



Agilent CytoGenomics 5.3

Feature Extraction for CytoGenomics

User Guide

Revision A0, November 2022

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

Notices

© Agilent Technologies, Inc. 2022

No part of this manual may be reproduced in any form or by any means (including electronic storage and retrieval or translation into a foreign language) without prior agreement and written consent from Agilent Technologies, Inc. as governed by United States and international copyright laws.

Manual Part Number

G1662-90072

Edition

Revision A0, November 2022

Agilent Technologies, Inc.
5301 Stevens Creek Blvd.
Santa Clara, CA 95051

Technical Support

For US and Canada

Call 800-227-9770 (option 3, 5, 3)
Or send an email to
informatics_support@agilent.com

For all other regions

Agilent's world-wide Sales and Support Center contact details for your location can be obtained at
www.agilent.com/en/contact-us/page.

Warranty

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

Technology Licenses

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

Restricted Rights Legend

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

Safety Notices

CAUTION

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

WARNING

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

In This Guide...

This *User Guide* shows you how to set up and run the Feature Extraction for CytoGenomics program automatically for a batch of image files and how to extract image files in real time.

1 Working with Feature Extraction

This chapter introduces the Feature Extraction analysis process and the tasks you can perform with the software to help you with your microarray experiments.

2 Extracting Microarrays Automatically

This chapter provides instructions on how to set up and run Feature Extraction on a batch of existing Agilent and/or non-Agilent image files. It also shows you how to set up and run Feature Extraction on Agilent image files as they are saved from the scanner in real time.

3 Creating Grid Files and Templates

Automatic extraction can take place because grid templates and grid files can be assigned to an image file. This chapter shows you how to create grid files and templates.

4 Changing Protocol Settings

Another reason why automatic extraction can take place is because a set of Feature Extraction parameters can be assigned to an image file as a protocol. This chapter gives instructions on how to change the parameter values for each step of the Feature Extraction process.

5 Changing Image Displays

Learn how to work with image displays in this chapter. This chapter also includes instructions on working with the frames and panes of the interface display.

Acknowledgments

Apache acknowledgment

Part of this software is based on the Xerces XML parser, Copyright (c) 1999-2000 The Apache Software Foundation. All Rights Reserved (www.apache.org).

JPEG acknowledgment

This software is based in part on the work of the Independent JPEG Group. Copyright (c) 1991-1998, Thomas G. Lane. All Rights Reserved.

Loess/Netlib acknowledgment

Part of this software is based on a Loess/Lowess algorithm and implementation. The authors of Loess/Lowess are Cleveland, Grosse and Shyu. Copyright (c) 1989, 1992 by AT&T. Permission to use, copy, modify and distribute this software for any purpose without fee is hereby granted, provided that this entire notice is included in all copies of any software that is or includes a copy or modification of this software and in all copies of the supporting documentation for such software.

THIS SOFTWARE IS BEING PROVIDED "AS IS", WITHOUT ANY EXPRESS OR IMPLIED WARRANTY. NEITHER THE AUTHORS NOR AT&T MAKE ANY REPRESENTATION OR WARRANTY OF ANY KIND CONCERNING THE MERCHANTABILITY OF THIS SOFTWARE OR ITS FITNESS FOR ANY PARTICULAR PURPOSE.

Stanford University School of Medicine acknowledgment

Non-Agilent microarray image courtesy of Dr. Roger Wagner, Division of Cardiovascular Medicine, Stanford University School of Medicine

This software contains material that is Copyright (c) 1994-1999 DUNDAS SOFTWARE LTD., All Rights Reserved.

LibTiff acknowledgement

Part of this software is based upon LibTIFF version 3.8.0.

Copyright (c) 1988-1997 Sam Leffler

Copyright (c) 1991-1997 Silicon Graphics, Inc.

Permission to use, copy, modify, distribute, and sell this software and its documentation for any purpose is hereby granted without fee, provided that (i) the above copyright notices and this permission notice appear in all copies of the software and related documentation, and (ii) the names of Sam Leffler and Silicon Graphics may not be used in any advertising or publicity relating to the software without the specific, prior written permission of Sam Leffler and Silicon Graphics.

THE SOFTWARE IS PROVIDED "AS-IS" AND WITHOUT WARRANTY OF ANY KIND, EXPRESS, IMPLIED OR OTHERWISE, INCLUDING WITHOUT LIMITATION, ANY WARRANTY OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

IN NO EVENT SHALL SAM LEFFLER OR SILICON GRAPHICS BE LIABLE FOR ANY SPECIAL, INCIDENTAL, INDIRECT OR CONSEQUENTIAL DAMAGES OF ANY KIND, OR ANY DAMAGES WHATSOEVER RESULTING FROM LOSS OF USE, DATA OR PROFITS, WHETHER OR NOT ADVISED OF THE POSSIBILITY OF DAMAGE, AND ON ANY THEORY OF LIABILITY, ARISING OUT OF OR IN CONNECTION WITH THE USE OR PERFORMANCE OF THIS SOFTWARE.

Content

1 Working with Feature Extraction for CytoGenomics

Feature Extraction Processes	14
Overview of Software Capabilities	17
Capabilities in the Feature Extraction for CytoGenomics software window	17
Feature Extraction for CytoGenomics algorithms for data analysis	18
Extraction results	19
What you can do to set up and run automatic batch extraction	20
Setup	20
Run extraction	20
What you can do with images before extraction	21
View image and change image display	21
Find grid and spots on the image manually	21
What you can do with protocol algorithms	22
What you can do with the protocol steps	22
What the protocol steps do for you	22
What you can do with .shp files of extracted images	24
What you can do with extraction results	25

2 Extracting Microarrays Automatically

Setting up and editing projects	28
To set up/edit a Standard Project to run existing image files	28
To set up/edit an On-Time project to run image files in real time	29
To create a project	30
To open a project	31
To save a project	31
Changing properties for individual projects	32
To display project properties for individual projects	32
To change general project property settings	33
To select the location for saving and/or sending output files	34
To change the FTP settings to send result files automatically to Rosetta Resolver	36
To set up to automatically assign protocols	36
To select to use a grid file	37
To select an external DyeNorm list file	37
To select to overwrite previous results	37
To select the incoming image folder for On-Time projects	37
To select a time to terminate real-time extractions	38

To turn off the option to automatically extract XDR images	38
Changing default project properties in Preferences	39
To open the Preferences dialog box	39
To display the General default project properties	39
To select to save the Project Run Summary Report	40
To select to overwrite previous results	40
To select a default folder for a new project	41
To display the default project properties for Output and Data Transfer	41
To select the default location for saving and/or sending output files	41
To change the default FTP settings for automatic data transfer	43
To change default run post-processing plug-ins	44
To display the default properties for standard projects	44
To change the default protocol and its location	44
To select to use a grid file	45
To change the Start Run Mode setting	45
To select an external DyeNorm list file	46
To select to create XDR extraction sets automatically	46
To display the on-time default properties	46
To select the default image folder	47
To select a default time to terminate real-time feature extraction	47
To select to automatically extract an XDR image pair	48
Working with extraction sets	49
What is an extraction set?	49
How does the software create an extraction set?	50
To add extraction sets (image files) to the standard project	50
To sort extraction sets	51
To remove an extraction set from the project	51
To display/edit the extraction Set Configuration table	52
Working with grid templates/files	54
Setting up and Changing Grid Templates and Files	55
To show or hide the Grid Template Browser	55
To add or change a grid template to an extraction set	55
To add or change a grid file to an extraction set	55
To add a grid template	55
To open grid mode from an extraction set	56
To display or change grid template properties	56
To add or change the default protocol in the grid template	57
To add or change the default one-color protocol in the grid template	57
To change the dye normalization gene list in the grid template	57
To create a new internal dye normalization gene list for a grid template	58
To edit an internal dye normalization list	61

To import external dye normalization list to create a new internal list	61
To export an internal dye normalization gene list	62
Working with protocols	63
To show or hide the FE Protocol Browser	63
To add a protocol to an extraction set	63
To change the protocol used in an extraction set	63
To display or change protocol properties	64
To import or export a protocol to or from the database	64
Working with QC metric sets	66
To show or hide the QC Metric Set Browser	66
To associate a QC metric set with a protocol	66
To import or export a QC Metric Set to or from the database	66
To view metric set properties	67
Running Feature Extraction	68
To start feature extraction	68
To stop feature extraction	68
To create a new project with remaining extraction sets if run fails	68
Displaying results	69
To display the text result file in Microsoft Excel	69
To display the Project Run Summary Report	70
To display the QC Report	72
To print the QC Report	73
To display visual results	73

3 Creating Grid Files

About grid templates and grid files	78
When to use manual gridding	79
Analyzing the initial grid for an Agilent image	80
To open the current extraction in grid mode	80
If automatic gridding for an Agilent image fails	82
Step 1. Create a new project for failed extraction sets	82
Step 1b. Manually create a new project for failed extraction sets	83
Step 2. Access grid mode from a project extraction set	83
Step 3. Position grids and save grid file	84
Step 4. Set up and rerun Feature Extraction	84
Creating and using a grid file for non-Agilent images overview	85
Step 1. Start grid mode from the non-Agilent image file	85
Step 2. Position grids, locate spots and save grid file	85
Step 3. Extract the image file with its new grid file	86

Starting grid mode for a non-Agilent image	87
To open an image file	87
To change the Portrait/Landscape view of the image	88
To open grid mode for a non-Agilent image	88
Viewing and changing grid geometry	93
Optimizing grid placement using the help wizard	95
To zoom in or out of the views	95
To undo or redo grid position movements	96
Step 1. Get nominal spot spacing manually	96
Step 2. Get nominal subgrid spacing manually	97
Step 3. Adjust the position of the main grid	98
Step 4. Adjust the subgrid	99
Step 5. Adjusting rotation and skew of subgrids	100
Step 6. Save grid file	100
Manual grid adjustments	101
To rotate the position of the main grid	101
To move or rotate the main grid with Grid Adjustment pane	102
To rotate the subgrid	102
To skew the subgrid corner	103
To position subgrids with Grid Adjustment pane	104
Saving a grid file	105
Calculating spot size and spot centroids	106
Locating spot centroids manually	107
To find a spot to adjust	107
To position the spot centroid	109
To ignore spots	109
To undo or redo spot position movements	110
To position spot centroids with the Grid Adjustment pane	110
Refitting a grid	112
Changing default settings for Grid Editing	114
Shortcuts to help you work with grids and find spots	115

4 Changing Protocol Settings

Changing protocol settings	118
Protocol templates	118
To create a new protocol from a protocol template	118
Creating a new protocol from an existing protocol	119
To open a protocol	119
To view or change the protocol properties	120

To change the protocol steps to run	121
To change the settings for each protocol step	123
To save a protocol with a new name	123
Changing Protocol Step 1: Place Grid	124
To access the settings to place a grid	124
To select the microarray format	124
To select a method for placing the grid	126
To select to use enhanced gridding	126
To select to enable background peak shifting	126
To select to use central part of pack for slope and skew calculations	127
To select to use the correlation method to obtain origin x of subgrids	127
Changing Protocol Step 2: Optimize Grid Fit	128
To access the settings to optimize the grid fit	128
To select the grid format	128
To change the settings to iteratively adjust corners when grid format is selected	129
Changing Protocol Step 3: Find Spots	130
To access the settings to find spots when spot format is automatically determined	130
To select the statistical method to reject pixel outliers	131
To select whether to use mean or median for data analysis	131
To enable the Enhance Spot Finding option	131
To access the settings to find spots when spot format is selected	132
To change values to find and size the spots.	132
To select a spot statistics method to define features	133
To select to autoestimate the local radius to define local background	134
Final calculations for protocol step to find and measure spots	135
Propagation of Find Spot errors to other algorithms	135
Changing Protocol Step 4: Flag Outliers	136
To access the settings to flag outliers	136
To change settings to flag population outliers	137
To change settings to flag non-uniform outliers	138
Changing Protocol Step 5: Compute Bkgd, Bias and Error	140
To access the settings to compute background, bias and error	140
To select the background subtraction method	141
To choose to remove the surface trend found in the data	142
To choose to adjust the background globally (2-color)	144
To choose to perform multiplicative detrending	144
To choose to use robust negative control statistics	145
To select the method to estimate the error (Error Model)	145
To choose to use or not use surrogates	147
Propagation of errors from the Compute Bkgd, Bias and Error algorithm to other algorithms	147

To change settings to test the significance of a feature	147
Changing Protocol Step 6: Correct Dye Biases	149
To access the settings to correct dye biases	149
To choose how to select dye normalization list	150
To choose a method to select a set of dye normalization probes	150
To select to omit background population outliers	151
To select to allow positive and negative controls	151
To select signal characteristics	151
To select a normalization correction method	151
Changing Protocol Step 7: Compute Ratios	153
To access the settings to compute ratios	153
To enter the value for the log ratio limit	154
Use of surrogates for log ratio and p-value calculations	154
Changing Protocol Step 8: Calculate Metrics	157
To access the settings to calculate metrics	157
To select to use spikein targets	157
To change the minimum population for replicate statistics	158
To select the grid test format	158
To change the p-value for differential expression	158
To change the percentile value	158
Changing Protocol Step 9: Generate Results	159
To access the settings to generate results	159
To select the type of QC Report	159
To select to generate a single file for the text output	160
To change the compression factor for the JPEG output	160

5 Changing Image Displays

Displaying and saving image files	162
To open an image file	162
To crop an image	163
To save a cropped image to a new file	164
To add or change the New Array Identifier	166
Loading Feature Extraction visual results	167
To view visual results for low-resolution images:	167
To view visual results for high-resolution images (2- by 3-micron or 20-bit):	168
To save an image (original or cropped) for import to another program	169
To display file information	170
Changing image display settings	171
To change image orientation (landscape/portrait)	171
To change the channels that display	172

To change the display color	173
To change between log and linear color scale	175
To change the intensity scale (color display range – CDR)	175
To change the magnification (zooming)	176
To change the window size to fit the image	178
To change the extraction results display	178
To change and save histograms and line graphs	179
To change default image viewing settings	182
To change the print settings	183
Hot Key actions for image displays	184
Changing Feature Extraction window displays	185
To move windows and panes	186
To select the layout to use when software starts	186

1 Working with Feature Extraction for CytoGenomics

Feature Extraction Processes	14
Overview of Software Capabilities	17
What you can do to set up and run automatic batch extraction	20
What you can do with images before extraction	21
What you can do with protocol algorithms	22
What you can do with .shp files of extracted images	24
What you can do with extraction results	25

This chapter gives you an introduction to the process of Feature Extraction for CytoGenomics and lets you know what you can do with the software to optimize the process. It is not necessary to run the Feature Extraction software separately in order to do feature extractions as part of Agilent CytoGenomics workflows.

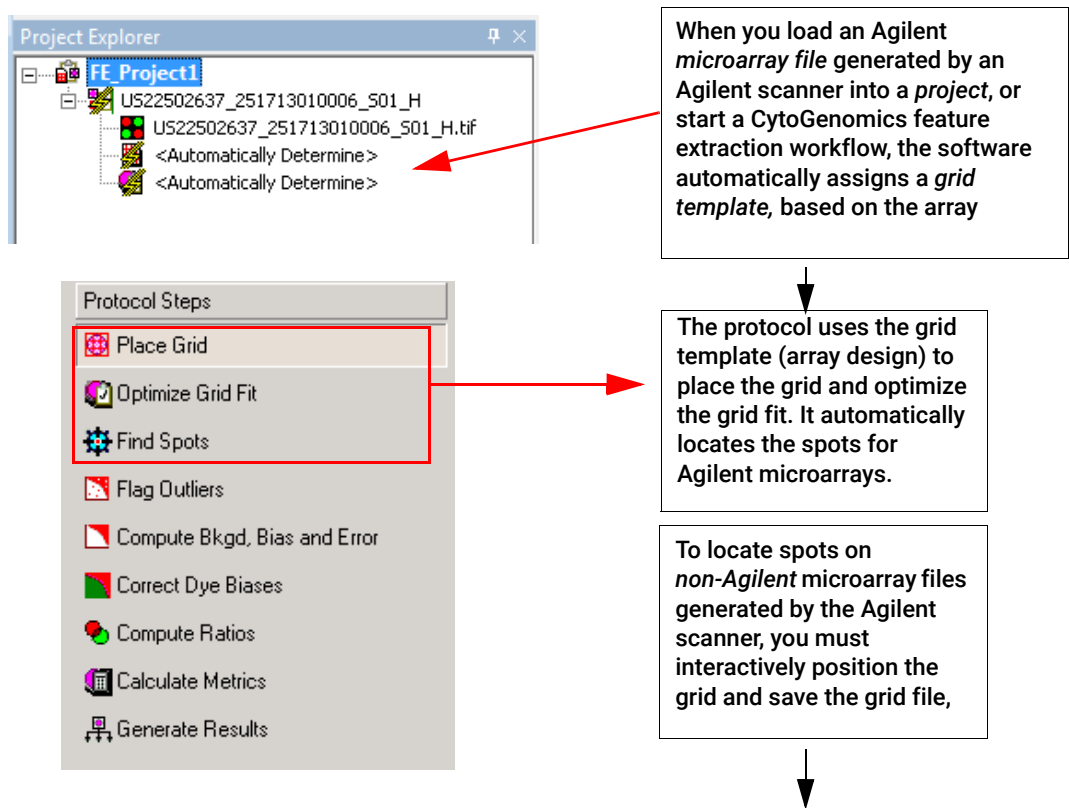
NOTE

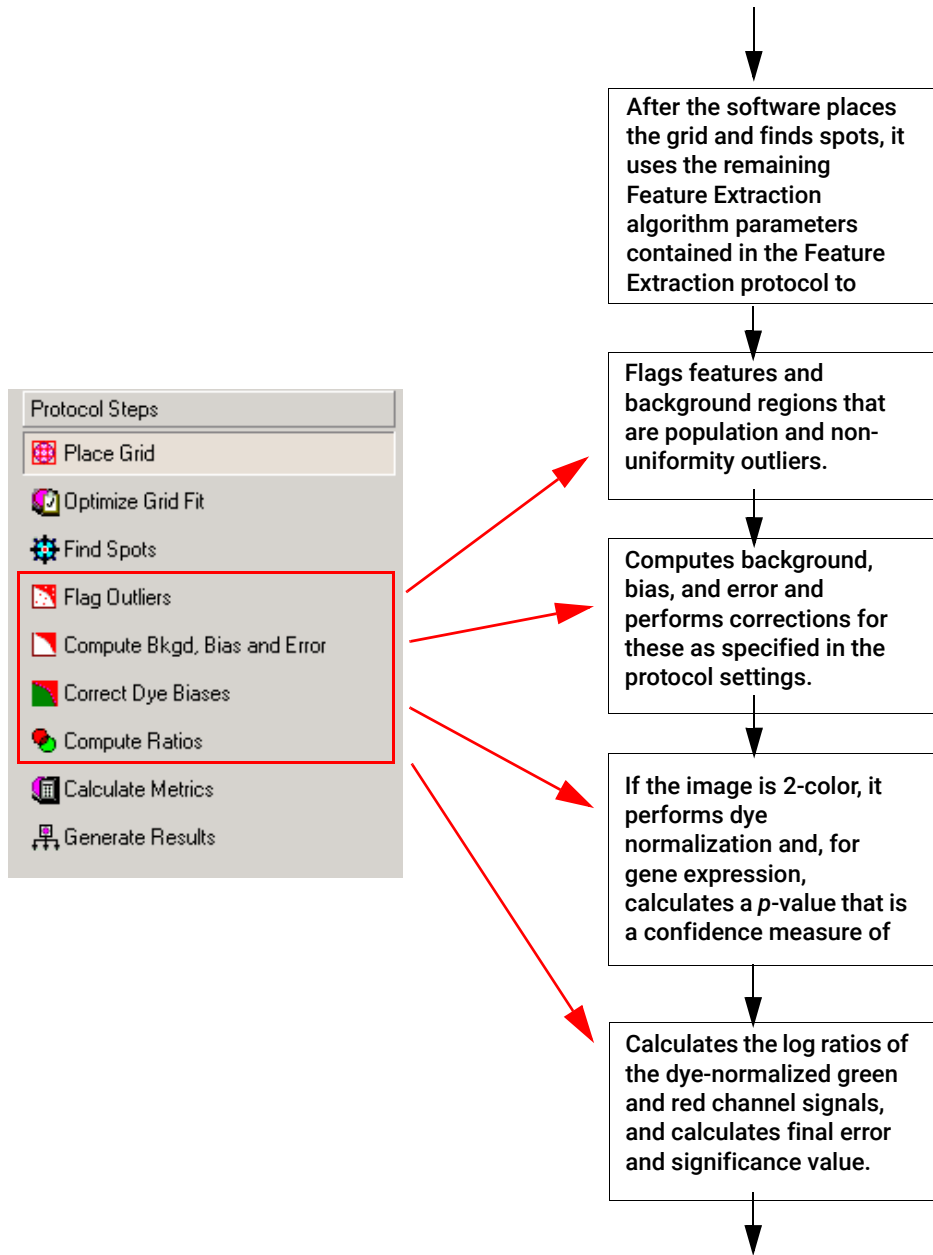
Note that Feature Extraction for CytoGenomics is not supported on Macintosh systems and is not installed with the Macintosh version of the Agilent CytoGenomics software. To run an analysis workflow, you must first extract the image file using the standalone Agilent Feature Extraction program (on a Windows PC), then use the extracted FE file to run a manual workflow in the Macintosh version of CytoGenomics.

