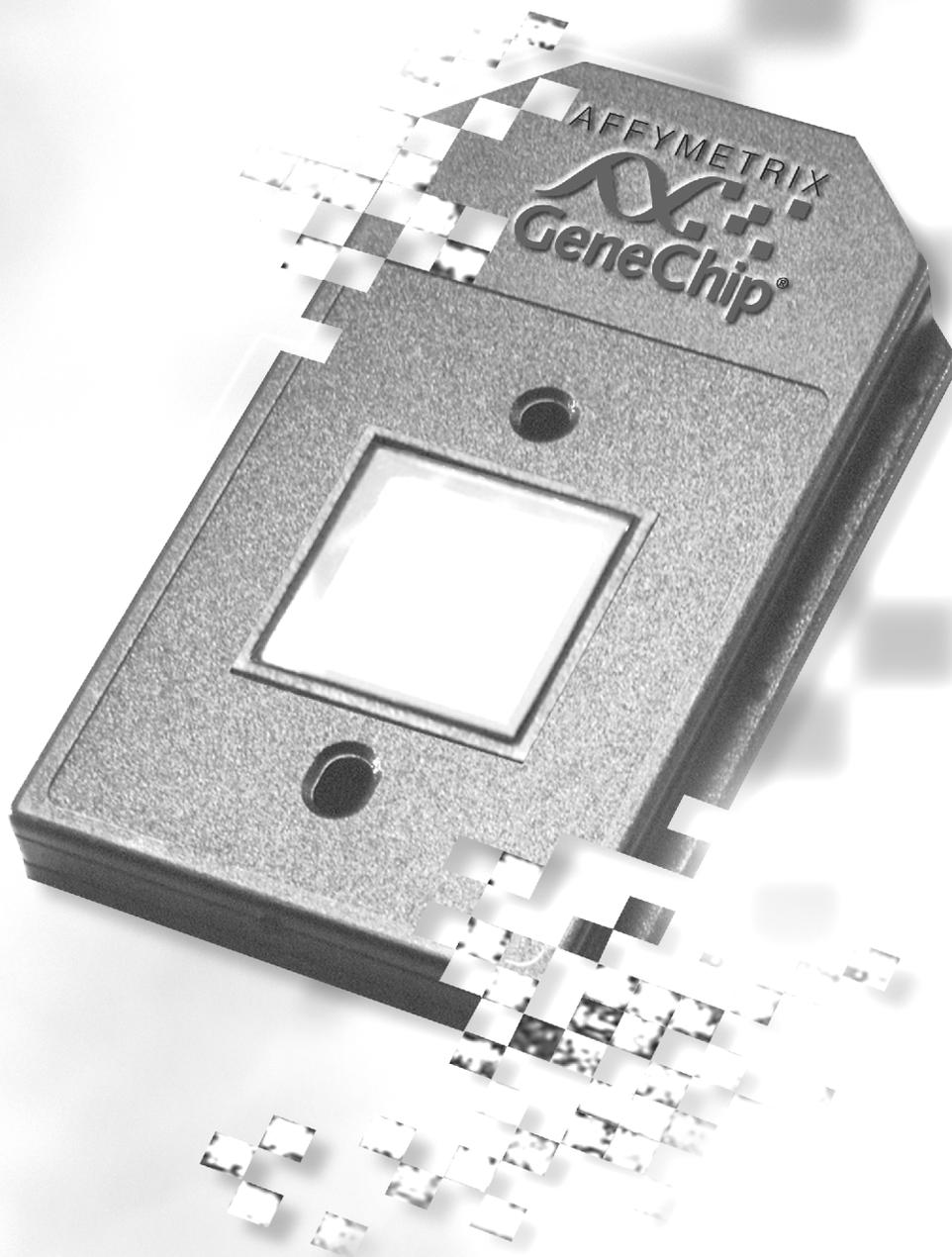
✓ NOTE

The N1 and N2 columns can be ignored in the customized user-defined HuSNP.tab file. A HuSNP™ marker may have a sense and an anti-sense block tiled on the array. When both sense and antisense blocks are present, the values are reported in the RAS1 and RAS2 columns. A RAS of -1000 indicates the block does not exist or is not used to determine the call. A RAS of -9 indicates the block failed quality control. If both blocks are present and used in the analysis, a new RAS value is derived from RAS1 and RAS2.



J

Appendix J



You can display toolbars with text labels (Figure J.1). To display the toolbar button labels, select **View** → **Toolbar** → **Text Labels** from the menu bar.

Main Toolbar



Figure J.1
Main toolbar

Table J.1
Main toolbar

Menu Bar Command	Toolbar Button	Function
File → Open		Displays the Open dialog box so that a file (for example, *.chp or *.cel) may be opened.
File → Save		Saves the open image, experiment, or report.
File → Print		Displays the Print dialog box.
Window → Data File Tree		Displays or hides the data file tree.
Window → Shortcut Bar		Displays or hides the shortcut bar.
Window → Status Log		Displays or hides the status log.

Table J.1
Main toolbar

Menu Bar Command	Toolbar Button	Function
Run → Experiment Info		Displays experiment window.
Run → Fluidics		Displays the Station Selection dialog box for the Affymetrix® Fluidics Station 400.
Run → Scanner		Displays the Scanner dialog box.
Run → Stop Scanner		Stops a scan in progress.
Run → Analysis		Runs an analysis on an open *.dat or *.cel file.
Edit → Image Settings		Displays the Image Settings dialog box.
Help → Contents		Displays Microarray Suite software help.

Expression Analysis Window (EAW) Toolbar



Figure J.2
EAW toolbar

Table J.2
EAW toolbar button functions

Menu Bar Command	EAW Toolbar Button	Function
Edit → Find		Displays the Find Probe Set dialog box.
Edit → Sort		Displays the Sort dialog box.
View → Hide Selected		Hides selected probe set(s) in the metrics or pivot table.
View → Hide Unselected		Hides unselected probe set(s) in the metrics or pivot table.
View → Unhide All		Displays all probe sets previously hidden in the metrics or pivot table.
Analysis → Options		Displays the Analysis Options dialog box.
Graph → Clear Selected Graphs		Clears selected graphs from the graph pane in the EAW.
Graph → Scatter Correlation Graph		Displays the Scatter Graph dialog box.
Graph → Series Graph		Displays the Select Series Graph Parameter dialog box.
Graph → Intensity Bar Graph		Plots the Intensity Bar Graph for the selected probe set(s).

Table J.2
EAW toolbar button functions

Menu Bar Command	EAW Toolbar Button	Function
Graph → Measured Images		Displays the hybridization intensity image data (from the *.dat) for the selected probe set(s).
Graph → Lasso Points		Changes the cursor to a drawing tool that enables the user to draw a circle around points in a scatter graph.

Sequence Analysis Window (SAW) Toolbar



Figure J.3
SAW toolbar

Table J.3
SAW toolbar button functions

Menu Bar Command	SAW Toolbar Button	Function
Edit → Find		Displays the Find Probe Set dialog box.
View → Amino Acids View → Nucleotides View → Amino Acids and Nucleotides		Toggles the view in the SAW between amino acids, nucleotides, or amino acids and nucleotides.
View → Intensity		Displays intensity graphs for all probe sets.
View → Hide		Hides selected probe sets.
View → Unhide All		Displays all previously hidden probe sets.
Analysis → Options		Displays the Analysis Options dialog box.