Feature Extraction Processes

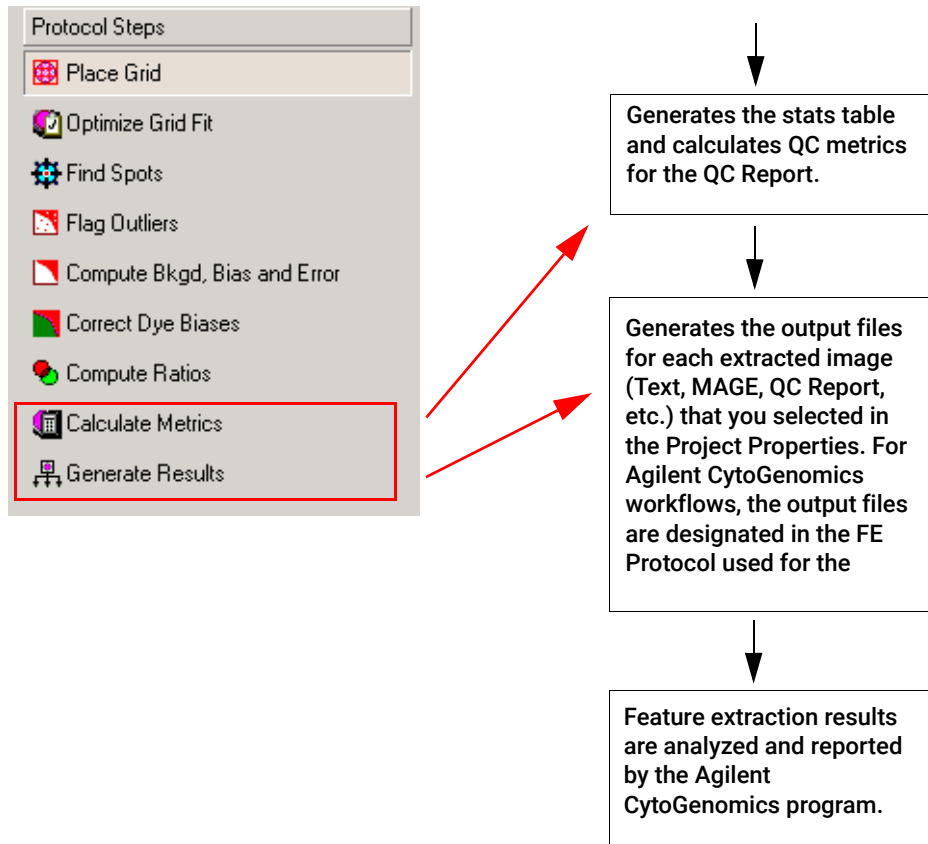
Feature Extraction consists of two major processes: image analysis to place the grid and locate spots and data analysis to define and measure spot features for gene expression.

Below is a flow chart that shows the process for producing reliable microarray results using Feature Extraction for CytoGenomics software on Agilent or non-Agilent microarrays scanned on the Agilent Scanner. The data analysis flow starts with removing outlier pixels and ends with transfer of Feature Extraction results into third-party data analysis programs to compare the results of different experiments. This flow chart shows the protocol steps for Feature Extraction using a CGH protocol.





Feature Extraction Processes



Overview of Software Capabilities

Capabilities in the Feature Extraction for CytoGenomics software window

Automatic batch extraction

You create projects to contain extraction sets. The extraction set includes the Agilent or non-Agilent image file, a grid template or grid file to locate spots and a protocol to specify the values for the parameters to carry out Feature Extraction data analysis.

A list of supported scans and array formats are listed in the following table. (Note that GenetiSure Pre-Screen microarray images generated by non-Agilent scanners for single cell applications is not supported by Agilent CytoGenomics.)

Table 1 Supported scans and array formats

Array	Format	Scanner	Scan type	Resolution	Support	Project run summary message
Agilent	65um	Agilent	16 bit	5um	Supported and tested	
Agilent	65um	Agilent	20 bit	5um	Supported and tested	
Agilent	30um	Agilent	16 bit	3um	Supported and tested	
Agilent	30um	Agilent	20 bit	3um	Supported and tested	
Agilent	65um	Agilent	16 bit	3um	Supported but not tested	Information message
Agilent	65um	Agilent	16 bit	2um	Supported but not tested	Information message
Agilent	65um	Agilent	20 bit	3um	Supported but not tested	Information message
Agilent	65um	Agilent	20 bit	2um	Supported but not tested	Information message
Agilent	30um	Agilent	16 bit	2um	Supported but not tested	Information message
Agilent	30um	Agilent	20 bit	2um	Supported but not tested	Information message
Agilent	65um	Agilent	16 bit	10um	Not Supported	Warning then proceed
Agilent	65um	Agilent	20 bit	10um	Not Supported	Warning then proceed
Agilent	30um	Agilent	16 bit	10um	Not Supported	Warning then proceed
Agilent	30um	Agilent	16 bit	5um	Not Supported	Warning then proceed
Agilent	30um	Agilent	20 bit	10um	Not Supported	Warning then proceed
Agilent	30um	Agilent	20 bit	5um	Not Supported	Warning then proceed
non-Agilent	N/A	Agilent	any	any	Limited Support	Information message
Agilent	30um	non-Agilent	2-color	2um	Select Support	Information message
Agilent	65um	non-Agilent	2-color	5um	Select Support	Information message

Explanation of support types “Not Supported” means that no support is provided by Agilent Technologies. “Limited Support” means that Agilent provides support on a best-effort basis to help you scan microarrays and extract images, however there is no guarantee or commitment. “Select Support” is provided for Agilent arrays scanned on selected non-Agilent scanners (specifically, the Roche NimbleGen MS 200 scanners, Molecular Devices GenePix scanners, select Innopsys InnoScan scanners, and select Tecan scanners.) Note that when scanning Agilent arrays on a non-Agilent scanner, you may need to manually grid the arrays if auto-gridding fails. See **“If automatic gridding for an Agilent image fails”** on page 82.

CAUTION

Although Feature Extraction for CytoGenomics 5.1 is capable of processing images of select Agilent arrays scanned on select NimbleGen, Molecular Devices, Innopsys, and Tecan scanners, Agilent can only guarantee the quality of the data output when processing images of Agilent arrays scanned on an Agilent scanner.

When you start an extraction, the software automatically extracts all the sets in the project and produces a project summary report and a QC report for each image. You can also set up a project to automatically extract Agilent microarray images in real time as soon as the files are made available by the Agilent scanner.

Image analysis

Before extraction you can examine the spots on your microarrays visually and compute statistics about the image, using the histogram and line plots provided in the software. You can interactively position grids to find the spots on the microarray and calculate the spot centroids and sizes.

Feature Extraction for CytoGenomics can handle high-resolution images (2- or 3-micron and 20-bit) from the Agilent G2565CA Scanner. To avoid using large amounts of memory to handle the high-resolution images, Feature Extraction for CytoGenomics loads a low-resolution preview image for first viewing.

When you create a cropped image to one-tenth of its original size or less, the program loads the original high-resolution image for the cropped image.

Histogram and line plots are always based on the original high-resolution image, not the preview one.

Feature Extraction for CytoGenomics algorithms for data analysis

For all images, the protocols define features and background regions, remove pixels, and flag features and background regions that may affect the reliability of the results. In addition, protocols set up the extraction to subtract background information from features and make a background adjustment on signals of low intensity.

The software then calculates a reliable log ratio, p -value, and log ratio error for each feature to give you a confidence measure in the measured log ratio.

Extraction results

Feature Extraction for CytoGenomics offers several types of outputs for each microarray extraction:

- Project Run Summary report
- MAGE (for Rosetta Resolver)
- JPEG (for Rosetta Resolver)
- Text (normal and .zip format)
- Shape
- Grid (used in Feature Extraction only)
- QC Report

You can export the text results containing all the parameters, statistical calculations and feature results into a spreadsheet.

To compare the results of different microarray experiments, you can also import various output types into GeneSpring, Agilent Genomic Workbench software, and third-party analysis programs. The text results can be saved in full, compact, QC, and minimal formats, and MAGE-ML results can be saved in both full and compact formats. For details on what information is included in these files, see the *Feature Extraction for CytoGenomics Reference Guide*.

You can load and display the results to see the outcome of the statistical and flagging algorithms for each feature and background.

What you can do to set up and run automatic batch extraction

Setup

See the Feature Extraction for CytoGenomics Quick Start Guide for definitions of project, extraction set, grid template and protocol.

You can perform the following tasks when you set up automatic batch extraction from within the Feature Extraction for CytoGenomics program:

- Create projects and save them with new names.
- Load image files to create extraction sets. For Agilent image files, the software automatically:
 - Loads the required grid template (with associated DyeNorm List, if available) and associated default protocol for each extraction set, if these are available in the database.
- Drag and drop scan files onto the project.
- Type or change the protocol or grid template for the project.
- Drag and drop grid templates or protocols onto each extraction set.
- Assign a grid file to an extraction set.
- Choose the type of result file to save.
- Set up to transfer the files via FTP after extraction.
- Delete extraction sets from the project. Add protocols, and QC metric sets to the database.
- Edit protocols and grid template properties.
- Change the order that extraction sets are processed.
- Choose to automatically stop Feature Extraction for CytoGenomics after a specified time when extraction is in real-time mode (On-Time Project).
- Load image files for supported types.

Run extraction

- Start and stop extracting.
- Monitor the extraction progress for each extraction.
- View a Summary Report of the extraction set (or batch of extraction sets).

What you can do with images before extraction

View image and change image display

For details on how to view and change image display, see Chapter 5, "Changing Image Displays".

- Change color and/or scale of the image for easier viewing.
- Select rectangular and elliptical regions of interest.
- Create histograms and line graphs.
- Create images for other programs.
- Print the images.
- Compute statistics about the image
 - Line plots
 - Histogram plots
- Type and save barcode information for Agilent images where the barcode reader has failed.
- Flip and rotate the image from landscape to portrait mode and vice versa, for scans of third-party arrays (cannot do for high-resolution images).

Find grid and spots on the image manually

For details on manual gridding, see Chapter 3, "Creating Grid Files".

For Agilent images, grid templates are applied automatically. You would not normally need to adjust the grid unless an error indicating a gridding failure appears in the Summary Report for an extraction. For non-Agilent images, you must create a grid file before you can extract the image.

- Interactively adjust the grids on the microarray scan.
- Interactively adjust the spots on the microarray (for third-party microarray images only).
- Calculate the spot size and compare the position of the nominal spot centroid (+) laid down by the grid with the found centroid position (X) for the spot (for third-party microarray images only).
- Move the found centroid position (X) to where you want it on the spot (for third-party microarray images only).
 - Feature Extraction for CytoGenomics uses these manually adjusted centroid positions to define features rather than those calculated by the Place Grid algorithm.
- Find features by typing feature number, gene name, or probe name (for third-party microarray images only). For Agilent arrays, if a shape file is loaded, you can find features by feature number.
- In Grid Mode, select spots to be ignored from analysis (for third-party microarray images only).

What you can do with protocol algorithms

Protocols contain the values for the parameters used by Feature Extraction for CytoGenomics algorithms to calculate intermediate and final results. See **“Feature Extraction Processes”** on page 14.

For details on protocols, see **Chapter 4**, “Changing Protocol Settings” in this User Guide and the Feature Extraction for CytoGenomics Reference Guide.

What you can do with the protocol steps

- Select the protocol steps that you want to use.
- Change the default settings for each algorithm in the protocols.
- Reload default settings after you have changed them.

What the protocol steps do for you

These protocol algorithms operate in the following order.

- Place Grid**
 - For Agilent microarrays using a grid template, automatically places a grid on the microarray.
 - For any microarray loaded with a grid file that you saved during Grid Mode (i.e. interactive gridding), automatically calculates centroid positions. Any spots that were manually adjusted are left unchanged.
 - For any microarray, provides two methods of auto-gridding. The default method is to “allow some distortion”. This lets the grid have a feature spacing somewhat different from the spacing found in the grid template. The second choice is “place and rotate only”. This assumes that the feature spacing from the grid template is accurate and only lets this grid move to the best location.
 - For any microarray, user can set the Enable Background Peak Shift flag in order to get better gridding. Some variations in the results are expected with and without use of this flag as the grid positions obtained differ.
 - For any microarray, user can select to use regions near to the center of packs for slope and skew calculations. This option is of use in cases where pack edges have dim spots and are failing to grid. When this is selected, the user can select to have the software to perform one extra step of correlation following the projection data analysis to get the origin.
- Optimize Grid Fit**
 - For microarrays with 30- or 65-micron feature size, iteratively adjusts the grid by looking at spots in the corners of the microarray to improve overall image registration.
- Find Spots**
 - Finds the center of each located spot and encircles the pixels belonging to the feature using one of two selected methods: CookieCutter or WholeSpot.
 - Calculates statistics for the pixel populations of the features and backgrounds, removing outlier pixels.
 - Calculates signal statistics for inlier pixels.
- Flag Outliers**
 - Flags features and background regions that are population and non-uniformity outliers.

What the protocol steps do for you

Compute Bkgd, Bias and Error

- If you select to subtract the background, subtracts background from the raw signal of each feature using a local or global background method that you select.
- Corrects for spatial gradient across each microarray and any dome effects.
- Adjusts the subtracted signal for inaccuracies in the noise level estimation.
- For all data, also calculates the error based on a selected error model.
- Determines if the feature signal is greater and significantly different from background using the error model calculation.
- Can calculate signals and background for low-intensity features using surrogates.

Correct Dye Biases

- Automatically determines if a dye normalization gene list exists for the grid template.
- Selects to use dye normalization list or generates one internally.
- Selects the features for dye normalization based on a feature selection method of your choosing.
- Performs dye normalization on the normalization feature set with a curve fit of your choosing. (SNP probes are not included in dye normalization calculations.)

Compute Ratios

- Calculates the log ratio of the dye-normalized green and red channel signals.
- Calculates a final error of the log ratio and a significance value (p -value) based on an error model that you selected in the Compute background step.

Calculate Metrics

- Generates the stats table and calculates QC metrics for the QC Report.
You must leave this option turned on to generate a complete stats table and QC report. Some values will be missing if this option is turned off.

See the *Feature Extraction for CytoGenomics Reference Guide* for details on QC metrics.

Generate Results

- Generates for each extracted image the output files (Text, MAGE, QC Report, etc.) that you selected in the Project Properties.

See the *Feature Extraction for CytoGenomics Reference Guide* to learn more about this text file.

What you can do with .shp files of extracted images

- Show and hide feature information on the image:
 - Feature number
 - Gene name
 - Raw intensities per color channel
 - Log ratio ($rProcessedSignal/gProcessedSignal$) for 2-color microarray images
- Show or hide outlier features on the image.
- Search and display one feature at a time using feature number.
- Get reference to annotation results from Help > Reference Guide.
- Print the images.
- View the pixels used in the calculation of a feature or set of features.

What you can do with extraction results

- View a QC report. You can use the QC report to evaluate if the grid was placed correctly (snapshot of four corners of the grid) and to show if the microarray scan data is reliable.
- View all the extraction parameters, statistical calculations and feature results created by the Feature Extraction for CytoGenomics software to know how reliable each feature result is.
- Export the text file containing the numbers described above to a spreadsheet or a database of your choice.
- Automatically export MAGE-ML files to Rosetta Resolver (FTP) or manually load the files into a database of your choice.

What you can do with extraction results

2 Extracting Microarrays Automatically

Setting up and editing projects	28
Changing properties for individual projects	32
Changing default project properties in Preferences	39
Working with extraction sets	49
Working with grid templates/files	54
Setting up and Changing Grid Templates and Files	55
Working with protocols	63
Working with QC metric sets	66
Running Feature Extraction	68
Displaying results	69

You set up to extract a microarray image through the use of *projects*. Projects let you automatically extract one or more image files either as:

- A batch of existing Agilent or non-Agilent files that you add to the project (Standard Project).
- One Agilent image at a time as scanning completes (On-time Project).

This chapter explains how to set up and edit projects, run feature extraction, and view the results.

For information on how to use the Auto-Processing function of Agilent CytoGenomics to automate Feature Extraction, see the Agilent CytoGenomics help system.

Setting up and editing projects

You set up to extract a microarray image through the use of *projects*. Projects let you automatically extract one or more image files either as:

- A batch of existing Agilent or non-Agilent files that you add to the project (Standard Project).
- One Agilent image at a time as scanning completes (On-time Project).

To set up/edit a Standard Project to run existing image files

You can run Feature Extraction for Agilent or non-Agilent microarrays automatically through the use of standard projects. To each standard project, you add one or more *extraction sets*, which consist of the *image file*, a *grid template or file* and a *protocol*. The protocol contains all the values and settings for the parameters used in each step of the Feature Extraction process.