Nucleotide Analysis Window (NAW) Toolbar



Figure J.4
NAW toolbar

Table J.4
NAW toolbar button functions

Menu Bar Command	NAW Toolbar Button	Function
Edit → Find		Displays the Find Probe Set dialog box.
View → Hide		Hides selected probe sets.
View → Unhide All		Displays all previously hidden probe sets.
Edit → Sort		Displays the Sort dialog box.
Graphs → Clear Selected Graphs		Clears selected graphs from the graph pane in the NAW.
Graphs → Draw Intensity Graphs		Displays intensity graphs for the selected probe set(s).
Graphs → Draw Measured Image		Displays the hybridization intensity image data (from the *.dat) for the selected probe set(s).
Analysis → Options		Displays the Analysis Options dialog box.

GT Viewer Toolbar



Figure J.5
GT Viewer toolbar

Table J.5
GT Viewer toolbar button functions

Menu Bar Command	GT Viewer Toolbar Button	Function
Edit → Add Cell File		Adds the selected HuSNP™ *.cel files (allele A and allele B) to the GT Viewer.
Edit → Remove Cell File		Removes the selected allele A and allele B *.cel files from the GT Viewer.
Edit → Start Analysis		Starts the GT Viewer analysis.
Edit → Stop Analysis		Stops the GT Viewer analysis.

Hybridization Analysis Window (HAW) Toolbar



Figure J.6
HAW toolbar

Table J.6
HAW toolbar button functions

Menu Bar Command	HAW Toolbar Button	Function
Edit → Find		Displays the Find Probe Set dialog box.
Edit → Copy		Copies user-specified selection in the HAW data table to the system clipboard.
Edit → Remove Experiments		Displays the Select Experiments to Remove dialog box.

Batch Analysis Toolbar



Figure J.7
Batch analysis toolbar

Table J.7
Batch analysis toolbar button functions

Menu Bar Command	Batch Analysis Toolbar Button	Function
Edit → Add Item		Displays the Open dialog box to select and add *.cel files to the Batch Analysis window.
Edit → Remove Item		Removes the selected *.cel files from the Batch Analysis window.
Edit → Start Analysis		Starts the batch analysis.
Edit → Stop Analysis		Stops the batch analysis.
View → Options		Displays the Batch Analysis Options dialog box.

Report Toolbar



Figure J.8
Report window toolbar

Table J.8
Report toolbar button functions

Menu Bar Command	Report Toolbar Button	Function
Edit → Cut		Removes a highlighted selection from the report.
Edit → Copy		Copies a highlighted selection in the report to the system clipboard.
Edit → Paste		Pastes a copied or cut selection at the current insertion point.
Edit → Find		Displays the Find dialog box.
Edit → Find Next		Performs a text search for the item specified in the Find dialog box.