To learn more about working with individual images and with the Feature Extraction software window, see [Chapter 5](#), “Changing Image Displays”.

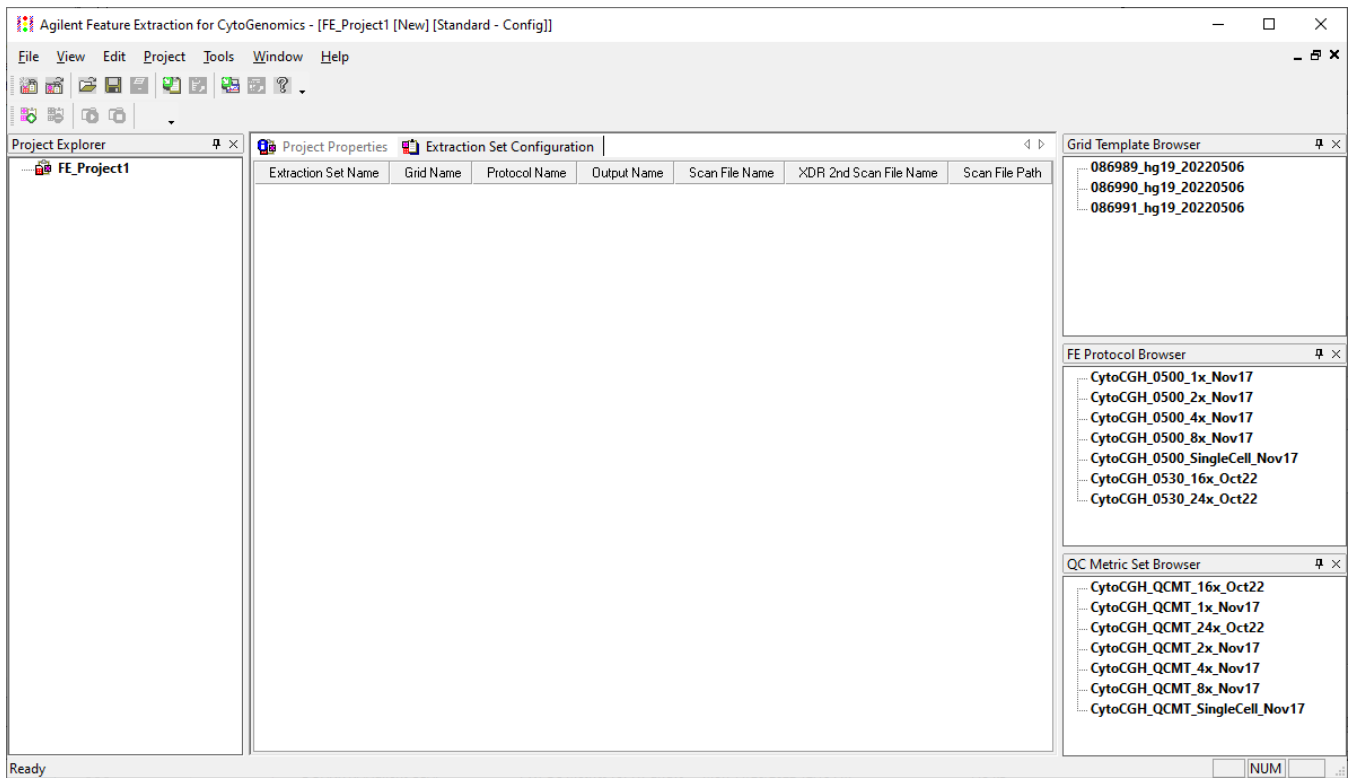


Figure 1 Feature Extraction Standard Project Window

To set up/edit an On-Time project to run image files in real time

*If you do not see the QC Metric Set Browser, click **View > QC Metric Set Browser**.*

- 1 Open a Standard Project, if one is not already open.
See **"Create a Standard project"** on page 30.
- 2 Change Project Properties.
See **"To display project properties for individual projects"** on page 32.
- 3 Add the image files to make up extraction sets.
See **"Working with extraction sets"** on page 49.
- 4 View the Extraction Set Configuration tab sheet to make certain all information is correct for each extraction set.
See **"To display/edit the extraction Set Configuration table"** on page 52.
- 5 For any extraction set with a non-Agilent microarray image, add an appropriate grid file.
See **"To add or change a grid file to an extraction set"** on page 55.
- 6 For any extraction set with a non-Agilent microarray image, add a protocol.
See **"To add a protocol to an extraction set"** on page 63.

Starting with Feature Extraction 10.5, protocols are linked to QC metric sets. A QC metric set is no longer assigned to the Project Properties tab. Instead, a QC metric set is assigned to a protocol. You can add or change the default QC metric set associated to the protocol.

See **"Working with QC metric sets"** on page 66.
- 7 Save the project.
See **"To save a project"** on page 31.

To set up/edit an On-Time project to run image files in real time

You can run Feature Extraction for Agilent microarrays in real time as image files are saved after scanning on the Agilent Scanner. The differences between a Standard project and an On-Time project are listed below:

- You can set up an On-Time project only for Agilent microarray images.
 - You do not add image files or extraction sets (image files, grid templates, and protocols) to the On-Time project.
 - You can type a maximum total processing time and a maximum time between extractions for an On-Time project.
- 1 Open an On-Time project.
See **"Create an On-Time project"** on page 30.
 - 2 Change Project Properties.
See **"To select a time to terminate real-time extractions"** on page 38.
 - 3 Save the project.
See **"To save a project"** on page 31.

To create a project

To create a project

Create a Standard project

To create a new standard project:

- Click **File > New > Standard Project**,
Or
Click the New project file icon.

When you create the next project after FE_Project1, the name is FE_Project2. The name of every project that you create after that follows in numerical order, whether you save FE_ProjectN or not or close FE_ProjectN or not.

Each time you restart the software, the first new project is FE_Project1, if that is the selection in the Preferences dialog box.

Create an On-Time project

- Click **File > New > On-Time Project**.

Notice that the On-Time Project, **Figure 2** does not have a Project Explorer and an Extraction Set Configuration tab.

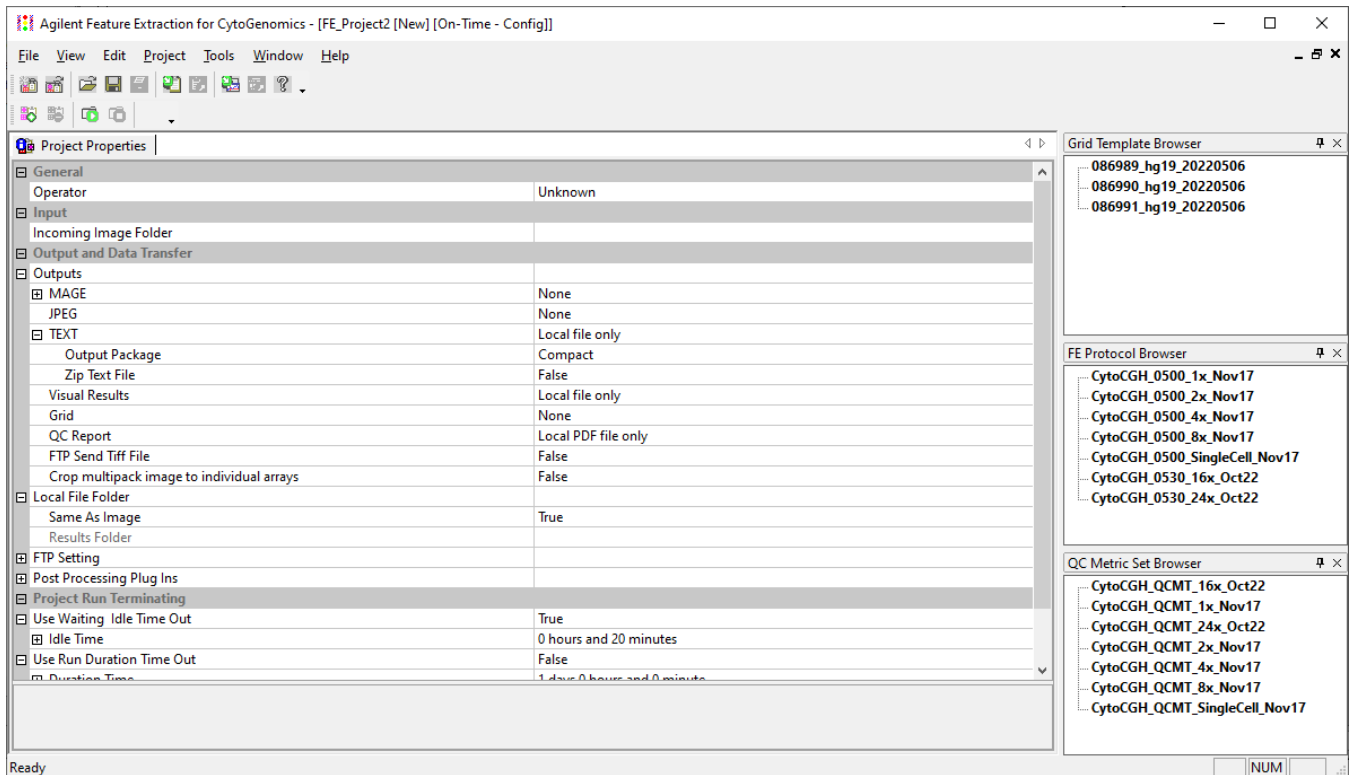


Figure 2 On-time project properties window

To save a project

To open a project

To open an existing Standard or On-Time project:

- 1 Click **File > Open > Project**,
Or
Click the Open project file icon.
- 2 Select the project name, and click **Open**.

To save a project

To save a Standard or On-Time project to a new name:

- 1 Click **File > Save As**.

The default name is FE_ProjectN, where N is the number of the currently loaded project.

- 2 Type the new name, then click **Save**.

If you are saving the project for the first time, you can click the Save icon and type any name you choose. The default name that appears is FE_ProjectN, where N is the number of the currently loaded project.

The default directory for projects is
C:\ProgramData\Agilent\FeatureExtractionCyto\FEProjects

To save a project to its current name:

- Click the Save icon.

If the current project has never been saved, the Save As dialog box opens.

Changing properties for individual projects

When you create a new project or open an existing project in the Feature Extraction for CytoGenomics software, you see a Project Properties tab. The default values found on this tab were set in the Preferences dialog box. For each project that you open, you can change these default properties.

If you are changing properties for every project that you run, you probably should change the default properties in the Preferences dialog box. See **“Changing default project properties in Preferences”** on page 39.

To display project properties for individual projects

Project Properties tab for Standard projects

- In the main window for Project setup, click the Project Properties tab.

The Project Properties tab for Standard projects with specified extraction sets looks like **Figure 3** below.

Project Properties Extraction Set Configuration	
General	
Operator	Unknown
Input	
Number of Extraction Sets Included	0
Output and Data Transfer	
Outputs	
MAGE	None
JPEG	None
TEXT	Local file only
Output Package	Compact
Zip Text File	False
Visual Results	Local file only
Grid	None
QC Report	Local PDF file only
FTP Send Tiff File	False
Crop multipack image to individual arrays	False
Local File Folder	
Same As Image	True
Results Folder	
FTP Setting	
Post Processing Plug Ins	
Automatic Protocol Assignment	
Highest Priority Default Protocol	Grid Template Default
Project Default Protocol	
Automatic Grid Template Assignment	
Use Grid file if available	False
External DyeNorm List File	
Overwrite Previous Results	False

Figure 3 Standard Project Properties tab, expanded

Project Properties tab for On-Time projects

When you create a new On-Time project or open an existing one, you see the Project Properties tab.

The Project Properties tab for On-Time projects looks like **Figure 4** below.

To change general project property settings

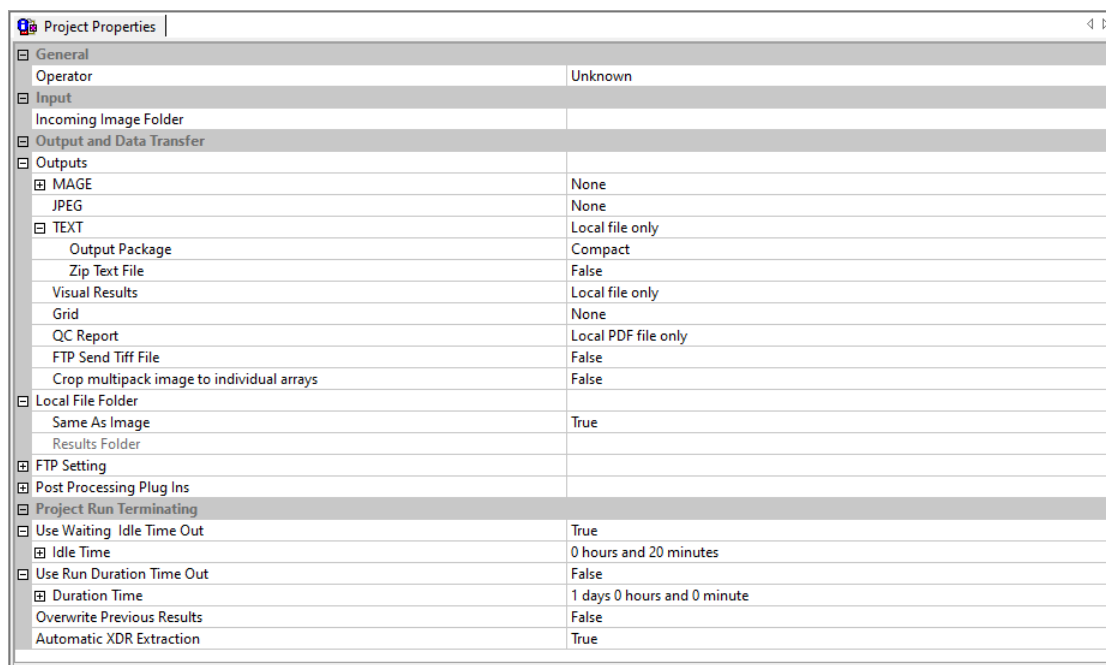


Figure 4 On-Time Project Properties tab, expanded

The parameters and settings in this tab are the same as those in a standard project with these differences:

- No capacity to select an available grid file.
- No setting for default protocol.
- No capacity to select an external DyeNorm list file.
- Capacity to set up a folder for images coming in from the scanner.
See **“To select the incoming image folder for On-Time projects”** on page 37.
- Capacity to terminate the run automatically.
See **“To select a time to terminate real-time extractions”** on page 38.
- Capacity to automatically extract XDR (extended dynamic range) images.
See **“To turn off the option to automatically extract XDR images”** on page 38.

To change general project property settings

Change to new values

- Double-click the cell to the right of the field name.
True changes to **False** and vice versa. For list selections, each double-click brings up a different choice in the list.
This does not work for check boxes or Browser cells.
- Or, click the arrow to the right of the cell, and select from the list.

To select the location for saving and/or sending output files

Reset changed values to default values

At any time, you can change a modified value back to the original default value.

- Right-click the cell on the right, and click **Reset to Default**.

Change the name of the operator

The only general property that you can change is the name of the Operator.

- Click the cell next to **Operator**, and type the name of the operator.

View the Number of Extraction Sets Included

You cannot change this field. If you add more extraction sets than are visible on screen, you can refer to this field to check the number that you have added.

To select the location for saving and/or sending output files

See the Feature Extraction for CytoGenomics Reference Guide for a listing of all the results that are generated in the output files.

This portion of the Project Properties tab lets you select the ultimate location(s) for the image and result files. See **"Select the location for output files"** on page 34 and **"Select the local file folder for saved results"** on page 35.

Select the location for output files

- Click the arrow to the right of the cell next to each of the result or image file types to select one of the following location(s).

Local file only

Select to save the files to a local file folder. See **"Select the local file folder for saved results"** on page 35.

FTP send only

Select to send the files to an external location. See **"To change the FTP settings to send result files automatically to Rosetta Resolver"** on page 36.

Both local file and FTP send

Select to save the files both to a local file folder and to an external location.

The following output files are generated and sent to the location(s) you specify:

MAGE result file

This file contains results in the MAGE-ML (XML) data transfer format, which lets you transfer the data into Rosetta Resolver and Array Express databases.

You can select to send a Full Output Package of MAGE-ML data or a Compact (reduced) set of data. And you can choose to send the data in a Compressed (ZIP) format, or not.

If MAGE output is selected, you also have the capacity to set and pass data ownership and permissions to Rosetta Resolver.

Expand the MAGE folder, and expand the Rosetta Resolver Data Access Setting folder to view the following fields:

If you do not type names for these settings, no access control information is entered into the MAGE-ML file.

Rosetta Resolver Data Access Settings: User Name Type the name of the user who will own the data transferred to the Rosetta Resolver database.

Rosetta Resolver Data Access Settings: Control Groups Type the names of the groups who have permission to use the data transferred to the Rosetta Resolver database. If there is more than one group, separate their names with a colon. For example, if you want the groups Public and Special to see the incoming extraction information, type `public:special`.

To select the location for saving and/or sending output files

JPEG image file To create a compressed version of the scan file, you can create a JPEG file. A JPEG file stores a likeness of the original image, which is used for visual reference. This option is for Rosetta Resolver use.

You can control the spatial compression of this JPEG output in the protocol. See **“To change the compression factor for the JPEG output”** on page 160 of this guide.

Text result file These results are in a tab-delimited text file that you can import into spreadsheets. The file contains input parameters, statistical calculations and feature results.

Compact mode contains all the columns needed for loading into GeneSpring, DNA Analytics, and Agilent Genomic Workbench. You can split this file into three files, each containing one of the sections. You can also import this file into GeneSpring data analysis software, either as a Full set of data, a Compact (reduced), QC, or Minimal. For a listing of the columns included in these formats, see Chapter 3 of the *Reference Guide*.

You can select to save the text result file in .zip format. The default for this selection is **False**.

Visual Results Visual results from the feature extraction process are saved to a .shp or “shape” file, which only the Feature Extraction for CytoGenomics program can read.

Grid This is the grid information saved to grid.csv and (for non-Agilent arrays) feat.csv files. It contains the output of all of the grid placement steps (e.g. the spot centroids and spot sizes).

QC Report A QC Report is saved to the results directory as a pdf file.

Send a Tiff file via FTP

Because Tiff files already are located in a local file folder, you can select to send them to an external location via FTP, or not.

- Double-click the cell to the right of FTP Send Tiff File, and select True or False.

Select the local file folder for saved results

Select to save results in the same local folder as the image This folder must be the folder that contains the images in the extraction sets.

- Under Local File Folder, double-click the cell next to Same As Image to select True, if necessary.

Select to save results in a local results folder This folder can be any folder that you choose.

- 1 Double-click the cell next to **Same As Image** to select **False**.
- 2 Click the cell next to **Results Folder**.
- 3 Click **Browse**.
- 4 Select the folder to contain the results.
- 5 Click **OK**.

To change the FTP settings to send result files automatically to Rosetta Resolver

If you use the same settings repeatedly, the values specific to your installation should be set up in the Preferences dialog box. See [“To change the default FTP settings for automatic data transfer”](#) on page 43.