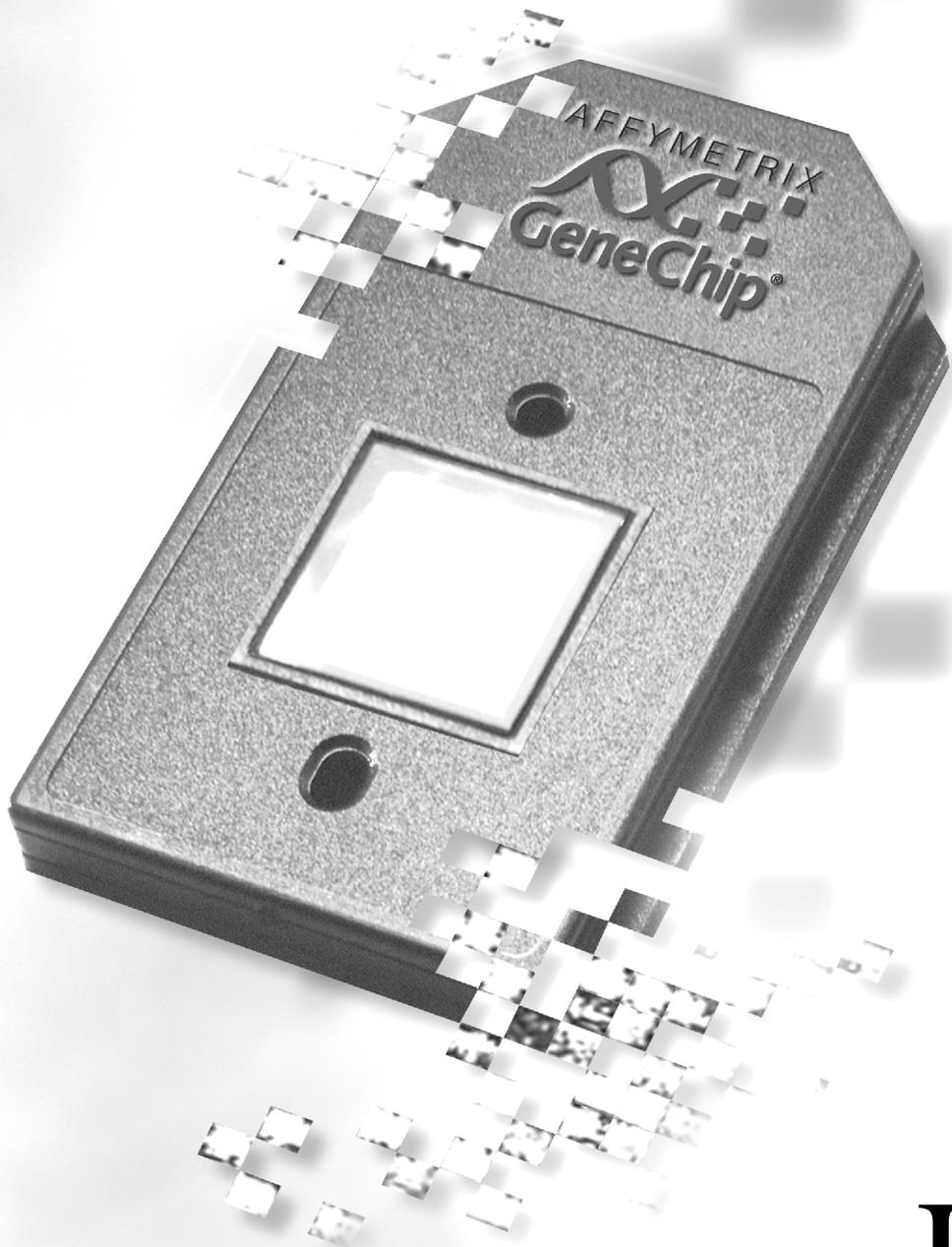
Publish Toolbar



Figure J.9
Publish window toolbar

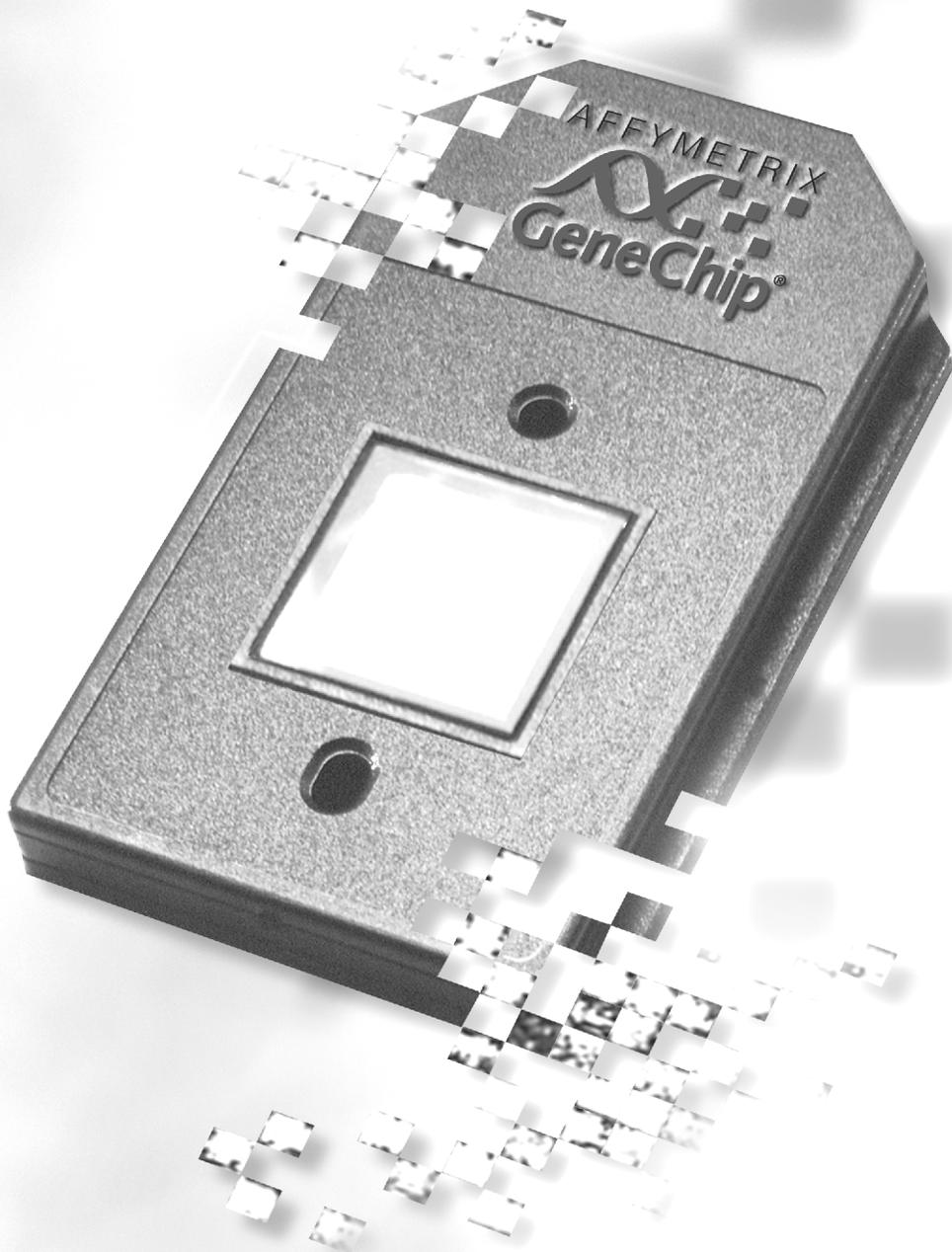
Table J.9
Publish toolbar button functions

Menu Bar Command	Publish Window Toolbar Button	Function
Publish → Add Item		Displays the Open dialog box to select and add experiment data files (<i>task items</i>) to the publish task.
Publish → Remove Item		Removes the selected task item(s) from the publish task.
Publish → Publish		Send the task to the LIMS server to be published during the time specified by the LIMS administrator.
Publish → Monitor → Cancel Publish		Cancels a task (changes the task status from WAIT to CANCELED).
Publish → Monitor → Restart Publish		Reinstates a canceled task to be published during the time specified by the LIMS administrator (changes the task status from CANCELED to WAIT).



K

Appendix K





Working With Windowpanes & Columns

Resizing Windowpanes

You can resize windowpanes.

1. Place the mouse pointer over a windowpane border so that it changes to a double arrow  (Figure K.1).

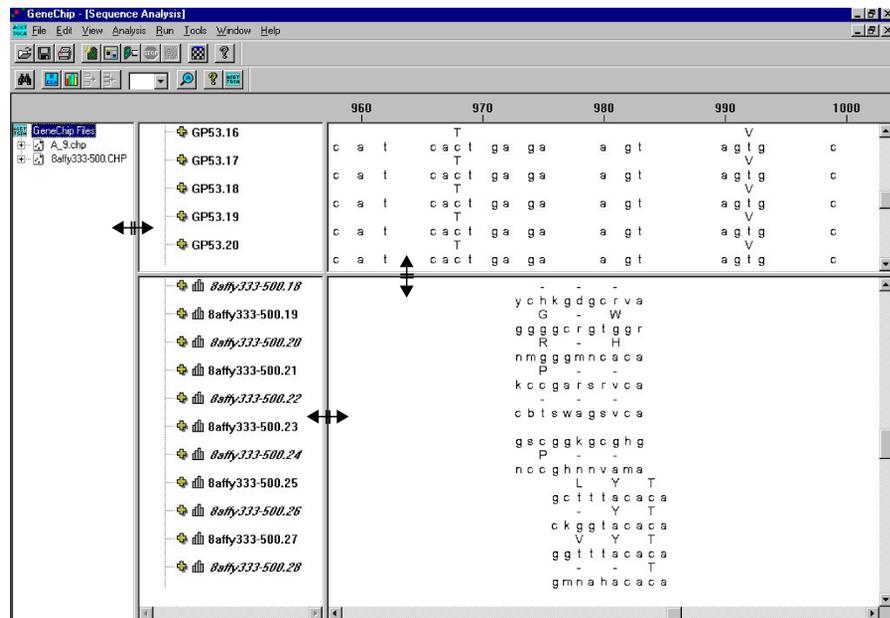


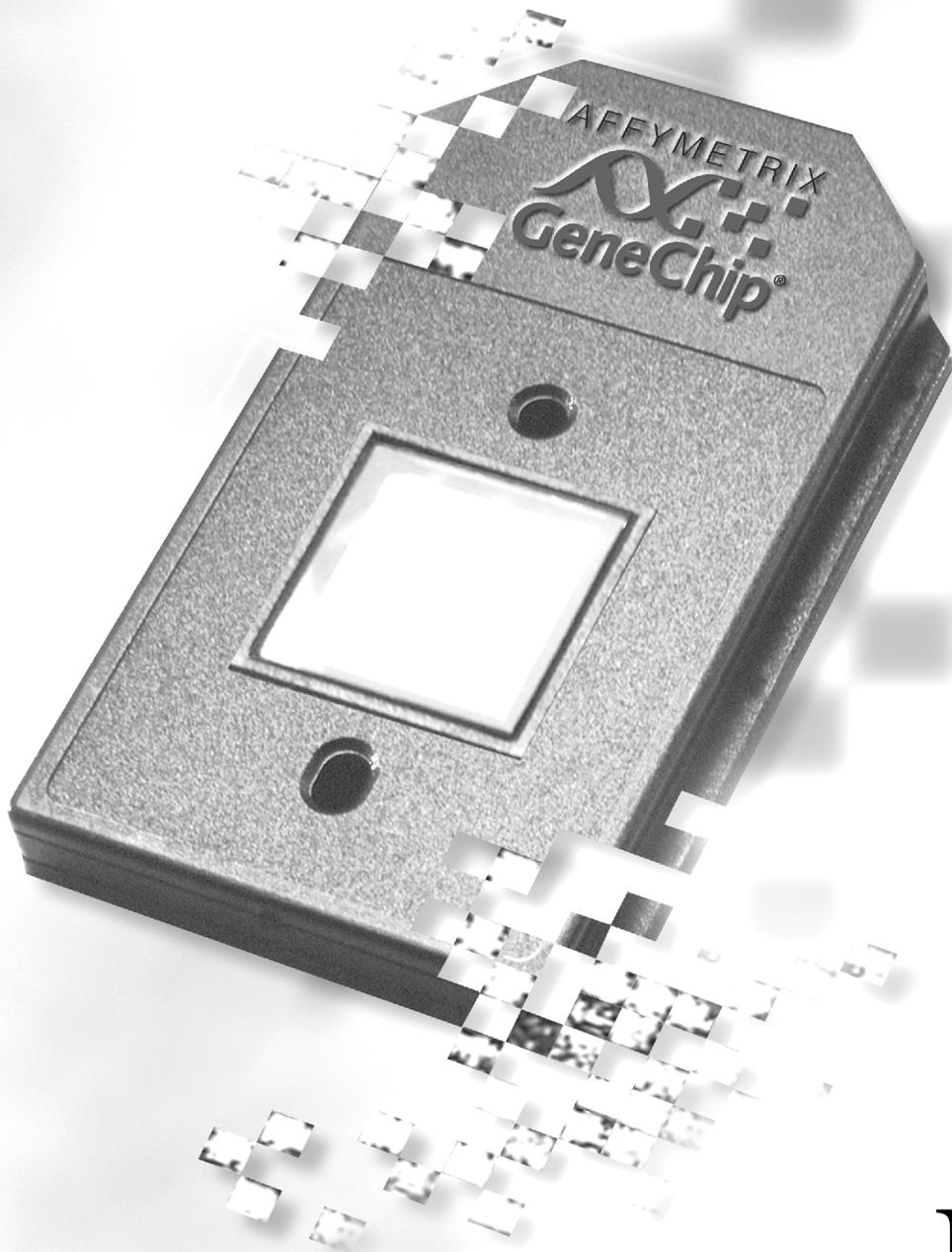
Figure K.1
Resize windowpanes horizontally or vertically