- In the Project Properties tab, expand the FTP Setting folder, and type the values for these FTP settings:

Machine Destination	This field is specific to your Rosetta Resolver server name. Please contact your system administrator for the name to use.
Port	Type the number of the FTP port on the server. The default is port 21.
User name	Type the user name on the server that allows FTP access to the import directory of the server. The default User Name is resolverftp for Rosetta Resolver systems.
Password	This field already contains the default password for Rosetta Resolver servers. If it has been changed, contact your administrator for the correct password.
File Destination Folder	Scan information is sent to this folder. For MAGE results, the default setting is mage (/import/mage on the server). Confirm with your system administrator that the FTP account uses the import directory of the server as the default directory for the transfer.

To set up to automatically assign protocols

Select the highest priority location for the default protocol

The software uses the protocol specified at this parameter step to complete an extraction set if a protocol is available in both the grid template and project. For more information on extraction sets, grid templates and protocols, see [“Working with extraction sets”](#) on page 49.

- 1 Click the arrow to the right of the cell next to **Highest Priority Default Protocol**.
- 2 Select **Grid Template Default** or **Project Default**.

Grid Template Default	If selected, the protocol specified in the grid template is assigned to the extraction set. If the grid template contains no protocol, then the software assigns the project default protocol. Use this selection if you add extraction sets to your projects that require different protocols.
Project Default	If selected, the protocol specified as the project default protocol is assigned to <i>all</i> extraction sets in the project. If none is specified, then the software assigns the grid template default protocol. Use this selection if you typically add extraction sets that use the same protocol, and this protocol is not the default Agilent protocol assigned to the Grid Template.

Change the project default protocol

- 1 Click the arrow to the right of the cell next to **Project Default Protocol**.
- 2 Select a protocol from the list.

If the protocol you want is not shown on the list, you must import the protocol into the database. See [“To import or export a protocol to or from the database”](#) on page 64.

To select to use a grid file

To select to use a grid file

- Double-click on the cell next to Use Grid file if available and select True or False from the list.

True If selected, the software assigns to the extraction set the grid file with the same name as the image file and contained in the image directory.

False If selected, the software does not use any grid files in the image directory even if one that meets the criteria exists.

To select an external DyeNorm list file

If the project calls for a DyeNorm list and the image is a non-Agilent array, use an external DyeNorm list. This list must have the column names in the following order: Probe Name, Gene Name, and Systematic Name. This is the same list order for a DyeNorm list exported by the DyeNormList Editor.

- 1 Click the cell next to **External DyeNorm List File**.
- 2 Click **Browse**.
- 3 Select the .txt file, and click **Open**.

To learn how to create an external DyeNorm list from the DyeNormList Editor, see **“To create a new internal dye normalization gene list for a grid template”** on page 58.

External DyeNorm lists are only used with non-Agilent microarrays. For Agilent microarrays, specify the dye normalization list in the grid template. See **“To change the dye normalization gene list in the grid template”** on page 57.

To select to overwrite previous results

- Double-click on the cell next to **Overwrite Previous Results** and select **True** or **False**.

True If selected, a project extraction can overwrite existing files without warning.

False If False, the project extraction does not start if output files exist with the same name in the same folder.

To select the incoming image folder for On-Time projects

This selection is for On-Time projects only.

- 1 To assign a folder for incoming images, click the Browse button in the cell to the right of Incoming Image Folder.
- 2 Select a folder for the incoming images, and click **OK**.

To turn off the option to automatically extract XDR images

To select a time to terminate real-time extractions

This task is for On-Time projects only.

- 1 Expand **Project Run Terminating**.
- 2 Set one or both of the following settings to **True** or **False**.

Use Waiting Idle Time Out

If True, you can set an Idle Time in hours and/or minutes. The Idle Time is the time between the arrival of one scan file and the arrival of the next scan file into the real-time image folder.

Use Run Duration Time Out

If True, you can set a Duration Time in days, hours and/or minutes. The Duration Time is the time you are willing to wait for the project feature extraction to complete.

- 3 If you set either or both settings to True, expand the Idle Time or Duration Time folder, if not expanded.
- 4 Type the times.

To turn off the option to automatically extract XDR images

This task is for On-Time projects only.

- Double-click the cell to the right of **Automatic XDR Extraction** to change the default setting from **True** to **False**.

If the setting is True, the paired XDR images from the scanner are extracted in a combined fashion and a single output file set covering a broader range of data is generated.

If the setting False, the two images are extracted separately and two output file sets are generated.

Changing default project properties in Preferences

You set the default project properties in the Preferences dialog box. The defaults that you type or change here appear in the Project Properties tab when you create a new project.

To open the Preferences dialog box

If you want to change the default settings and values that appear in the Project Properties sheet for all projects, open the Preferences dialog box.

- Click **Tools > Preferences**.

The Preferences dialog box opens to the folder that was last opened.

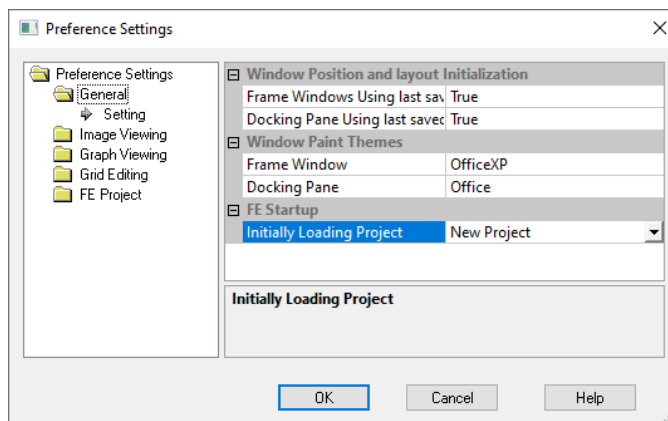


Figure 5 Preference Settings window

To display the General default project properties

- 1 Open the Preferences dialog box. See [“To open the Preferences dialog box”](#) on page 39.
- 2 Open the FE Project folder.
- 3 Select **General**.

To select to save the Project Run Summary Report

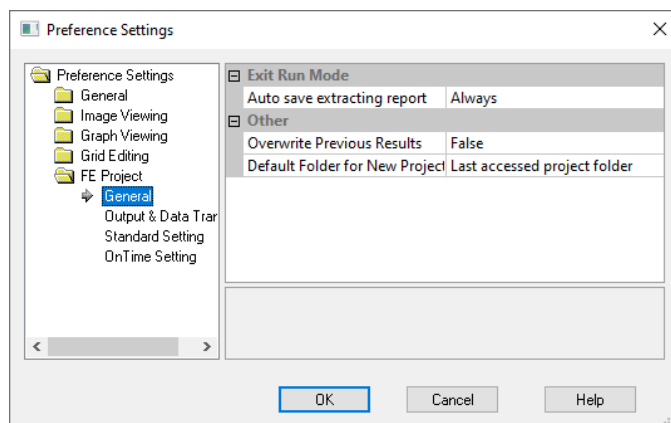


Figure 6 Preference Settings—FE Project General

To select to save the Project Run Summary Report

The Project Run Summary Report contains a history of the extractions run for a specific project. The Project Run Summary Report is initially saved to a temporary .rtf file (YourComputerName_LastBatchReport.rtf) in the directory containing the project files.

This file is overwritten every time an extraction is run on a project and contains only the latest summary report. If you want to save the summary report for each project for later viewing, then you must specify that intention here.

- 1 Display the General default project properties. See [“To display the General default project properties”](#) on page 39.
- 2 Click the cell next to **Auto save extracting report**.
- 3 Select from the list.

Always If selected, the summary report will always be saved after generation as a separate .rtf file in the project file directory. You can view this report for a specific project even after extractions are run on other projects.

Never If selected, the summary report for that specific project is never saved as a separate .rtf file after generation. You are able to view the report for the project by opening the temporary .rtf file after the project is run. After another project is run, you are not able to access the report for review.

Prompt If selected, when you exit the project, a prompt appears to ask if you want to save the summary report as a separate file.

To select to overwrite previous results

- 1 Display the General default project properties. See [“To display the General default project properties”](#) on page 39.
- 2 Double-click on the cell next to Overwrite Previous Results to select True or False.

True If selected, a project extraction can overwrite existing files without warning.

False If False, the project extraction does not start if output files exist with the same name in the same folder.

To select a default folder for a new project

To select a default folder for a new project

This option defines the initial folder into which you save a new project in the Save As dialog box.

- 1 Display the General default project properties. See **“To display the General default project properties”** on page 39.
- 2 Click on the cell next to **Default Folder for New Project**.
- 3 Select one of these options from the list.

Last accessed project folder Folder you accessed to save the previous project.

FE Default Project Folder Folder set by the application. If you let the installer choose the default installation folder, the path of this folder is C:\ProgramData\Agilent\FeatureExtractionCyto\FEProjects.

To display the default project properties for Output and Data Transfer

- 1 Open the Preferences dialog box. See **“To open the Preferences dialog box”** on page 39.
- 2 Open the FE Project folder.
- 3 Select **Output & Data Transfer**.

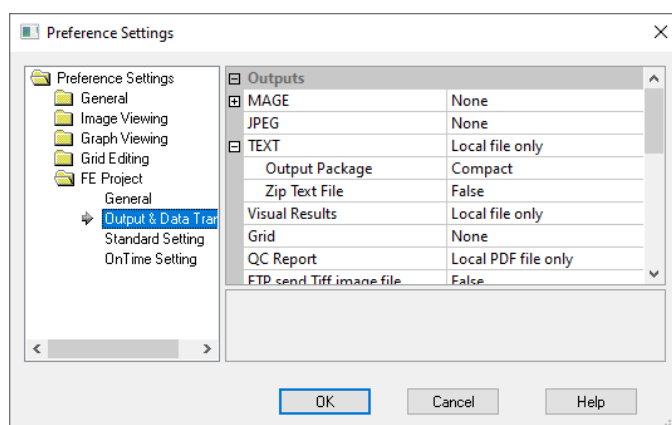


Figure 7 Preference Settings—Output & Data Transfer

To select the default location for saving and/or sending output files

- Display the default project properties. See **“To display the default project properties for Output and Data Transfer”** on page 41

See the Feature Extraction for CytoGenomics Reference Guide for a listing of all the results that are generated in the output files.

This portion of the Preferences dialog box lets you select the default location(s) for the image and result files. See **“Select the default location for output files”** on page 41 and **“Select the default local file folder for saved results”** on page 43.

Select the default location for output files

- Click the arrow to the right of the cell next to each of the result or image file types to select one of the following location(s).

To select the default location for saving and/or sending output files

Local file only Select to save the files to a local file folder. See **“Select the default location for output files”** on page 41.

FTP send only Select to send the files to an external location. See **“To change the default FTP settings for automatic data transfer”** on page 43.

Both local file and FTP send Select to save the files both to a local file folder and to an external location.

The following output files are generated and sent to the location(s) you specify:

MAGE result file This file contains results in the MAGE-ML (XML) data transfer format, which lets you transfer the data into Rosetta Resolver and Array Express databases.

You can select to send a Full Output Package of MAGE-ML data or a Compact, or reduced, set of data. And you can choose to send the data in a Compressed (ZIP) format, or not.

If MAGE output is selected, you also have the capacity to set and pass data ownership and permissions to Rosetta Resolver.

Expand the MAGE folder, and expand the Rosetta Resolver Data Access Setting folder to view the following fields:

If you do not type names for these settings, no access control information is entered into the MAGE-ML file.

Rosetta Resolver Data Access Settings: User Name Type the name of the user who will own the data transferred to the Rosetta Resolver database.

Rosetta Resolver Data Access Settings: Control Groups Type the names of the groups who have permission to use the data transferred to the Rosetta Resolver database. If there is more than one group, separate their names with a colon. For example, if you want the groups Public and Special to see the incoming extraction information, type `public:special`.

JPEG image file To create a compressed version of the scan file, you can create a JPEG file. A JPEG file stores a likeness of the original image, which is used for visual reference.

You can control the spatial compression of this JPEG output in the protocol. See **“To change the compression factor for the JPEG output”** on page 160 of this guide.

Text result file These results are in a tab-delimited text file that you can import into spreadsheets. The file contains input parameters, statistical calculations and feature results.

You can split this file into three files, each containing one of the sections. You can also import this file into GeneSpring data analysis software, either as a *Full* set of data, *Compact* (reduced), *QC*, or *Minimal* set. You can also choose to create a .zip file containing the text results.

NOTE

A Compact data set contains sufficient columns for use by Rosetta Resolver, GeneSpring GX, CGH Analytics or ChIP Analytics software. See the Feature Extraction for CytoGenomics Reference Guide for a list of these columns.

Visual Results Visual results from the feature extraction process are saved to a .shp or “shape” file, which only the Feature Extraction for CytoGenomics program can read. The compact version of the text file has the columns containing the final processing steps of the data (and other columns needed for import to downstream SW packages). It excludes some columns with intermediate data calculations.

Grid Choose to save the grid information to grid.csvfiles (for Agilent and non-Agilent grids) and feat.csv files (for non-Agilent grids). It contains the output of all of the grid placement steps (e.g., the spot centroids and spot sizes).

QC Report A QC Report is saved to the results directory.

To change the default FTP settings for automatic data transfer

Send a Tiff file via FTP

Because Tiff files already are located in a local file folder, you can select to send them to an external location via FTP, or not.

- Double-click the cell to the right of **FTP Send Tiff File**, and select **True** or **False**.

Select the default local file folder for saved results

Select to save results in the local image folder This folder must be the folder that contains the images in the extraction sets.

- Under Local File Folder, double-click the cell next to Same As Image to select True, if necessary.

Select to save results in a local results folder This folder can be any folder that you choose.

- 1 Double-click the cell next to **Same As Image** and select **False**.
- 2 Click the cell next to **Results Folder**.
- 3 Click **Browse**.
- 4 Select the folder to contain the results.
- 5 Click **OK**.

To change the default FTP settings for automatic data transfer

- 1 Display the default project properties for Output and Data Transfer. See [“To display the default project properties for Output and Data Transfer”](#) on page 41.
- 2 In the Project Properties tab, expand the FTP Setting folder, and type the values for these FTP settings:

Machine Destination	This field is specific to your Rosetta Resolver server name. Please contact your system administrator for the name to use.
Port	Type the number of the FTP port on the server. The default is port 21.
User name	Type the user name on the server that allows FTP access to the import directory of the server. The default User Name is resolverftp for Rosetta Resolver systems.
Password	This field already contains the default password for Rosetta Resolver servers. If it has been changed, contact your administrator for the correct password.
File Destination Folder	Scan information is sent to this folder. For MAGE results, the default setting is mage (/import/mage on the server). Confirm with your system administrator that the FTP account uses the import directory of the server as the default directory for the transfer.

To change default run post-processing plug-ins

- 1 Display the default project properties for Output and Data Transfer. See **“To display the default project properties for Output and Data Transfer”** on page 41.
- 2 In the Project Properties tab, expand the Post Processing Plug Ins folder, and type or select the plug-in you want to run.

The plug-ins are executed independently, so you can run the same or different plug-ins after each event.

- Run After Each Array** Select a plug in to run after each array is extracted. All array related arguments are passed to the executable called.
- Run After Each Slide** Select a plug in to run after each slide is extracted, where a slide contains 1 or more arrays. Arguments pointing to files relating to the slide are passed to the executable called.
- Run After Each Batch** Select a plug in to run after each batch of extractions is completed. Arguments pointing to files relating to the batch are passed to the executable called.

To display the default properties for standard projects

- 1 Open the Preferences dialog box. See **“To open the Preferences dialog box”** on page 39.
- 2 Open the FE Project folder.
- 3 Click **Standard Setting**.

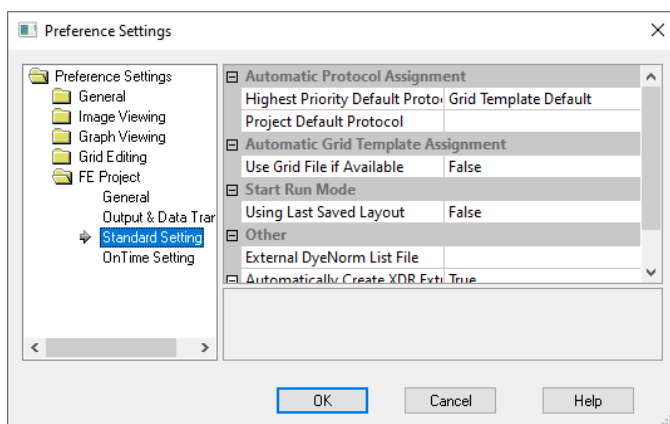


Figure 8 Preference Settings— Standard Project

To change the default protocol and its location

Select the default highest priority location for the default protocol

The software uses the protocol assigned to this location to complete an extraction set if a protocol is available in both the grid template and project. To learn more about extraction sets, grid templates and protocols, see **“Working with extraction sets”** on page 49.

To select to use a grid file

- 1 Display the default properties for standard projects. See [“To display the default properties for standard projects”](#) on page 44.
- 2 Click the arrow to the right of the cell next to **Highest Priority Default Protocol**.
- 3 Select **Grid Template Default** or **Project Default**.

Grid Template Default

If selected, the protocol specified in the grid template is assigned to the extraction set. If the grid template contains no protocol, then the software assigns the project default protocol. Use this selection if you add extraction sets to your projects that require different protocols.

Project Default

If selected, the project default protocol is assigned to the extraction set. If none is specified, then the software assigns the grid template default protocol. Use this selection if you typically add extraction sets that use the same protocol, and this protocol is not the default Agilent protocol assigned to the Grid Template.

Change the default protocol for all projects

- 1 Display the default properties for standard projects. See [“To display the default properties for standard projects”](#) on page 44.
- 2 Click the cell next to **Project Default Protocol**.
- 3 Select a protocol from the list.

To select to use a grid file

- 1 Display the default properties for standard projects. See [“To display the default properties for standard projects”](#) on page 44.
- 2 Double-click on the cell next to **Use Grid file if available** to select **True** or **False** from the list.

True If selected, the software assigns to the extraction set the grid file with the same name as the image file and contained in the image directory.

False If selected, the software does not use any grid files in the image directory even if one that meets the criteria exists.

To change the Start Run Mode setting

- 1 Display the default properties for standard projects. See [“To display the default properties for standard projects”](#) on page 44.

Number of extracting threads

This setting is fixed to one thread. You can no longer change the number of extracting threads.

- 2 Double-click the cell next to **Using Last Saved Layout** to select **True** or **False** from the list.

The default setting is False. If True is selected, the software saves the layout changes (i.e. placement of panels and browsers) for the Standard Project and reloads this layout the next time you open Feature Extraction for CytoGenomics.

To select an external DyeNorm list file

If the project calls for a dye normalization list and the image is from a non-Agilent array, use an external DyeNorm list. This list must have the column names in the following order: Probe Name, Gene Name, and Systematic Name. This is the same list order for a DyeNorm list exported by the DyeNormList Editor.

- 1 Display the default properties for standard projects. See [“To display the default properties for standard projects”](#) on page 44.
- 2 Click the cell next to **External DyeNorm List File**.
- 3 Select the external dye normalization list .txt file, and click **Open**.

External DyeNorm lists are only used with non-Agilent microarrays. For Agilent microarrays, specify the dye normalization list in the grid template. See [“To change the dye normalization gene list in the grid template”](#) on page 57.

To learn how to create an external DyeNorm list from a DyeNormList Editor, see [“To create a new internal dye normalization gene list for a grid template”](#) on page 58.

To select to create XDR extraction sets automatically

The default for this setting is True. When set to True, the software will automatically create one extraction set for the two images (_H and _L) scanned with extended dynamic range (XDR) turned on in the scanner.

- 1 Display the default properties for standard projects. See [“To display the default properties for standard projects”](#) on page 44.
- 2 Double-click the cell to the right of **Automatically Create XDR Extraction Set** to select either **True** or **False**.

If you want to be able to add just one of the two images to the extraction set and have the second image loaded automatically, follow step 3.

- 3 Double-click the cell to the right of **Search XDR image file pair on name matching only** to select either **True** or **False**.

True is the default selection. The pair images must be in the same folder for the software to load the second image automatically to the extraction set.

To display the on-time default properties

- 1 Open the Preferences dialog box. See [“To open the Preferences dialog box”](#) on page 39.
- 2 Open the FE Project folder.
- 3 Click **OnTime Setting**.

To select the default image folder

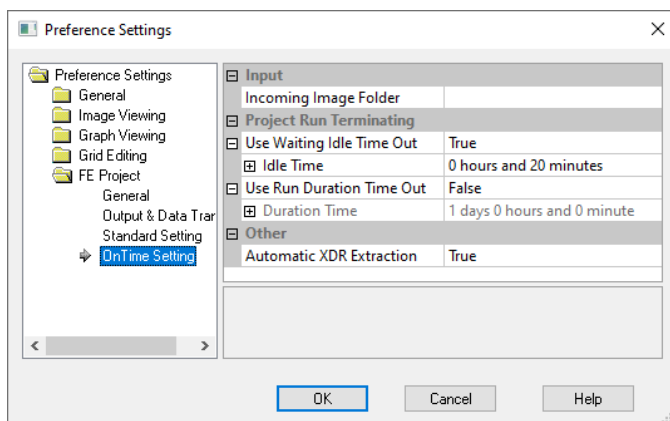


Figure 9 Preference Settings – On-Time Project Properties

To select the default image folder

- 1 Display the on-time default properties. See [“To display the on-time default properties”](#) on page 46.
- 2 Click on the cell next to **Incoming Image Folder**.
- 3 Click **Browse**.
- 4 Double-click the image folder, and click **OK**.

To select a default time to terminate real-time feature extraction

- 1 Display the on-time default properties. See [“To display the on-time default properties”](#) on page 46.
- 2 Expand **Project Run Terminating**.
- 3 Set one or both of the following settings to **True** or **False**.

Use Waiting Idle Time Out

If True, you can set an Idle Time in hours and/or minutes. The Idle Time is the time between the arrival of one scan file and the arrival of the next scan file into the real-time image folder.

Use Run Duration Time Out

If True, you can set a Duration Time in days, hours and/or minutes. The Duration Time is the time you are willing to wait for the project feature extraction to complete.

- 4 If you set either or both settings to True, expand the Idle Time or Duration Time folder, if not expanded.
- 5 Type the times.

To select to automatically extract an XDR image pair

The default setting is True. That is, the software will automatically extract both the High and Low signal images as a single extraction set when the XDR option is turned on in the scanner.

- 1 Display the on-time default properties. See **“To display the on-time default properties”** on page 46.
- 2 Double-click the cell to the right of **Automatic Extraction of XDR images** to change from **True** to **False**, or vice versa.

Working with extraction sets

What is an extraction set?

To each standard project you add one or more *extraction sets*, which consist of the *image file*, a *grid template or file* and a *protocol*. The protocol contains the values and settings for the parameters used in each step of the feature extraction process.

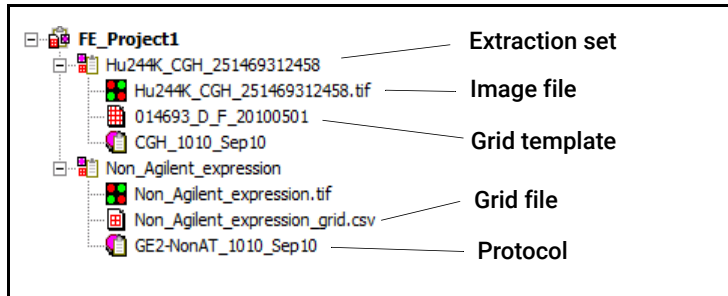


Figure 10 Extraction sets in Project Explorer

extraction set

Each grouping of microarray image (image file), grid template (or grid file) and protocol within a standard project. This term is used both before and after extraction. An XDR extraction includes two image files (_H and _L) and is represented with an icon that marks it as an XDR extraction. See [Figure 11](#).

20-bit single scans replace XDR dual scans. To determine if a scan is 20-bit, click the image file info button, click the Scan Statistics tab, and view the Saturation Value (about 777,000 for 20-bit).

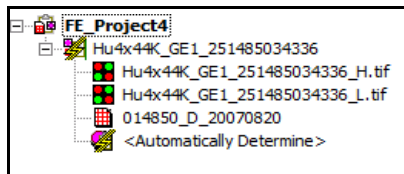


Figure 11 XDR extraction image files marked with _H and _L

grid template

Grid information from Agilent array design files stored in the database. A grid template includes feature annotation and general geometry about the microarray (number of rows, columns, subgrids, feature spacings), which is used to locate spots before data analysis takes place. Although not specific to any image, grid template information is usually applied to the image for which the template was designed.

grid file

You can also extract using a grid file, which has the nominal and “found” spot locations specific to the image from which the grid file was created. The grid file contains the locations of the “found” spots if the “Calculate spot size and centroids” option is run before the file is saved. You create grid files for non-Agilent images and specific Agilent images that have failed the gridding process.

protocol

A list of steps and parameter values that define the data analysis algorithms and calculations used for feature extraction of a microarray image or images.

How does the software create an extraction set?

Step a After you add an Agilent image file to the project, the software assigns a grid template from the database by matching the array design file's AMADID to the microarray barcode.

AMADID stands for **A**gilent **M**icro**A**rray **D**esign **I**Dentifier. The first 6 digits of the array design file indicate the AMADID. Given a design file 012097_D_20070820, 012097 is the AMADID and 20070820 is YYYYMMDD that the design file was generated.

For most Agilent microarray images, the grid template is automatically assigned, if the grid template is in the database. You may need to assign a grid template/grid file manually if, for example, the Agilent microarray image failed auto-gridding on a previous run.

A grid file must be used for a non-Agilent extraction set or if you had to manually grid the microarray.

Step b The software also adds a protocol to form an extraction set.

The added protocol is either the default protocol entered with the grid template or specified for the project. You can select which protocol is the first chosen by the software, the protocol in the grid template or the one in the project. If the added grid template or file does not contain a protocol, the software uses the default protocol for the project.

To add extraction sets (image files) to the standard project

You can add an image file or extraction set in one of four ways:

- Drag and drop the image file onto **Project Explorer**.
- Select the extraction set through the FE Project toolbar.
- Select the image file through the Edit menu.
- Select the image file through the Project Explorer shortcut menu

To add image files through Windows Explorer:

- 1 Open Windows Explorer.
- 2 Select the file or files that you intend to add.
- 3 Drag them to the Project Explorer and then drop them.

You now see a list of extraction sets.

- 4 Expand the extraction sets.

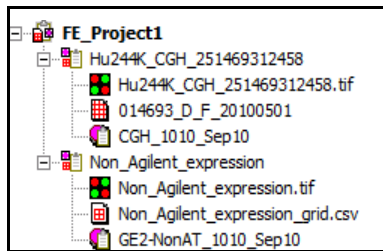



Figure 12 Project window after adding image files (extraction sets)

To sort extraction sets

To add extraction sets through the toolbar:

- 1 Click the Add Extraction Set(s) icon, .
- 2 Select the image files you want to add, then click **Open**.

You can select one or more images for extraction. Use the Shift key to select contiguous files or the Ctrl key to select separate files.

To add image files through the Edit menu:

- 1 Click **Edit > Add Extraction Set(s)**.
- 2 Select the image files you want to add, then click **Open**.

To add image files through the Project Explorer shortcut menu:

- 1 Right-click in **Project Explorer**.
- 2 Click **Add Extraction**.
- 3 Select the image files you want to add, then click **Open**.

You can also add extended dynamic range (XDR) extraction sets. These consist of a pair of images, one of high signal and the other of low signal, scanned with the XDR option turned on in the scanner. Feature Extraction for CytoGenomics extracts both images as a set and creates a single combined set of output files with data covering the combined dynamic range of both scans before it moves on to the next extraction set.

You must make sure that the preferences options for creating and extracting an XDR extraction set are turned on. The default selections are True (ON).

To sort extraction sets

If you want to run extraction sets in an order different from their order in Project Explorer, sort them using these instructions or drag and drop the extraction set in any order in Project Explorer.

- 1 Click **Edit > Sort Extraction Sets**.
- 2 Then select from the list – By Name, By Image, By Grid or By Protocol.

The list is sorted alphabetically and numerically for any of the choices.

To remove an extraction set from the project

To delete an extraction set from Project Explorer

- 1 Select an extraction set in Project Explorer.
You can select only one extraction set in Project Explorer.
- 2 Right-click **Project Explorer**, and click **Delete**.

The extraction set is immediately deleted.

Or

Click the Delete Extraction Set(s) icon, or click **Edit > Delete Extraction Set(s)**.

To display/edit the extraction Set Configuration table

If an extraction set in Project Explorer is selected, a message appears to ask if you want to delete the single set in Project Explorer. Click **Yes**.

If an extraction set or sets in the Extraction Set Configuration table is selected, a message appears to ask if you want to delete the extraction sets in the table.

To delete an extraction set or sets from the Extraction Set Configuration table

- 1 Select an extraction set or sets in the Extraction Set Configuration table.

You can select more than one by holding down the Ctrl key while you click additional extraction sets.

- 2 Right-click the selected extraction sets, and click **Delete Extraction set(s)**.

The extraction sets(s) are immediately deleted.

Or

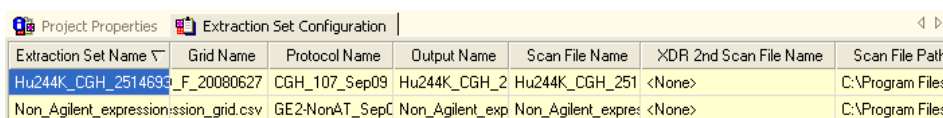
Click the Delete Extraction Set(s) icon, or click **Edit > Delete Extraction set(s)**.

If an extraction set or sets in the Extraction Set Configuration table are selected, a message opens to ask if you want to delete the extraction sets in the table. Click **Yes**.

If different extraction sets are selected in both Project Explorer and the Extraction Set Configuration table, a message opens to ask if you want to delete the extraction set in Project Explorer. Click **Yes**. To delete the selected sets in the Extraction Set Configuration table, you must repeat step 2.

To display/edit the extraction Set Configuration table

- 1 Click the Extraction Set Configuration tab.
- 2 Drag the column marker on the header to expand each cell.



Extraction Set Name	Grid Name	Protocol Name	Output Name	Scan File Name	XDR 2nd Scan File Name	Scan File Path
Hu244K_CGH_2514693	F_20080627	CGH_107_Sep09	Hu244K_CGH_2	Hu244K_CGH_251	<None>	C:\Program Files\
Non_Agilent_expression:ission_grid.csv	GE2-NonAT_SepC	Non_Agilent_exp	Non_Agilent_expre	<None>	<None>	C:\Program Files\

Figure 13 Extraction Set Configuration table

The Extraction Set Configuration tab is a table that lets you review your extraction sets all in one place.

- Extraction Set Name** The name of the extraction set. This name is automatically set to be the same as that of the scan file.
- Grid Name** The name of the grid template or grid file belonging to the extraction set.
- Click the cell and click the arrow to add or change this template or file from this column.
 - Double-click the cell to view the Grid Template Properties dialog box.
- Protocol Name** Protocol used with the grid template or project and belonging to the extraction set.
- Click the cell and click the arrow to add or change the protocol from this column.
 - Double-click the cell to view the Protocol Editor.

To display/edit the extraction Set Configuration table

Output Name	The name of the results file. This name is set to the name of the extraction set name with the name of the protocol added.
Scan File Name	The name of the scan file that belongs to the extraction set. <ul style="list-style-type: none">• Double-click this name to see the image in the scan file.
Scan File Path	The directory that contains the scan file.

Working with grid templates/files

A microarray design is the layout of probes on a single glass slide with a particular format. Each array design contains the biological and control probe groups, as well as a required Agilent quality control grid (grid template). When working in Feature Extraction for CytoGenomics, the Grid Template Browser displays the array designs from which the grid templates are derived. Designs that are displayed in the Grid Template Browser are the default designs that are included with the Agilent CytoGenomics software or were imported manually from a hard disk location.

Setting up and Changing Grid Templates and Files

To show or hide the Grid Template Browser

When you first open the Feature Extraction for CytoGenomics software after installation, the Grid Template Browser appears on the right-hand side of the main window.

If the Browser is closed, you can re-open it:

- Click **View > Grid Template Browser**.

To add or change a grid template to an extraction set

If you want to add an array design/grid template or use a different grid template with the image file in the extraction set, follow these instructions:

- 1 Select a grid template from the list of Grid Templates.
- 2 Drag and drop the grid template onto the extraction set.

The new grid template appears in the extraction set.

If you add or change an array design/grid template that does not match the design file for an Agilent image, the software highlights the template in orange in the Extraction Set Configuration tab to warn you. Extraction can still be run on this set, if you did not select for Feature Extraction for CytoGenomics to automatically assign a grid template, but gridding or annotation may be incorrect.

To add or change a grid file to an extraction set

Although you can add a grid file for Agilent microarrays, you must add or create a grid file for non-Agilent microarrays.

- 1 In the Project Explorer, right-click the second line of the extraction set, and click **Select grid file from the menu**.
- 2 Select the .csv file that you want to use with the image file, and click **Open**.

To add a grid template

Grid templates (design files) cannot be imported from within the Feature Extraction for CytoGenomics program; they must be imported as design files from within the Agilent CytoGenomics program. To add an array design/grid template to the Grid Template Browser from a file:

- 1 Launch Agilent CytoGenomics 5.3.
- 2 From the Supporting Files screen, click **Import Design**.
- 3 Select the design/grid template file (format *.xml) that you want to import into the CytoGenomics database.

To open grid mode from an extraction set

Once the design is imported into CytoGenomics, it appears in the Grid Template Browser of Feature Extraction for CytoGenomics. You may need to click **Tools > Grid Template > Refresh** to refresh the Grid Template Browser.

Grid templates are created from either array design files or grid files. They do not exist as objects in the file system.

See the CytoGenomics help system for more information on importing design files.

To open grid mode from an extraction set

If automatic grid placement fails during extraction, you may want to create and use a grid file in place of the grid template, or edit the grid file, depending on which is associated with the extraction set. Or, you may just want to “fine tune” a grid placement. You can more easily access grid mode from Project Explorer in configuration mode than from the image display.

For non-Agilent images, you can also access grid mode from the image display. See “Step 1. Start grid mode from the non-Agilent image file” on page 85.

- 1 In Project Explorer, right-click the grid template or grid file in the extraction set.
- 2 Click **Open Current Extraction in Grid Mode**.
The image file for the extraction set appears in grid mode, ready for gridding.
See **Chapter 3**, “Creating Grid Files” in this guide to learn how to set up or edit a grid file.
- 3 If you want to add a Design ID not in the list, type the ID in the Add Design ID field, and click **Add**.
- 4 To check the release dates, click **Check Latest Release Date**, Or
- 5 To add the most recent array design/grid template to the database, click **Update DB with New Release**.

If you want to add the latest release(s) to the Feature Extraction database without overwriting the most recent version, *clear* the check box for Update Grid Template in FE database by replacing the most recent version.

To display or change grid template properties

You can only change grid template properties such as default protocols and associated dye normalization lists. Array design properties (such as geometry) cannot be changed.

- 1 Double-click the grid template/array design whose properties you want to display or change
- 2 Display the grid template information.

You can display all of the grid information. You can also change the default protocol, the default 1-color protocol, and the dye normalization gene list.

To add or change the default protocol in the grid template

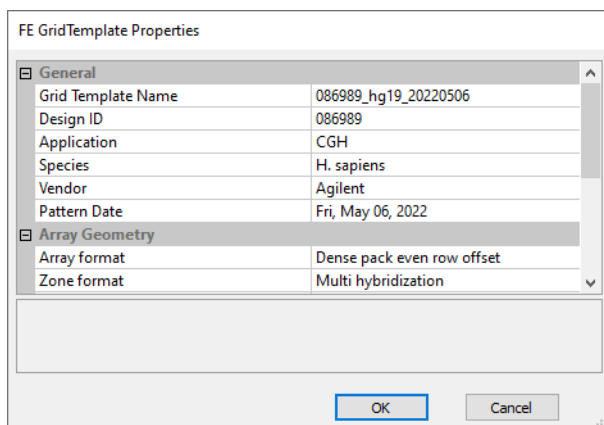


Figure 14 Grid Template Properties window

To add or change the default protocol in the grid template

- 1 Double-click the grid template whose protocol you want to add or change.
- 2 Click the cell to the right of Default protocol, and select the protocol from the list.

If you don't see the protocol you want from the list, you must import the protocol into the database first. See ["To import or export a protocol to or from the database"](#) on page 64.

To add or change the default one-color protocol in the grid template

Agilent design files and grid templates are the same for 2-color and 1-color or 2-color experiments. To make sure that the software can use a 1-color protocol with 1-color data, both 2-color and 1-color default protocols are specified. The software uses the 1-color protocol when the 1-color scan file is loaded.

- 1 Double-click the grid template whose protocol you want to add or change.
- 2 Click the cell to the right of Default OneColor protocol, and select a 1-color protocol from the list.

If you don't see the protocol you want from the list, you must import the protocol into the database first. See ["To import or export a protocol to or from the database"](#) on page 64.

To change the dye normalization gene list in the grid template

If the protocol specifies use of a dye normalization gene list and the extraction set includes a grid template, you create or select the gene list in the Grid Template Properties window.

- 1 Double-click the grid template whose dye normalization gene list you want to add or change.
- 2 Click the arrow to the right of the cell next to **Dye Normalization Gene List**.

To create a new internal dye normalization gene list for a grid template

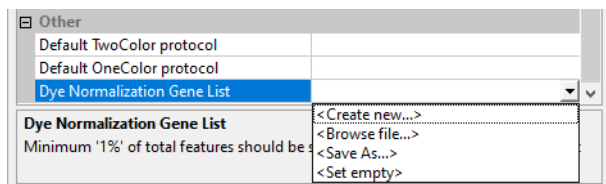


Figure 15 Dye Normalization Gene List options

3 Select one of the following options:

Dye Normalization Gene List

This is blank if no DyeNorm Gene Lists are available or if you select **Set Empty**.

If you previously created an internal dye normalization gene list, or selected an external gene list, select a name from this list. Once selected, you can double-click the gene list name to open the DyeNormList Editor to display or edit the list.