2. Drag the border to resize the windowpane.

Resizing or Hiding Columns in the EAW or NAW

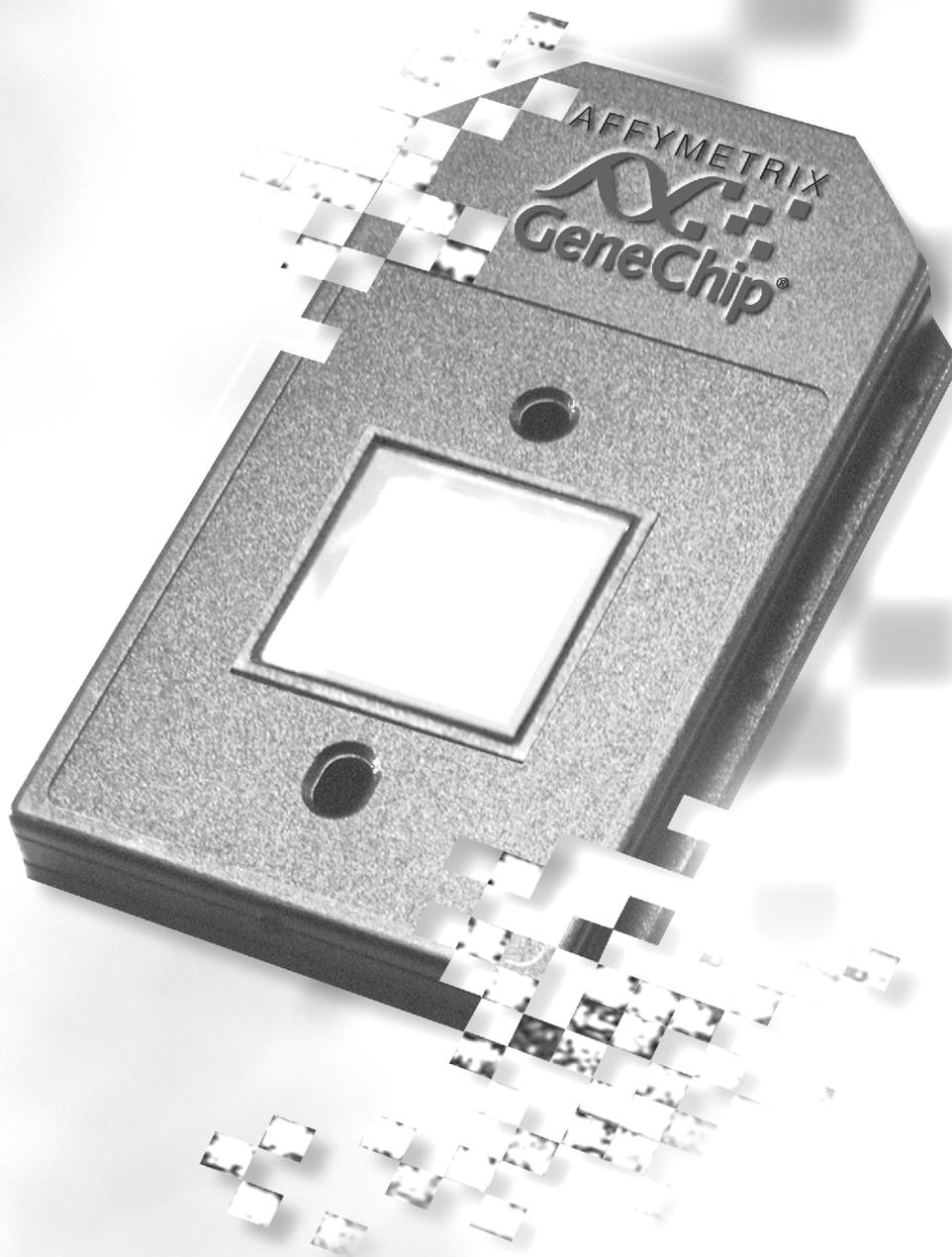
You can resize or hide columns in the EAW or NAW. Microarray Suite stores the column settings for future sessions on a per user basis (identified by the logon name) so that one user's settings do not affect the settings of another.

- 1.** Position the mouse over the left or right cell border in the column header so that it changes to a double arrow \leftrightarrow .
- 2.** Drag the cell border to resize the width of the column.
- 3.** To hide a column, drag the left or right cell border of the column header until the column width is reduced to zero. Alternatively, right-click the column header in the EAW and click **Hide Column** in the shortcut menu.