Create New Select this option if you want to create a new internal list from the existing probes. See **“To create a new internal dye normalization gene list for a grid template”** on page 58.

Browse file Select this option to browse and select an external gene list for the default list. External dye normalization lists are not imported into the database.

Save As Saves the selected dye normalization gene list as an external list. When saving, you can change the file name of the external list.

Set Empty Select this option if you do not want Feature Extraction to use a dye normalization gene list with this grid template.

To create a new internal dye normalization gene list for a grid template

- 1 Double-click the grid template for which you want to create a dye normalization gene list.
- 2 Click the cell to the right of **Dye Normalization Gene List**.
- 3 Click the arrow and click **Create New**.

NOTE

You can create a dye normalization gene list using the Gene/Probe View, or using the Chromosome View, but not both.

To create a new internal dye normalization gene list for a grid template

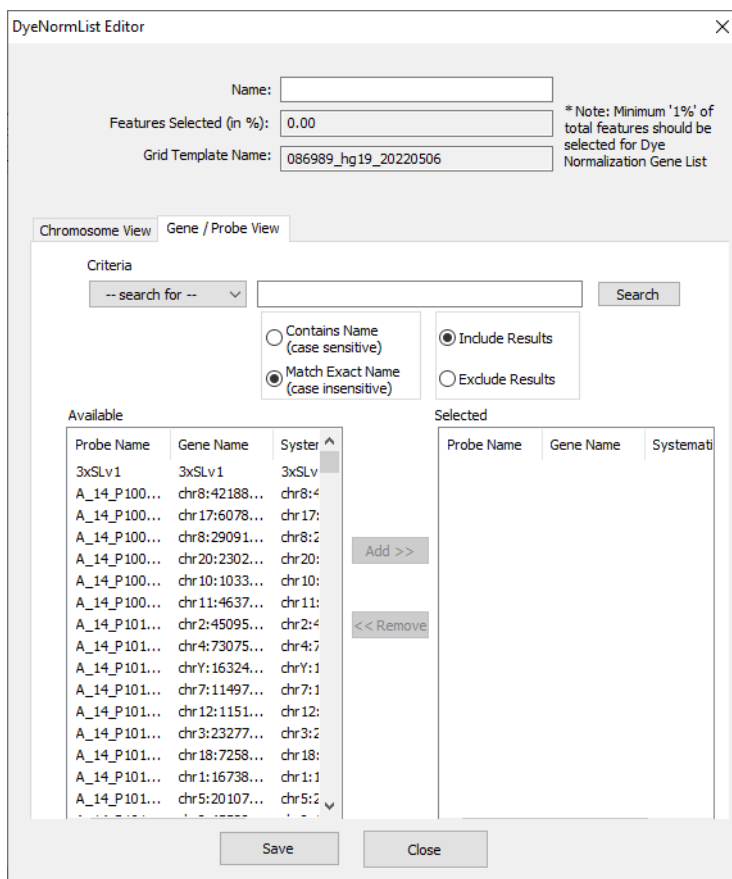


Figure 16 DyeNorm List Editor—Gene/Probe View

Add genes to list by search criteria

To streamline the process of selecting genes for the list, you can search for genes according to Probe Name, Gene Name or Systematic Name. SNP probes are not included.

- 1 If necessary, click the Gene/Probe View tab.
- 2 Click the box next to **Name** and type a name for the dye normalization list.
- 3 Under **Criteria**, click **Search for**, and then select **Probe Name**, **Gene Name** or **Systematic Name** as the type of name for the search.
- 4 Select how you want to use the name in the search. You can search according to the exact name, part of the name, or using a regular expression.
- 5 Type the name or part of the name for the search.
- 6 Select to include or exclude results.
- 7 Click **Search**.

Name Type the name for the dye normalization list. The list you create will be saved as an internal dye normalization list with this name.

Features Selected (in %) The dye normalization list must contain at least 1% of the total features in the array. This indicates the % of the features currently selected for the list.

Grid Template Name This shows the grid template with which you are associating this dye normalization list.

To create a new internal dye normalization gene list for a grid template

Contains Name (case sensitive)

If selected, you can search for part of the name, or you can use a regular expression. The case must match exactly. Regular expressions give you advanced search capabilities. For example,

- If you want to search all probes you can type the regular expression as “.”(dot) and it will search all the probes.

Match Exact Name (case insensitive)

If selected, you must type the exact probe name, gene name or systematic name, (depending on the type of name you selected in step 1.) For this search, the case does not need to match.

Include Results

If selected, all the results found are transferred from the Available Genes on Array list to the Selected Genes from Array list.

Exclude Results

If selected, all the results *not* found are transferred from the Available Genes on Array list to the Selected Genes from Array list.

Add chromosomes to list

- 1 Click the Chromosome View tab.

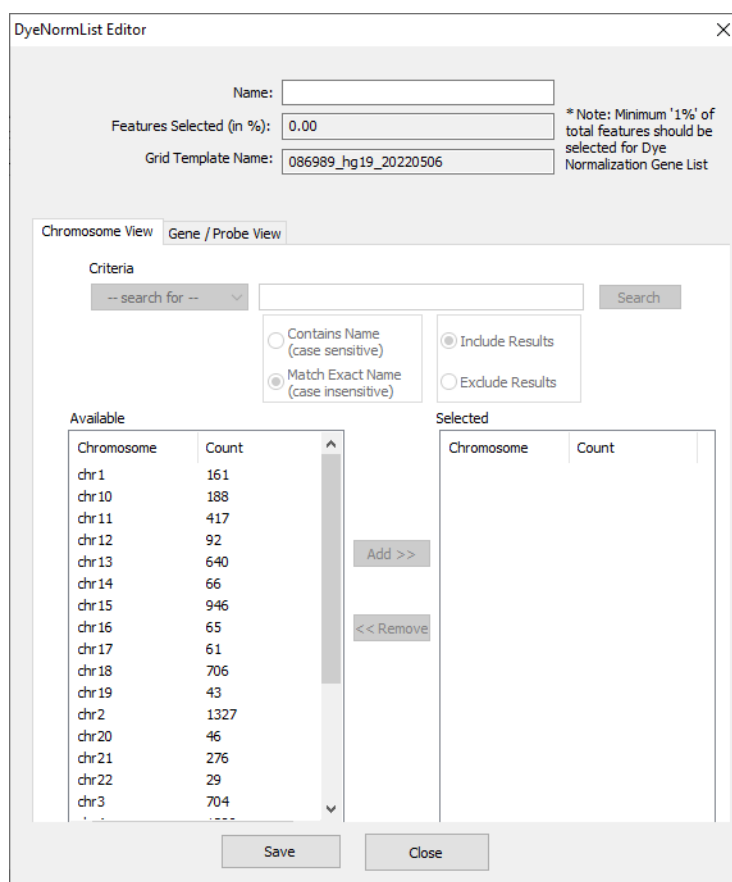


Figure 17 DyeNorm List Editor—Chromosome View

- 2 Select the chromosome(s) from the Available list, and click **Add>>**. See **“Add/remove genes or chromosomes to list interactively”** on page 61.

To edit an internal dye normalization list

Add/remove genes or chromosomes to list interactively

- 1 Select the gene(s) or chromosome(s) of interest from the Available list.
For multiple selection, hold down the Shift key and click the first name and then the last name or hold the Ctrl key and click for non-contiguous names.
- 2 Click **Add >>**, or
Click **<<Remove** to transfer genes from the Selected list to the Available list.

If you attempt to transfer a large number of genes or chromosomes from list to list, the process can take time.

Save dye normalization gene list to internal file

- 1 Type a name for the dye normalization list in the Name field.
- 2 Click **Save**.

When you click **Save**, a list is created using the Name specified, and is placed in the internal list available to the grid template.

You may have to refresh the screen in order to see the name of the new list in the Grid Template Properties window.

If you want to save the internal dye normalization list as an external file, see [“To export an internal dye normalization gene list”](#) on page 62.

To edit an internal dye normalization list

- 1 Double-click the grid template whose dye normalization gene list you want to change.
- 2 Click the cell to the right of **Dye Normalization Gene List**.
- 3 Select an existing gene list.
- 4 Double-click the list to bring up the DyeNormList Editor.
- 5 Add and remove any genes or chromosomes that you want.
- 6 Click **Save As**.

Once an internal dye normalization gene list is created, you cannot change it. You must save changes using a different name.

To import external dye normalization list to create a new internal list

- 1 Double-click the grid template whose dye normalization gene list you want to add or change.
- 2 If you want to create an internal list from an external .txt file, click the cell next to **Dye Normalization Gene List**, and click the down arrow.
- 3 Click **Browse file**, select the external .txt file and click **OK**.

This file must be in a special format: 3 columns and a carriage return with Probe Name, Gene Name and Systematic Name listed in that order. The file must not have a row of column headers.

- 4 Click the down arrow again, and select imported gene list as the default list.

To export an internal dye normalization gene list

If you need to make changes to the gene list, see [“To edit an internal dye normalization list”](#) on page 61.

To export an internal dye normalization gene list

After you create an internal dye normalization gene list file, you can export it as a .txt file to any folder you choose.

- 1 Double-click the grid template whose dye normalization gene list you want to export.
- 2 Click the down arrow to the right of the cell next to **Dye Normalization Gene List**, and select the internal gene list.
- 3 Click the down arrow again, and click **Save As**.
- 4 Select a folder you want to save the file to, rename the file, if necessary, and click **Save**.

The file is saved as an external .txt file.

Working with protocols

To show or hide the FE Protocol Browser

When you first open the Feature Extraction software after installation, the FE Protocol Browser appears on the right-hand side of the main window.

If the Browser has been closed, you can re-open it:

- Click **View > FE Protocol Browser**.

To add a protocol to an extraction set

By default, the program will automatically determine the protocol to use for the extraction. If you want to manually add a protocol to an extraction set, follow these instructions:

- 1 Click to select a protocol in the FE Protocol Browser.
- 2 Hold down the mouse button and drag and drop the protocol onto the third line in the extraction set.

The new protocol appears in the extraction set.

To change the protocol used in an extraction set

If you want to use a different protocol than that associated with the grid template or project, follow these instructions:

- 1 Select a protocol from the FE Protocol Browser.
- 2 Hold down the mouse button and drag and drop the protocol onto the extraction set.

The new protocol appears in the extraction set, and replaces the previous protocol.

To display or change protocol properties

- 1 Double-click a protocol in the FE Protocol Browser to open the Protocol Editor.

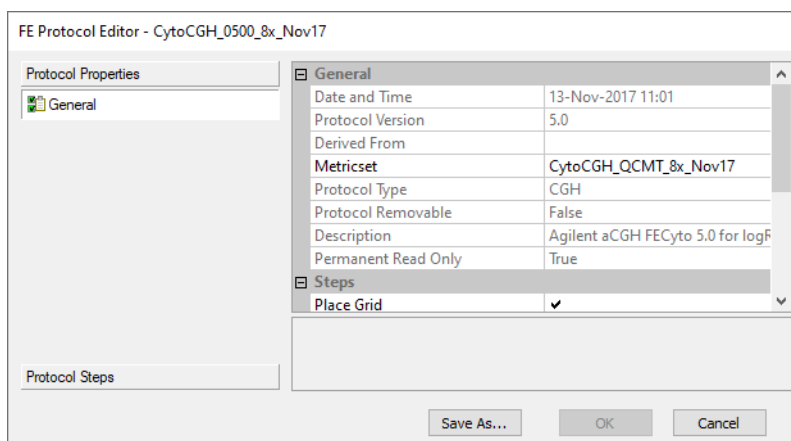


Figure 18 FE Protocol Editor—Protocol Properties

- 2 Click the Protocol Properties button which shows the General tab.

Protocols and QC metric sets are linked. Each protocol has a metric set assigned to it.

You can view information about the protocol. Under the General section, you can change the QC metric set that is associated with the protocol. Under the Steps section, you can remove protocol steps from the extraction run.

- 3 To change the QC metric set assigned to the protocol, click the cell to the right of **Metricset**, and select a QC Metric Set or **Set empty**.
- 4 To remove a protocol step from the run, clear the check box next to the step.
- 5 Click **Save As** to save the changes with a new protocol name.

See [Chapter 4](#), “Changing Protocol Settings” for more details.

To import or export a protocol to or from the database

To import a protocol to the database:

- 1 Click **Tools > FE Protocol > Import**.

Or

Right-click the FE Protocol Browser, and click **Import**.

- 2 Select a protocol from the list, and click Open.

To export a protocol from the database:

- 1 Click **Tools > FE Protocol > Export**.

OR

Right-click the FE Protocol Browser, and click **Export**.

- 2 Select a protocol from the list, and click **Save**.

A copy of the protocol is exported, but the protocol remains in the database.

To import or export a protocol to or from the database

NOTE

When importing a protocol from a previous version of Feature Extraction for CytoGenomics (i.e., version 4.0 or earlier), the enhanced gridding features are disabled by default. For instructions on enabling the enhancements, see **[“To select to use enhanced gridding”](#)** on page 126.

Working with QC metric sets

By default, the program calculates a set of default statistics that are displayed in the QC report. QC metric sets give you additional statistics for evaluating the feature extraction process on your microarrays. These statistics also appear in the QC report and in the Stats table of the text results file.

QC metric sets are linked with protocols. Each protocol has a metric set assigned to it. When the protocol is used by the software, the QC metric set is also used to generate metrics about the extraction set.

Agilent-created QC metric sets are installed with the Feature Extraction for CytoGenomics software. You can also create your own QC metrics sets using the Quality tools in Agilent CytoGenomics. When you create a QC metric set in Agilent CytoGenomics, it will automatically be available in the QC Metric Set Browser of Feature Extraction. See **“To import or export a QC Metric Set to or from the database”** on page 66.

If you wish to use your custom QC metric set with a protocol, you must associate it to the protocol. See **“To associate a QC metric set with a protocol”** on page 66.

To show or hide the QC Metric Set Browser

This Browser contains the most recent Agilent QC metric sets after the software is installed. In addition, it shows custom QC metric sets that were created in the Quality tab of Agilent CytoGenomics. To show or hide the QC Metric Set Browser:

- Click **View > QC Metric Set Browser**.

To associate a QC metric set with a protocol

You associate a QC metric set with a protocol in the FE Protocol Editor.

- 1 Double-click on the protocol to associate a QC metric set with. This opens the FE Protocol Editor.
- 2 Click the cell next to **Metricset**, and click the down arrow.
- 3 Select a QC metric set from the list.

To import or export a QC Metric Set to or from the database

To import a QC Metric Set to the database:

- 1 Click **Tools > QC Metric Set > Import**.

Or

Right-click the QC Metric Set Browser, and click **Import**.

- 2 Select a QC Metric Set from the list, and click **Open**.

To view metric set properties

To export a QC Metric Set from the database:

- 1 Click **Tools > QC Metric Set > Export**.

Or

Right-click on a QC metric set in the QC Metric Set Browser, and click **Export**.

- 2 Select a QC metric set from the list, and click **Save**.

A copy of the QC metric set is exported, but the metric set remains in the database.

To view metric set properties

- 1 In the QC Metric Set Browser, double-click a metric set you wish to view.

The QC Metric Set Properties for the selected metric set is displayed. Names of all the metrics included with this metric set are displayed, along with threshold values, if provided. For details on default metric set properties, see the *Feature Extraction for CytoGenomics Reference Guide*.

For details on the logic used for metric set evaluation, see the Feature Extraction for CytoGenomics Reference Guide.

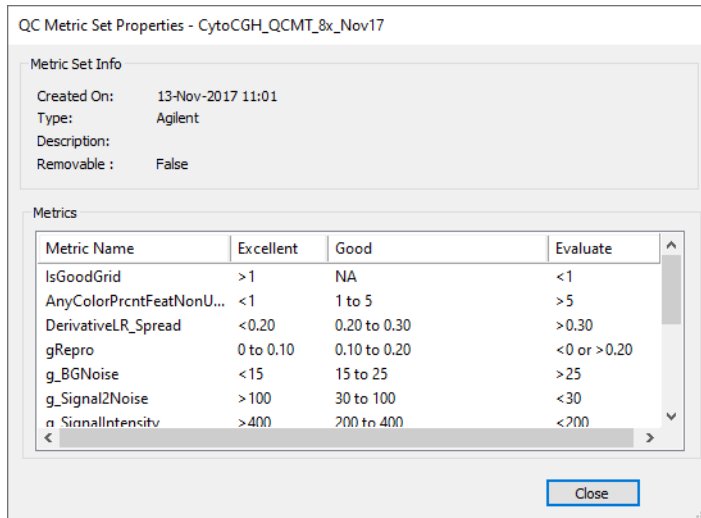


Figure 19 Metric set properties

- 2 Click **Close** to close the dialog box.

Running Feature Extraction

To start feature extraction

- 1 Click the Project Run mode On/Off icon in the toolbar, or click **Project > Start Extracting**.
If you did not save the project, a message appears that asks if you want to save the project. You must save the project in order to run feature extraction.
- 2 Click **Yes**.
- 3 Click **Save** if you want to save the project to the default FE Project N.fep name, or type a new name, and click **Save**.

During extraction, information on what steps are performed is displayed in the Running Monitor at the bottom of the window. A Project Run Summary report is generated in the Summary Report tab when the extraction is complete.

To stop feature extraction

- Click the Stop Extracting icon, or click **Project > Stop**.

To create a new project with remaining extraction sets if run fails

If a run fails due to a gridding problem, a message appears and provides an easy way to create a new extraction set containing the failed extraction sets. See **“Step 1. Create a new project for failed extraction sets”** on page 82 for details.

If a run fails for whatever reason, you can create a project containing the failed or not yet processed extraction sets left over from the previous project run. The following requirements must be met:

- Project must be a Standard FE project.
- Project must still be in the Run mode.
- Project has just stopped extracting and contains failed extraction set(s) or ones not yet run.

- 1 Click **Project > Create Project for Leftover Set(s)**.

A message appears asking if you want to switch back to Config mode and continue.

- 2 Click **Yes**.

Feature Extraction for CytoGenomics creates a new project with all unfinished extraction sets from the current project. Feature Extraction for CytoGenomics also closes the project and starts the new project in its Config mode.

The name for the newly created project is “Parent project name_2”.

You must save the new project before you can extract or exit the project.

Displaying results

A feature extraction run on any image produces the following files:

- A Project Run Summary report .rtf file
- Text result .txt files (generated per microarray)
- For Agilent arrays, a .pdf QC report (generated per microarray) or .html QC report (if you are running with a temporary license).
- A .shp visual result file

NOTE

An Agilent multiplex array is one that has multiple Hyb/Wash chambers for separate samples on a single slide. These were formerly referred to as “multipack” arrays.

For Agilent multiplex arrays, text & QC reports are generated for each subarray, but just one .shp and Project Run Summary is produced per extraction of the image.

This section describes how to view these results.

To display the text result file in Microsoft Excel

The result files are named with the following format:

Single pack: extraction set name_protocol name.

Multiplex (formerly known as Multipack): extraction set name_protocol name_pack number, indicated by row number.

- 1 Go to the folder specified for containing the results.
- 2 Right-click the text results file (protocol name on end), and click **Open with**.
- 3 Click **Microsoft Excel**.