L

Appendix L





Base Codes and Amino Acid Abbreviations

IUPAC Base Codes

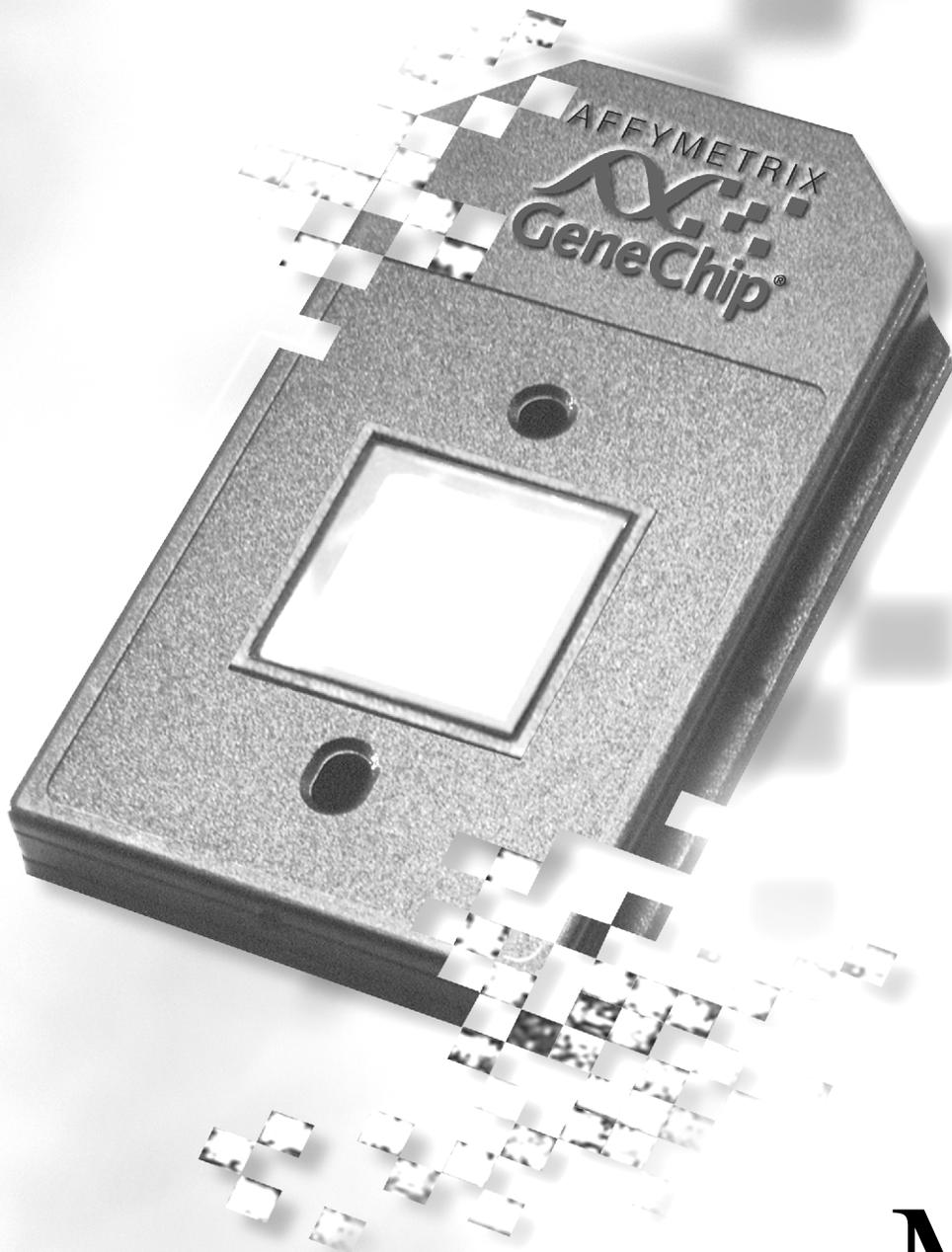
Table L.1
Base Codes

IUPAC Code	Group	Base(s)
A	A	Adenine
C	C	Cytosine
G	G	Guanine
T	T	Thymine
M	A or C	aMino
R	A or G	puRine
W	A or T (U)	Weak interaction (2 H bonds)
Y	C or T (U)	pYrimidine
S	C or G	Strong interaction (3 H bonds)
K	G or T(U)	Keto
V	A or C or G	not-T or not-U (since V follows U)
H	A or C or T(U)	not-G (since H follows G)
D	A or G or T(U)	not-C (since D follows C)
B	C or G or T(U)	not-A (since B follows A)
N	A, C, G or T(U)	aNy

Amino Acid Abbreviations

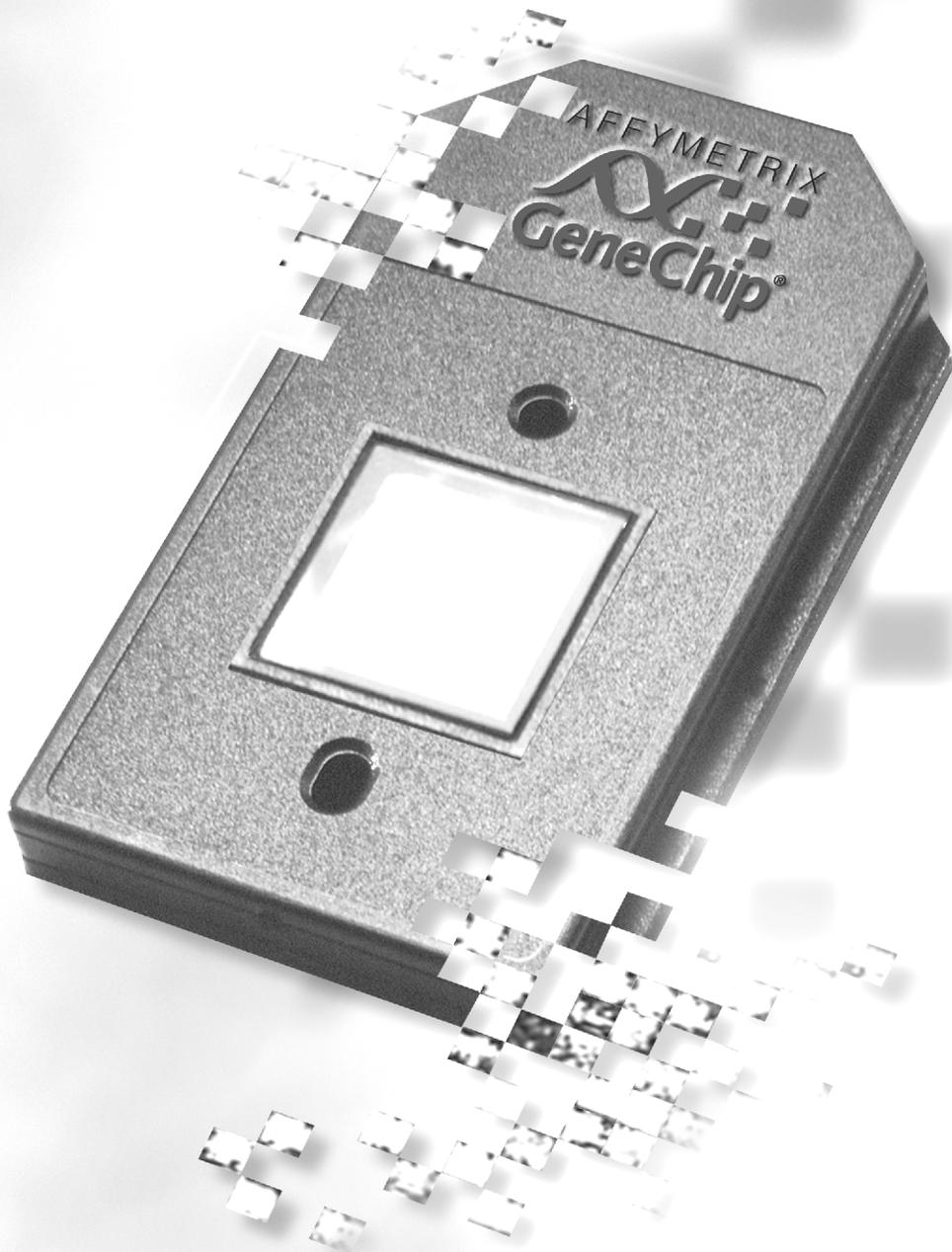
Table L.2
Amino acid abbreviations

One Letter	Three Letter	Amino Acid
A	Ala	Alanine
B	Asx	Asparagine or Aspartic acid
C	Cys	Cysteine
D	Asp	Aspartic acid
E	Glu	Glutamic acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	Hi	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine
Z	Glx	Glutamine or Glutamic acid