The text results will open in the spreadsheet. The following figure shows all three tables (FEPARAMS, STATS, and FEATURE) in the file. You can also choose to create a different text file for each table. See [“To select to generate a single file for the text output”](#) on page 160 for more information.

To display the Project Run Summary Report

TYPE	text	text	text	integer	text	float	float	text	integer	text	text	integer	integer	integer
FEPARAM	Protocol	Protocol	Scan_Sca	Scan_Nur	Scan_Dat	Scan_Mici	Scan_Mici	Scan_Orig	Scan_Nurr	Grid_Namr	Grid_Date	Grid_Nurr	Grid_Nurr	Grid_Nurr
DATA	CGH_105	#####	Agilent Te	2	#####	3		3 c121ba31	1	021529_D	#####	1	1	1068
TYPE	float	float	float	integer	integer	float	float	float	integer	integer	float	float	integer	float
STATS	gDarkOffs	gDarkOffs	gDarkOffs	gDarkOffs	gSaturation	gDarkOffs	gDarkOffs	gDarkOffs	gDarkOffs	gSaturation	gAvgSig2f	gAvgSig2B	gNumSaf	gLocalBG
DATA	24.3684	24.3684	1.99337	1048576	65510	24.5267	24.5267	2.32225	1048576	65510	1.75806	1.28801	8	11.5053
TYPE	integer	integer	integer	integer	integer	text	text	float	float	float	float	float	float	float
FEATURE	FeatureNu	Row	Col	SubTypeM	ControlTyp	ProbeNan	Systemati	PositionX	PositionY	LogRatio	LogRatioE	PValueLog	Processr	Processr
DATA	1	1	1	0	1	HsCGHBr	HsCGHBr	606.969	388.687	5.45E-02	6.17E-02	3.77E-01	5.56E+03	6.31E+03
DATA	2	1	2	66	1	DarkCorn	DarkCorn	628.333	388.609	0.00E+00	6.20E-01	1.00E+00	3.25E+01	3.35E+01
DATA	3	1	3	66	1	DarkCorn	DarkCorn	649.5	388.676	0.00E+00	6.20E-01	1.00E+00	3.25E+01	3.35E+01
DATA	4	1	4	0	0	A_16_P01	chr5:0750	670.509	388.632	-1.85E-02	6.70E-02	7.82E-01	7.81E+02	7.48E+02
DATA	5	1	5	0	0	A_16_P01	chr4:1145	691.426	388.926	4.03E-02	7.10E-02	5.71E-01	5.49E+02	6.02E+02
DATA	6	1	6	0	0	A_16_P16	chr4:0413	712.577	388.923	-3.54E-02	6.62E-02	5.93E-01	8.69E+02	8.01E+02
DATA	7	1	7	0	0	A_16_P0C	chr2:2077	733.673	388.909	-2.02E-02	6.23E-02	7.46E-01	1.99E+03	1.90E+03
DATA	8	1	8	0	0	A_16_P3E	chr7:0777	755.158	388.684	-3.58E-02	6.69E-02	5.92E-01	8.04E+02	7.41E+02
DATA	9	1	9	0	0	A_16_P0C	chr1:1500	776.038	388.865	5.78E-02	6.35E-02	3.63E-01	1.22E+03	1.40E+03
DATA	10	1	10	0	0	A_16_P2C	chr1:7071	797.5	388.722	7.14E-02	6.41E-02	2.85E-01	1.08E+03	1.28E+03
DATA	11	1	11	0	0	A_14_P13	chr5:0986	818.75	388.607	1.77E-02	6.38E-02	7.82E-01	1.14E+03	1.19E+03
DATA	12	1	12	0	0	A_16_P02	chr10:024	839.825	388.614	6.04E-02	7.01E-02	3.89E-01	5.64E+02	6.48E+02
DATA	13	1	13	0	0	A_16_P3E	chr9:1294	860.903	388.403	2.80E-02	7.35E-02	7.03E-01	4.83E+02	5.15E+02
DATA	14	1	14	0	0	A_16_P0C	chr1:1490	882.317	388.733	4.42E-02	6.25E-02	4.79E-01	1.80E+03	2.00E+03
DATA	15	1	15	0	0	A_16_P3E	chr2:1976	903.154	388.846	1.83E-03	6.27E-02	9.77E-01	1.56E+03	1.57E+03
DATA	16	1	16	0	0	A_16_P17	chr6:0804	924.333	388.741	-1.08E-01	6.48E-02	9.54E-02	1.24E+03	9.66E+02
DATA	17	1	17	0	0	A_16_P1E	chr7:0829	945.407	389.237	-5.00E-02	6.25E-02	4.23E-01	2.01E+03	1.79E+03
DATA	18	1	18	0	0	A_16_P1E	chr5:0368	966.793	388.948	-3.44E-02	6.24E-02	5.82E-01	1.92E+03	1.78E+03
DATA	19	1	19	0	0	A_16_P3E	chr1:1006	988.075	388.679	7.13E-02	6.50E-02	2.73E-01	9.03E+02	1.06E+03
DATA	20	1	20	0	0	A_18_P23	chr3:1710	1009.34	389.16	7.80E-02	7.30E-02	2.86E-01	4.67E+02	5.59E+02
DATA	21	1	21	0	0	A_16_P17	chr5:1399	1030.32	388.877	6.44E-02	6.22E-02	3.00E-01	2.30E+03	2.67E+03
DATA	22	1	22	0	0	A_16_P1E	chr4:0266	1051.31	388.741	8.57E-03	6.25E-02	8.91E-01	1.74E+03	1.77E+03

Figure 20 Portion of text results file in Microsoft Excel

For some microarrays, the results will not all fit in an Excel spreadsheet (only the first 65536 rows in Excel 2003 and earlier). Results can be viewed in GeneSpring or DNA Analytics.

To display the Project Run Summary Report

When you start the extraction run, the Summary Report tab appears showing the project start time. Once extraction is completed, you can see information about the extraction set in run summary. This summary is automatically saved as a Project Run Summary report that you can access later.

- 1 Go to the folder containing the project whose report you want to display.

The default folder for Feature Extraction projects is C:\ProgramData\Agilent\FeatureExtractionCyto\FEProjects.

- 2 Double-click the file named YourComputerName_LastBatchReport.rtf to display the last saved report for the project.

Or, double-click the file named FE Project N_timestamp.rtf to display a summary report for a specific run of the project.

To display the Project Run Summary Report

```
Project Run Summary: AgilentExampleImages

Project started on Fri, Sep 10, 2010 at 16:32:35.

D:\AGW Data\Hu22K_GE2_251209710036.tif Fri Sep 10 16:32:35 2010
INFO: Grid Template in use : 012097_D_20070820
INFO: Protocol in use: GE2_1010_Sep10
INFO: MetricSet in use: GE2_QCMT_Sep10.
INFO: Agilent supports but has not tested this configuration, please consult the support matrix in the Feature Extrac
INFO: FE automatically determined that ArrayFormat is Single Density
INFO: FE automatically determined that GridFormat is 65 micron feature size
INFO: FE automatically determined that SpotFormat is Single Density
INFO: FE automatically determined that Scanner is Agilent Scanner
INFO: FE automatically determined that GridTestFormat is 65 micron feature size
INFO: There are 1829 (Red) negative features (Non-Controls).
INFO: There are 5445 (Green) negative features (Non-Controls).
INFO: 16 (Red) and 1 (Green) saturated features
INFO: 4 (Red) and 8 (Green) feature non-uniformity outliers
INFO: 98 (Red) and 73 (Green) feature population outliers
INFO: 9 (Red) and 0 (Green) background non-uniformity outliers
INFO: 48 (Red) and 0 (Green) background population outliers
INFO: QCMetrics Totals: Found 12 Metrics In Good Range
INFO: Extraction completed successfully.
QCREPORT: <D:\AGW Data\Hu22K_GE2_251209710036_GE2_1010_Sep10.pdf>
D:\AGW Data\Hu4x44K_GE1_251485034336_H.tif Fri Sep 10 16:33:31 2010
Processing 4 arrays in single scan.
INFO: Grid Template in use : 014850_D_20070820
INFO: Protocol in use: GE1_1010_Sep10
INFO: multiple extractions
INFO: MetricSet in use: GE1_QCMT_Sep10.
INFO: FE automatically determined that ArrayFormat is 65 micron feature size
INFO: FE automatically determined that GridFormat is 65 micron feature size
```

Figure 21 Project Run Summary

You can open the QC Report by double-clicking the blue QC Report link in the Project Run Summary.

Evaluation of grid placement failures

If you see one of these warnings in the Project Summary Report, check the QC report and the .shp file to confirm the success or failure of the grid placement.

- Grid failure error. There are many checks that are protocol dependent. These include number of features not found, number non-uniform local background regions, negative control standard deviation, and %CV of replicated features. Depending upon the thresholds set for these metrics in the protocol, the array may or may not show a grid failure error.
- The average difference between the nominal grid location and the spot centroid is greater than 10 microns.

If the grid location does not fit within the image, the software will stop the extraction and produce an error in the Project Run Summary.

If automated extraction fails for an Agilent image

Several options are available to troubleshoot the failure. See **“If automatic gridding for an Agilent image fails”** on page 82 for instructions for a successful grid placement and extraction.

If there is a warning that there is no barcode in the MAGE-ML output

Absence of a barcode may be due to several reasons:

- Barcode reader failure
- Scanner cannot read the Agilent microarray barcode for some reason
- Scan region set up to scan a non-Agilent microarray does not contain the barcode

To display the QC Report

For Feature Extraction to load MAGE-ML results files into Rosetta Resolver, they must have a barcode or microarray identifier associated with them. You can add this identifier in the Scan Image Properties dialog box. See **“To add or change the New Array Identifier”** on page 166.

To display the QC Report

The report generated depends on the protocol used, and is specified in the Generate Results settings for the protocol. See **“Changing Protocol Step 9: Generate Results”** on page 159.

- 1 Go to the folder specified for containing the results.
- 2 Double-click the QC Report .pdf or .html file.

The QC Report can be generated in .pdf or .html format. If the Feature Extraction license is a demo license, only an .html file is generated.

You can also access the QC Report from the Project Run Summary. Double-click the blue link next to QC Report. See **Figure 21** on page 71.

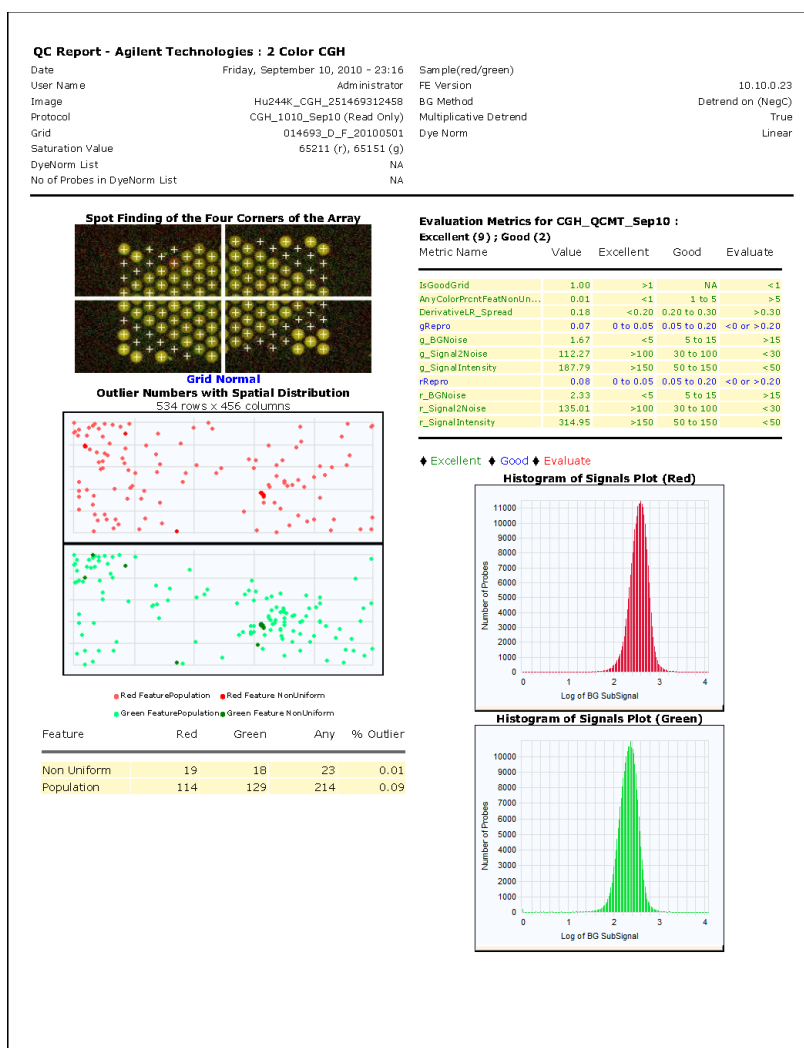


Figure 22 Streamlined CGH QC Report—Page 1

To print the QC Report

If the QC report is open in Internet Explorer, follow these instructions:

- 1 Click **File > Page Setup**.
- 2 Make sure that it prints in Portrait mode.
- 3 Select A4 size paper.
- 4 Click **Printer > Properties**, and select for color printing.
- 5 Exit until you see **Page Setup**, then close Page Setup.
- 6 Click **File > Print Preview**.
- 7 Adjust the margins to display all the images.

You can alternatively use another Internet browser that has a Shrink to Fit option for printing, such as Mozilla Firefox.

- 8 Click **Print**.

To display visual results

After extraction is completed, Feature Extraction generates one visual shape result file (.shp) for each tiff image file (.tif). For multiplex Agilent microarrays (more than one microarray on a single glass slide), a single .shp file is for the entire tiff image and covers all the arrays on the slide.

The shape file is only viewable in Feature Extraction. It provides visual results about the extracted microarray image. You can view extraction results such as feature cookies, gene names, raw signals, log ratios, outliers, and pixels used to quantify features (available for high-resolution images).

Load the visual results file (.shp) from Project Explorer

- 1 Right-click a processed extraction set.
- 2 Click **View Visual Result**.

The image is loaded, along with its shape file into the Image workspace. If no shape file is present, the menu item is grayed out.

Load the visual results file (.shp) from the Image workspace

- 1 Click **Open Image**, select the .tif file, and click **Open**.

You can also open the image from Project Explorer (right-click the image name and select **View Visual Result**) or the Extraction Set Configuration tab sheet (double-click the **Scan File Name**).

- 2 Click Feature Extraction > Load Visual Result.
- 3 Select the .shp file associated with the image, and click Open.

Work with the visual results

By default, when you load a visual results file, Feature Extraction shows the image in View Outliers Only mode, displaying information about outlier features only. To view results for non-outlier features, see **"View results on low-resolution images:"** on page 74, and **"View results on high-resolution images (2- by 3- micron or 20-bit):"** on page 74.

For more details on how to load the shape file from the Image workspace, see "Loading Feature Extraction visual results" on page 167.

To display visual results

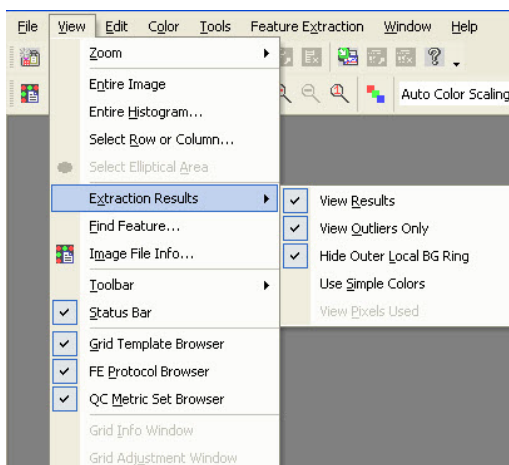



Figure 23 View Extraction Results—Default Settings

View results on low-resolution images:

- 1 After .shp file is loaded to the image, zoom in until you see enough of the spots to help you evaluate the results.
- 2 Click the cropping mode icon. The cursor turns from the cropping cursor to the arrow cursor.
- 3 Click **View > Extraction Results**, and clear the **View Outlier Only** check box.
- 4 On the image, pass cursor to feature of interest and hold for a few seconds. A description of the feature appears.

To find a specific feature, use Ctrl-F. Type the Feature number of the specific feature you want to find.

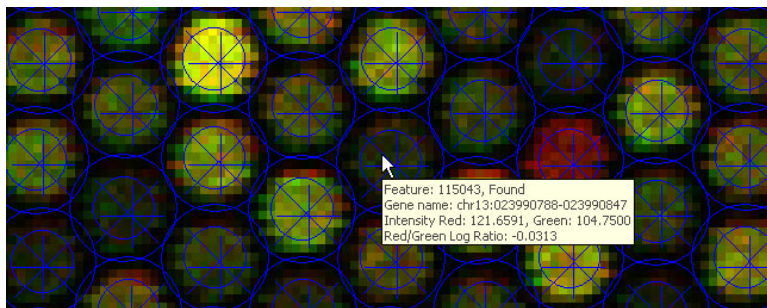


Figure 24 Visual results for extracted microarray image

View results on high-resolution images (2- by 3- micron or 20-bit):

After you first load the image, you, in fact, are viewing a preview image. You can only view the original high-fidelity image if you crop the size of the image to one-tenth or less of the original size.

- 1 After .shp file is loaded to the image, zoom in until you see enough of the spots to help you identify the region of interest.
- 2 Click the left mouse button at the point where you want to start forming a rectangle for cropping.
- 3 With the crop cursor, draw a rectangle equal to or less than one-tenth the size of the image.

Note that the crop cursor is black if the size is within one-tenth (200% to 1000%) of the original size, but it turns red if you are drawing outside the high fidelity range.

To display visual results

- 4 Then release the mouse button.

Note the status bar at the bottom of the window. It should now read “High Fidelity”.

- 5 Click **View > Extraction Results**, and clear the check box **View Outlier Only**.

- 6 On the image, pass cursor to feature of interest and hold for a few seconds.

A description of the feature appears like in **Figure 24** on page 74.

View Pixels Used

This option is available for low-resolution and high-resolution images that are in “High Fidelity” mode (i.e. when the image is zoomed in to one-tenth of its original size).

The View Pixels Used option displays the pixels used to quantify the feature signal within the cookie area. Pixels that are determined to be outliers are shown in black.

- 1 Complete steps 1-5 from “**View results on high-resolution images (2- by 3- micron or 20-bit):**” on page 74.

- 2 Under **View > Extraction Results**, select **View Pixels Used**.