M

Appendix M





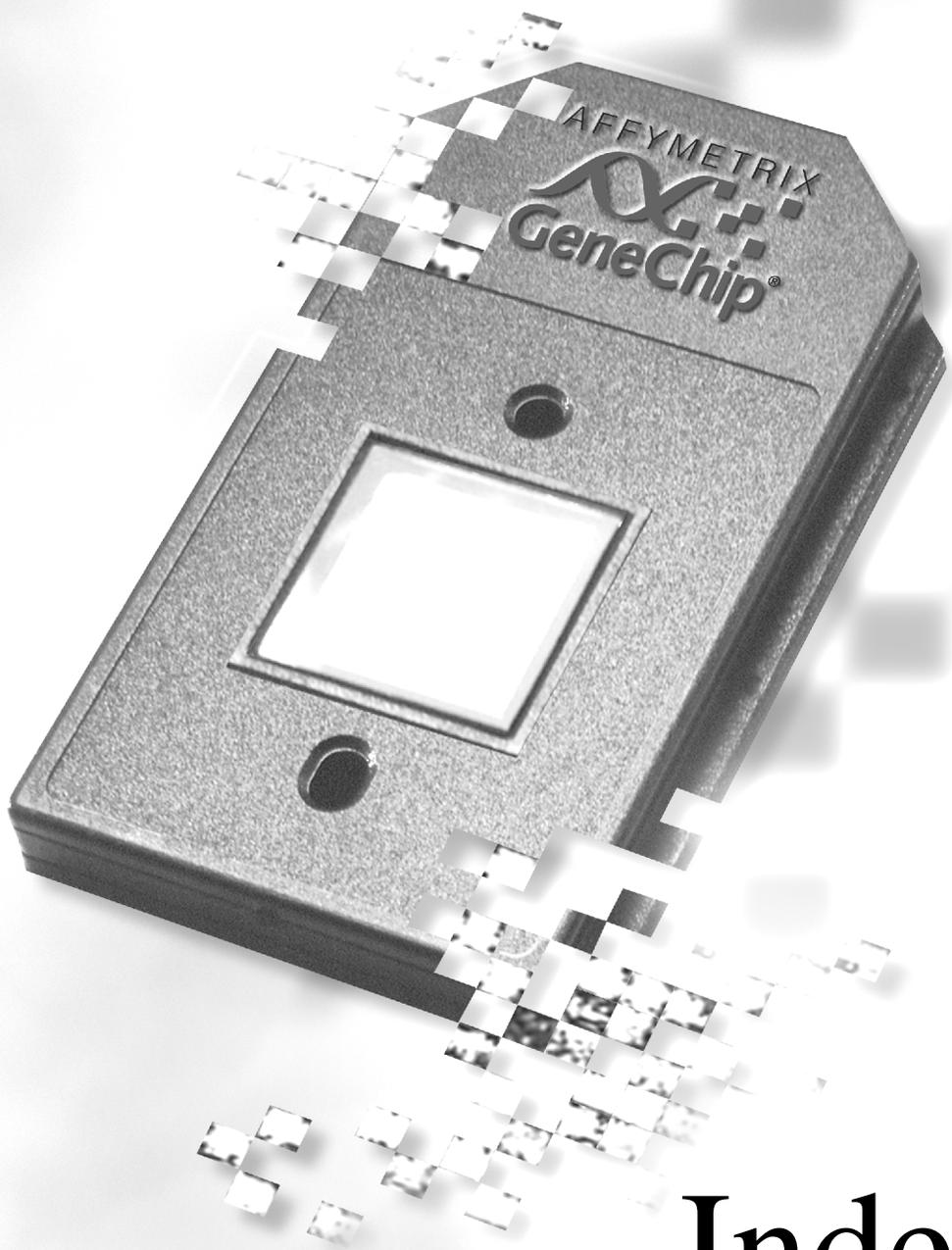
Hot Key Descriptions

Table M.1
Hot key descriptions

Menu Bar Command	Hot Key
File > Open	Ctrl + O
File > Print	Ctrl + P
Edit > Copy	Ctrl + C
File > Save	Ctrl + S
Edit > Image Settings	S
Edit > Delete	Del
Edit > Find Next	F3
Edit > Unmask All Cells	Shift + U
View > Grid	G
View > Probe Cell Data	C
View > Corner > Upper Left	F5
View > Corner > Upper Right	F6
View > Corner > Lower Left	F7
View > Corner > Lower Right	F8
View > Image > Measured	M
View > Image > Difference	D
View > Image > Average	A
View > Next Highlight	N
View > Previous Highlight	P
View > Clear Highlights	E
View > Probe Tiling	T

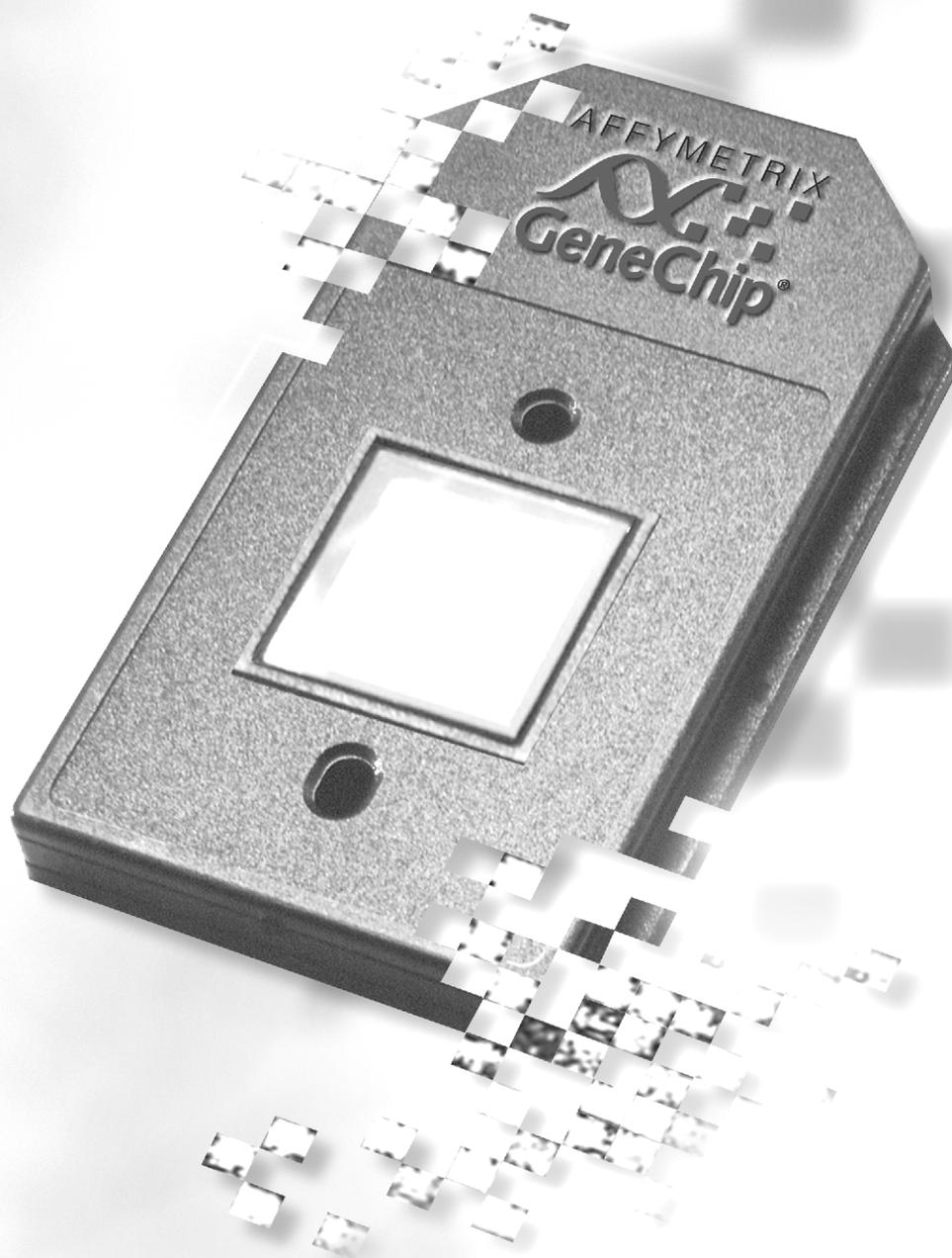
Table M.1
Hot key descriptions

Menu Bar Command	Hot Key
	1, 2, 3, etc. (toggle between open image windows)
	I (Image window zoom in)
	O (Image window zoom out)
	Shift+O (Image window full zoom out)
	L (intensity autoscale)
Edit > Mask Cells	X
Edit > Unmask Cells	U
Help	F1
Scroll up one page	Page Up
Scroll down one page	Page Down
Scroll left one page	Ctrl + Page Up
Scroll right one page	Ctrl + Page Down
Scroll up 1/10 page	Up arrow
Scroll down 1/10 page	Down arrow
Scroll left 1/10 page	Left arrow
Scroll right 1/10 page	Right arrow



Index

Index



A

Affymetrix technical support
26

algorithm

block 257
cell analysis 120
consensus rules 212
Expression 399
genotyping 257
mixture detection 210

algorithm settings

expression 441
genotyping 503
mixture detection 493

alignment bars 227

amino acid codes 536

amino acid data (SAW)

comparing 222–225
finding 235
going to 235
hiding 225
viewing 217

analysis information 169

exporting 161

analysis settings 362

B

baseline file 439

batch analysis

renaming results 308–310

batch analysis (disk files mode)

resuming 311
running 311
selecting baseline 300

batch analysis (disk files modes)

selecting files 299

batch analysis (LIMS mode)

resuming 311
running 311
selecting baseline 306

selecting files 302

selecting parameter set 305

Batch Analysis Window 297

removing files 310

Batch Analysis window

toolbar 525

batch file

exporting 23, 311
importing 23, 313

block algorithm 257

bookmarks 139

creating 140
editing 141
viewing 140

C

cell analysis algorithm 120

cell coordinates 353

cell image view 128

cell intensity file (*.cel) 120

client software

installing MAS 29

codes

amino acid 536
IUPAC base 536

columns

resizing, hiding, or reordering (EAW or NAW) 532

configuring

Fluidics Station 400 361
Microarray Suite software 64

consensus base calls

editing 240

consensus rules algorithm 212

controls

See expression report controls.

cross hybridization probe

mask 459
combining masks 466

creating 459

editing 464

CYP450

intensity graphs 269

options 271

plotting 271

report 277

printing 372

tabular data 265

D

data file tree 60

display options 63

data mining 49

database

LIMS 48

publish 49, 317

defaults

analysis settings 362

configuring fluidics stations 361

file locations 357–360

disk files mode 358

LIMS mode 359

image settings 352–356

descriptions 167

difference image view 128

disk files mode 64, 358

documentation

conventions used 24

online 25

E

editing

consensus base calls 240

experiment information 81

expression report 204

Entrez 167

experiment data files 67

experiment data storage

disk files mode 64

- LIMS mode 65
- experiment information**
 - editing 81
 - viewing 81
- experiment setup**
 - disk files mode 74–76
 - LIMS mode 76–80
- exporting data**
 - analysis information 161
 - batch file 311
 - genotype data 267
 - hybridization data 291
 - metrics table 163–165
- Expression algorithm** 399
- expression analysis**
 - absolute 151–153
 - analysis information 169
 - comparison 153–158
 - graphs
 - intensity bar graph 191
 - scatter graph 175
 - series graph 186, 191
 - measured images 172
 - metrics
 - metrics table 162
 - overview 146–149
 - pivot table 159
 - report 195–201, 328–334
 - settings 150
 - tabular data 159–162
- expression analysis settings**
 - baseline 439
 - normalization 422
 - probe masks 431
 - scaling 412
 - user-modifiable parameters 441
- Expression Analysis Window (EAW)** 159
 - toolbar 519
- expression report** 195–201, 328–334
 - controls 203
 - editing 204
 - find feature 204
 - generating 195
 - parameters 199–201

- printing 372
- settings 197

F

- file locations settings** 65–66
- file mode** 358
- file types, See Microarray Suite files.**

files

- importing (SAW) 241
- opening 61, 62
- setting locations 65–66
- viewing information 62

- filters (LIMS mode)** 67

find feature

- EAW 169
- expression report 204
- HAW 292
- NAW 264
- probe set information 169
- SAW 235

fluidics protocol

- bypassing steps 104
- editing 105
- resuming 103
- running 103
- setup 101

- fluidics protocol files** 67

Fluidics Station 400

- bypassing protocol steps 104
- configuring 98, 361
- editing a protocol 105
- priming 99
- protocol setup 101
- resuming a protocol 103
- running a protocol 103

G

genotype analysis

- exporting data 267
- graphs
 - CYP450 intensity graphs 269
 - HuSNP intensity graphs

- 273
- reports
 - CYP450 report 277
 - mapping report 280
- running an analysis 258
- tabular data
 - CYP450 265
 - HuSNP 266

- genotyping algorithm** 257

graphs

- CYP450 intensity graphs 269
- HIV PRT Plus intensity graph 228
- HuSNP intensity graphs 273
- intensity bar graph 191
- p53 score graph 232
- scatter graph 175–186
- series graph 186, 191

grid 356

- manual adjustment 119
- viewing 118

- GT Viewer** 23, 282, 345

- report 282
- selecting *.cel 282

H

- highlights** 126, 355

- HIV PRT Plus intensity graph**

- 228
 - options 229
 - plotting 229

- HIV PRT Plus report** 247

- generating 247
- printing 372

HuSNP

- intensity graphs 273
- tabular data 266

- HuSNP.tab** 510

- Hybridization Analysis Window (HAW)** 289, 291

- find feature 292
- printing 371
- toolbar 524

- hybridization data**

- copying 291

- exporting 291
- removing 292
- hybridization probe mask** 470
 - combining masks 478
 - creating 471
 - editing 475

I

- image data file** 110
- image settings** 121–127, 352–356
 - cell coordinates 353
 - color 353
 - current image only 122
 - global 123
 - grid 356
 - highlights 126, 355
 - intensity 123, 352
 - pixel coordinates 124, 353
- Image window** 115
 - opening multiple windows 117
 - task bar 117
- importing files**
 - Batch Analysis Window 313
 - SAW 241
- installing**
 - Microarray Suite 29
- intensity** 123, 352
 - cell image view 128
 - difference image view 128
 - measured image view 127
- intensity bar graph** 191–195
 - clearing 195
 - options 192
 - plotting 191
- Internet Browser window** 180

L

- library files** *See file types, probe information files.*
- LIMS**
 - databases 48
 - filters 67
 - overview 48

- workflow monitor 49
- LIMS mode** 64, 359

M

- mapping report** 280
 - printing 372
- masks**
 - See probe masks.* 431
- measured image view** 127
- measured images**
 - expression probe array 172
 - HuSNP probe array 274
- metrics**
 - See expression analysis metrics.*
- metrics table** 162
 - exporting 163–165
- Microarray Suite**
 - installing 29
- Microarray Suite files**
 - experiment data files 67, 394
 - fluidics protocol files 67, 394
 - probe information (library) files 67

Microarray Suite software

- configuring 64
- disk files mode 64, 358
- LIMS mode 64, 359
- overview 43
- starting 53
- walkthrough 43–47

mixture detection algorithm

- 210
- settings 493–499

mutation & polymorphism analysis

- analysis 212
 - amino acids, *See amino acid data (SAW).*
 - HIV PRT Plus intensity graph 228
 - reports 337–342

N

- NetAffx** 167

- normalization** 422–430
 - all probe sets 423
 - selected probe sets 424
 - user defined 422

normalization mask file

- creating 425
- editing 429
- selecting 427

Nucleotide Analysis Window (NAW) 259

- find feature 264
- hiding data table rows 262
- locating a polymorphism 261
- printing 370
- removing an experiment 263
- sort order 263
- toolbar 522
- view mutations only 260

nucleotide data (SAW)

- comparing 219–222
- finding 235
- going to 235
- hiding 225
- viewing 217

O**online documentation** 25

- outliers** 141
 - masking 142
 - unmasking 142
 - viewing 142

P

- p53 report** 242–245, 342–344
 - generating 242
 - printing 372

p53 score graph

- plotting 233

pivot table 159

- display options 166
- exporting 168

- pixel coordinates** 124, 353

polymorphism

- locating in the NAW 261

polymorphism and mutation**analysis**

- consensus rules algorithm 212
- editing consensus base calls 240
- mixture detection algorithm 210

priming 99**printing**

- analysis output files
 - expression analysis 368
 - genotype analysis 370
 - hybridization analysis 371
 - mutation & polymorphism analysis 369
- cell intensity file 367
- image data file 367
- reports 372

probe array

- highlighting tiles 138
- scanning 107–110
- tiling information 135

probe cells

- calculating average intensity 121
- cell image view 128
- data 129–134
- difference image view 128
- measured image view 127
- viewing data 136

probe information (library) files

67

probe mask

- cross hybridization 459
- hybridization 470
- spike 481

probe masks 431–438

- creating 432
- editing 436
- selecting 434

probe set

- descriptions 167
- obtaining information 167

probe set information

- finding (EAW) 169
- hiding (EAW) 171
- reordering or removing (EAW) 170
- sorting (EAW) 169

publish database

- security 23

publish database, selecting 317**publishing**

- canceling a task 323
- intensity data 322
- overwriting data 322
- publishing a task 323
- restarting a task 323
- specifying a task 319

R**relative allele signal (RAS)** 23,

- 282, 286, 345, 346, 509
- change (delta RAS) 286, 346, 509

reports

- CYP450 report 277
- expression analysis 195–201, 328–334
- GT viewer 282, 345
- HIV PRT Plus report 247, 340
- mapping report 280, 335
- p53 report 242, 342
- printing 372

S**sample history** 85

- file view 85
- process view 86

sample registration 76**scale factor** 413**scale factor mask file**

- creating 416
- editing 420
- selecting 418

scaling 412–421

- all probe sets 414
- selected probe sets 415
- user defined 413

Scanner

- scanning a probe array ??–110

scanner

- options 109
- scanning a probe array 107–??
- shutting down 111
- stopping a scan 110

scanning 107–110**scatter graph**

- magnifying 178
- options 182
- plotting 176
- selecting points 181

security 23**Sequence Analysis Window (SAW)** 214–227

- alignment bars 227
- comparing
 - amino acid data 222–225
 - comparing nucleotide data 219–222
- display options 238
- editing consensus base calls 240
- expanded information 215
- finding a nucleotide or amino acid 235
- going to a nucleotide or amino acid 235
- hiding tiles, bases, or rows 225
- importing files 241
- printing 369
- removing data, tiles, or rows 227
- re-ordering probe array type or experiment 226
- toolbar 521
- viewing nucleotides or amino acids 217
- viewing the complement of a tile number 218

series graph 186–191

- options 189

plotting 188

shortcut

keys, See hot keys.

shortcut bar 54

SNP, See HuSNP

sorting

probe set information 169

spike probe mask 481

combining masks 487

creating 482

editing 485

starting

Microarray Suite software
53

statistical expression 22

**statistical expression algo-
rithm** 22, 199–201, 332–334

status log 63

T

task bar 117

technical support 26

tiles

highlighting 138

tiling information 135

toolbars

Batch Analysis Window 525

Expression Analysis Win-
dow 519

GT Viewer 523

Hybridization Analysis Win-
dow 524

main 517

Nucleotide Analysis Win-
dow 522

publish 527

report 526

Sequence Analysis Window
521

Hybridization Analysis Win-
dow (HAW) 291

Image window 115

Internet Browser 180

Nucleotide Analysis Win-
dow (NAW) 259

Sequence Analysis Window
(SAW) 214–227

workflow monitor 49, 89

cell analysis 92

grid alignment 91

hybridization 91

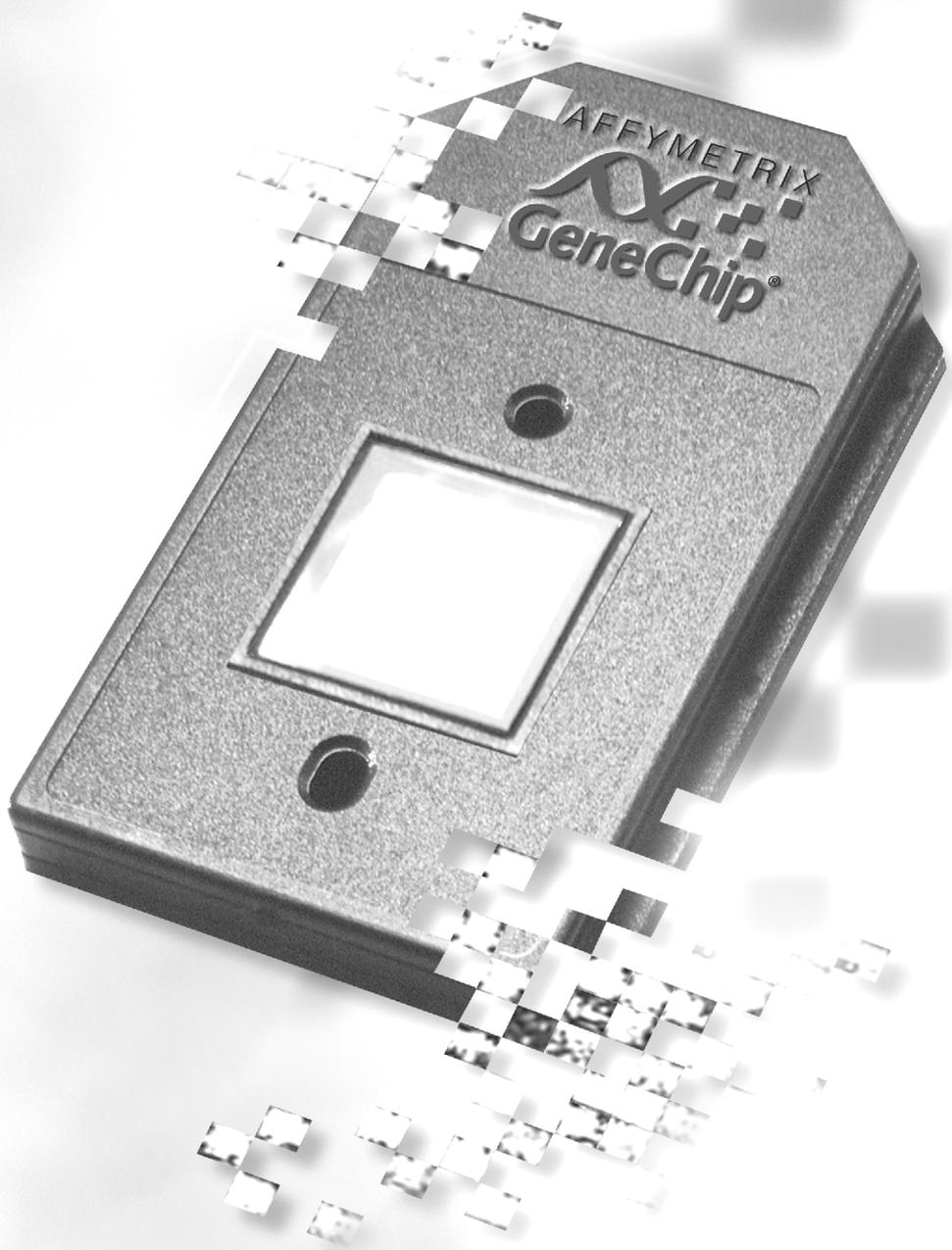
probe array analysis 93

scan 91

W

windows

Expression Analysis Win-
dow (EAW) 159



AFFYMETRIX
GeneChip®