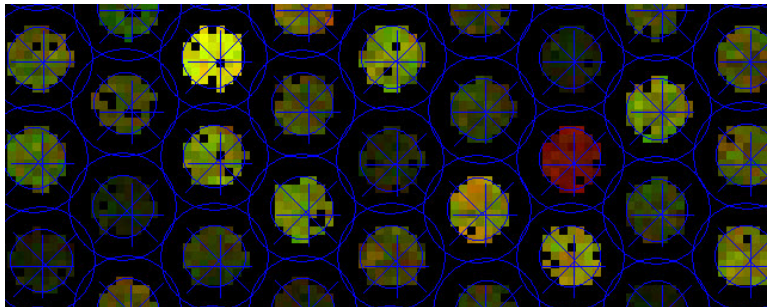


Figure 25 View Pixels Used results on a high-resolution image in high-fidelity mode

To display visual results

3 Creating Grid Files

About grid templates and grid files	78
When to use manual gridding	79
Analyzing the initial grid for an Agilent image	80
If automatic gridding for an Agilent image fails	82
Creating and using a grid file for non-Agilent images overview	85
Starting grid mode for a non-Agilent image	87
Viewing and changing grid geometry	93
Optimizing grid placement using the help wizard	95
Manual grid adjustments	101
Saving a grid file	105
Calculating spot size and spot centroids	106
Locating spot centroids manually	107
Refitting a grid	112
Changing default settings for Grid Editing	114
Shortcuts to help you work with grids and find spots	115

Agilent Feature Extraction for CytoGenomics software uses grid templates and grid files to determine where spots are located on a microarray before extraction.

This chapter shows you how to troubleshoot failed Agilent image extractions and create grid files for Agilent and non-Agilent images.

Each section of this chapter describes the actions you take to properly position a main grid, subgrids, and ultimately the spot centroids. The section also gives you background information to help you position these grids and spot centroids.

About grid templates and grid files

Feature Extraction for CytoGenomics automates the process of locating microarray spots through the use of grid templates or grid files, which contain spacing information required to place grids for specific Agilent microarray designs. Grid templates are derived from array design files of the same name that reside in the shared Feature Extraction/Agilent CytoGenomics database and are displayed in the Feature Extraction Grid Template Browser. When you import, add, or update a grid template, you actually import, add, or update the entire array design file.

A grid template contains spacing information for the specified format as well as a unique set of probes and annotations for the microarray design that it matches. A grid file is similar to a grid template, except it contains information specific for a single scan of a given microarray slide.

If gridding fails for a particular Agilent microarray, you may have to create an Agilent grid file to use for the extraction. Agilent grid files are saved using the image name followed by `_grid.csv`. If you specify a grid file in a project for extraction of an Agilent image, you must also specify the grid template to be used for the rest of the extraction steps.

For non-Agilent microarrays in a project, you must create an Agilent grid file for the specific non-Agilent microarray you intend to extract. A non-Agilent grid file is stored as a `.CSV` (comma separated value) file and is not saved in the database. When you save a non-Agilent grid file, two `.CSV` files are actually created that together define the grid. They are named using the image name followed by `_grid.csv` or `_feat.csv`. Both files must be present in order to define the grid.

When you create a non-Agilent grid file, you must locate the spot *centroids*, or geometric centers, on the microarrays. To locate spot centroids you interactively position a main grid and subgrids on the microarray. After the subgrids are well-placed, you calculate spot size and centroids.

The software algorithm should locate most, if not all, of the spots where you want them. Any that are incorrect can be moved using a "Spot navigator" (see **"Locating spot centroids manually"** on page 107). This ensures that the Feature Extraction for CytoGenomics software identifies the correct feature and local background.

You cannot, nor do you need to, adjust spot centroids for Agilent images because Agilent array printing is accurate once an initial grid is found.

When to use manual gridding

Table 2 Uses of manual gridding

Image Type	Reason to use manual gridding	See
Agilent Image	Look at how the grid file or template has placed the grid	"Analyzing the initial grid for an Agilent image" on page 80
Agilent Image	If a gridding error appears in the Summary Report after Feature Extraction	"If automatic gridding for an Agilent image fails" on page 82
Non-Agilent Image	Create a grid file to be used with the image for Feature Extraction	"Creating and using a grid file for non-Agilent images overview" on page 85
Agilent or non-Agilent image	Grid is marked as "unplaced" in Extraction Set	"Optimizing grid placement using the help wizard" on page 95
Non-Agilent Image	Evaluation of grid placement	"Starting grid mode for a non-Agilent image" on page 87
non-Agilent image	Spot spacing or location of centroids is not correct	"Calculating spot size and spot centroids" on page 106 "Locating spot centroids manually" on page 107
Agilent or non-Agilent image	Grid placement is not correct after automatic placement	"Manual grid adjustments" on page 101
Agilent (rarely) or non-Agilent image	Grid does not fit the spots on the grid (spacing is incorrect in the Geometry)	"Refitting a grid" on page 112

Analyzing the initial grid for an Agilent image

When you add an Agilent image to a project, the software assigns a grid template, based on the barcode (AMADID) and a protocol. The protocol uses the grid template to automatically find the spots. Normally, you will not need to do any grid adjustments for an Agilent image, however if you want to display the placement of the grid, or if the grid is labeled as “unplaced” you can open the extraction in grid mode and make adjustments.

To open the current extraction in grid mode

For Agilent images, the grid template associated with the image is automatically assigned to the extraction set when the grid template is available in the Grid Template Browser.

You can toggle back and forth from grid mode to the project explorer from the Window menu on the toolbar.

- 1 Open a project file, or add an extraction to an existing project file.
- 2 If no grid is open, drag and drop a grid template from the Grid Template Browser, or right click on the grid icon and click **Select Grid File** to select an existing grid file.
- 3 Select a protocol.

You must select a protocol that matches the image for grid mode. If not, the grid and spot placement may not be correct.

- 4 In the Project Explorer, right-click on the extraction set and select **Open Current Extraction in Grid Mode**.

The software attempts to place the grid and the image opens in grid mode. In grid mode, you can view the grid placement and make adjustments if necessary.

See [“Optimizing grid placement using the help wizard”](#) on page 95.

To open the current extraction in grid mode

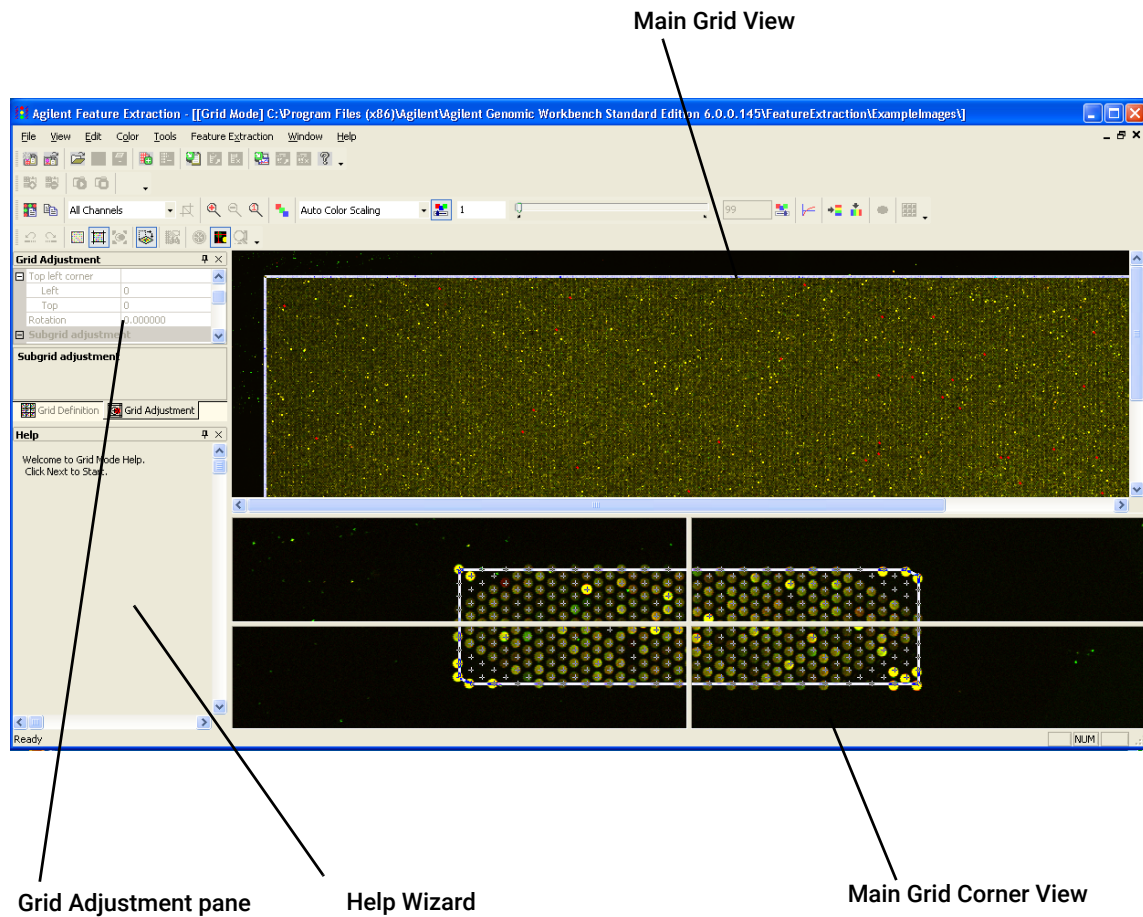


Figure 26 Extraction opened in grid mode

If automatic gridding for an Agilent image fails

If the gridding fails during feature extraction, you will see a red error message in the Summary Report that indicates why the gridding failed. See **“To display the Project Run Summary Report”** on page 70 for details on summary reports. If gridding fails, follow the steps below to adjust the grid and rerun an extraction set.

Step 1. Create a new project for failed extraction sets

Step 2. Access grid mode from a project extraction set

Step 3. Position grids and save grid file

Step 4. Set up and rerun Feature Extraction

Step 5. Adjusting rotation and skew of subgrids

Step 1. Create a new project for failed extraction sets

When you attempt to close a project or exit the Feature Extraction software after an extraction where one or more extractions failed due to a gridding error, the software will ask if you want to create a new project containing the failed extraction sets.

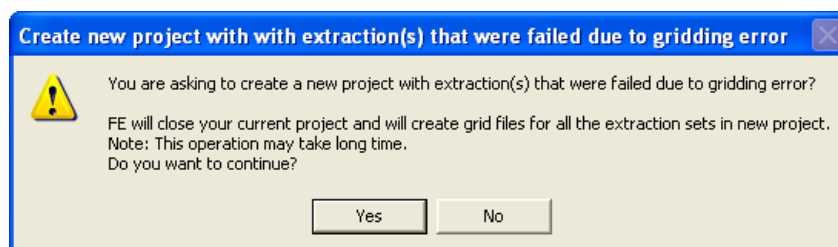


Figure 27 Failed extractions due to gridding error

If you click **No**, the project or Feature Extraction software will close.

If you click **Yes**, the software will

- Create a grid file for all the failed extractions (due to gridding error)
- Create a project for all failed extractions (due to gridding error)
- Created grids are automatically assigned to their respective extraction sets and their status is changed to “unplaced”.

The name for the newly created project is “Parent project name_2”. Every time a new project from leftover sets is created, the software appends _2 to the parent name.

You must save the new project before you can extract or exit the project.

Step 1b. Manually create a new project for failed extraction sets

An *unplaced* grid is one whose x- and y-coordinates, skew, and rotation are equal to zero.

When you run the project containing the failed extraction sets, Feature Extraction will check the grid placement status of the first extraction set. If the grid is “unplaced” a message will open asking you if you want to open grid mode. Click **Yes** to open the extraction in grid mode where you can place and save the grid. If you click **No**, the software will check the next extraction set in the project until all extraction sets have been checked.

Step 1b. Manually create a new project for failed extraction sets

If a run fails for any reason, you can create a project containing the failed or not yet processed extraction sets left over from the previous project run. The following requirements must be met:

- Project must be a Standard FE project
- Project must still be in the Run mode
- Project has just stopped extracting and contains failed extraction set(s) or extraction set(s) not yet run.

1 Select **Project > Create Project for Leftover Set(s)**.

A message appears asking if you want to switch back to Config mode and continue.

2 Click **Yes**.

Feature Extraction creates a new project with all unfinished extraction sets from the current project. Feature Extraction also closes the current project and starts the new project in Config mode.

The name for the newly created project is “Parent project name_2”. Every time a new project from leftover sets is created, the software appends _2 to the parent name.

You must save the new project before you can extract or exit the project. To proceed, open the extraction set in grid mode by following Step 2, below.

Step 2. Access grid mode from a project extraction set

This is the quickest way to access grid mode after automatic grid placement has failed. It is the only way to access grid mode for Agilent images.

After gridding, the *extraction set* consists of the image file, the grid template and grid file, and the protocol.

In the Project Explorer on the left side of the project window, right-click on any object in the extraction set (extraction set name, image, grid template or file, protocol).

3 Select **Open Current Extraction in Grid Mode**.

The image file for the extraction set appears in grid mode, ready for positioning of the main grid.

Also, the grid file appears under the original grid template in the extraction set.

Step 3. Position grids and save grid file

- 1 Open the extraction in grid mode.
- 2 Follow the help wizard steps to:
 - a Position the main grid.
 - b Position subgrids, if necessary.

The spot size and spot centroid options in the Help Wizard are used for non-Agilent images only.

See **“Optimizing grid placement using the help wizard”** on page 95.

- 3 Save the grid file.

See **“Saving a grid file”** on page 105.

The grid is labeled as “unplaced” in the Project Explorer until the grid is positioned and grid mode is closed.

Step 4. Set up and rerun Feature Extraction

- 1 Close the grid mode window.
- 2 Save the project.
- 3 Click the Project Run mode On/Off icon to start the extraction.

Creating and using a grid file for non-Agilent images overview

Follow these instructions to create and use a grid file for non-Agilent images. You must create a grid file for non-Agilent images to run Feature Extraction.

Step 1. Start grid mode from the non-Agilent image file.

Step 2. Position grids, locate spots and save grid file.

Step 3. Extract the image file with its new grid file.

Step 1. Start grid mode from the non-Agilent image file

1 Open the non-Agilent image file.

See **“To open an image file”** on page 87.

2 Open grid mode and select the gene list type and protocol to use.

See **“To open grid mode for a non-Agilent image”** on page 88.

Grid mode cannot be accessed from an open Agilent image. For Agilent images, you must access grid mode from within an extraction set. See **“Step 2. Access grid mode from a project extraction set”** on page 83.

3 Click **OK**.

4 In the Geometry Information dialog box, click **Auto Fit**.

The image file for the extraction set appears in grid mode, ready for positioning of the main window. See **“Geometry information”** on page 91.

Step 2. Position grids, locate spots and save grid file

If you are using a non-Agilent microarray, the decision to adjust spots, calculate spot sizes and centroids, or position spots manually is based on visual inspection of how the grid is fitted.

1 Follow the help wizard steps in the lower left corner of the grid mode window to:

a Determine nominal spot and subgrid spacing manually. (Optional and only for third-party arrays.)


b Determine nominal subgrid spacing. (Optional and for third-party arrays with subgrids only.)

c Position main grid and subgrids, if necessary.

See **“Optimizing grid placement using the help wizard”** on page 95.

The spot size and spot centroid options are used for non-Agilent images only.

2 Click the Adjust Spots icon  to enable spot location adjustment.

3 Click the Calculate spot size and centroids  icon and calculate the spot sizes and centroids.

See **“Calculating spot size and spot centroids”** on page 106.

4 Visually inspect the spot centroids and if necessary, position spots manually.

Step 3. Extract the image file with its new grid file

See **“Locating spot centroids manually”** on page 107.

- 5 Save the grid file.

See **“Saving a grid file”** on page 105.

Step 3. Extract the image file with its new grid file

To test the new grid file, create the project and add the extraction set.

- 1 Select **Feature Extraction > Create a New FE Project**.

A new Feature Extraction project is created with the image file already added.

- 2 In the Project Explorer, right-click on the grid icon below the image file, and click **Select Grid File**.
- 3 In the Select Grid file for Extraction Set dialog box, select your saved grid file, and then click **Open**.
- 4 From the FE Protocol Browser, drag the protocol for the image type to the extraction set.
- 5 Save the project and run Feature Extraction to see if the grid works.

Starting grid mode for a non-Agilent image

You can view and adjust grid placement using grid mode. For non-Agilent images, grid mode can be accessed from either an open image file or from an extraction set that has an associated grid and protocol.

To open an image file

You can also open an image file by dragging the file to the Feature Extraction desktop icon.

- 1 Click on the toolbar Open image file button, or click **File > Open > Image**, to display the Open Dialog Box.
- 2 In the Open Dialog Box, double-click on a .tif file from the scanner, or select a file and click **Open**.

An image of the scanned microarray contained in the file opens in its own window.

You cannot access grid mode for Agilent images from an open image. See **“Step 2. Access grid mode from a project extraction set”** on page 83.

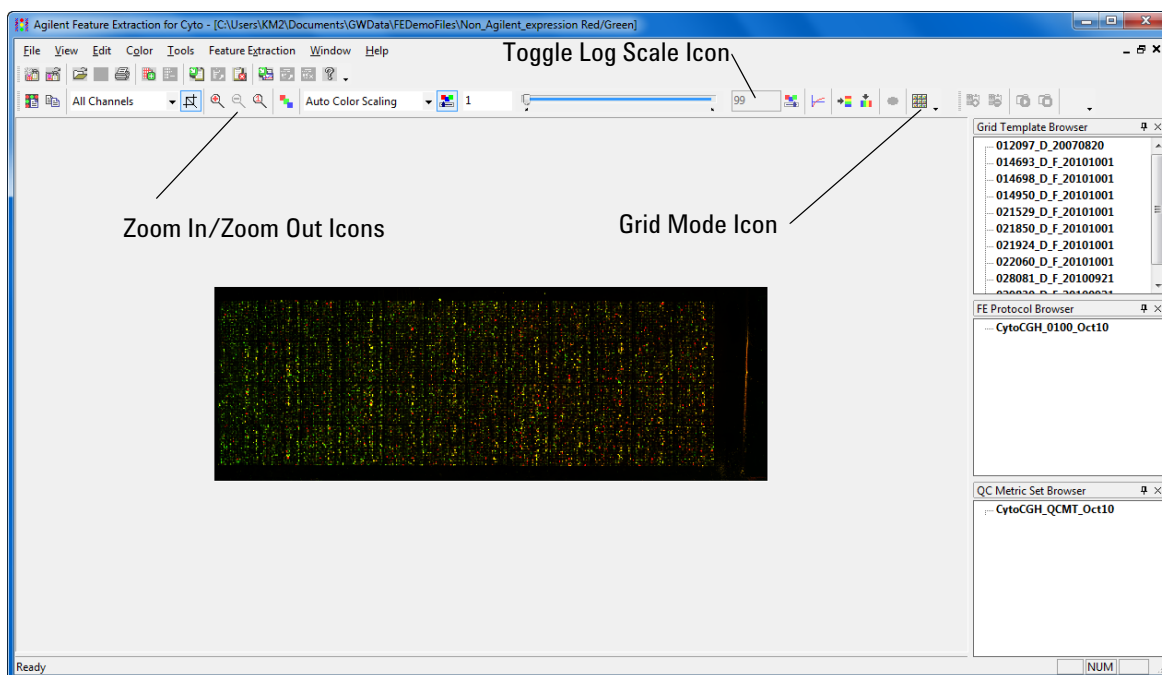


Figure 28 TIFF image for non-Agilent microarray

To change the Portrait/Landscape view of the image

- 3 To see the spots more clearly, click the Toggle Log Scale icon and click the Zoom In icon to increase the image size.

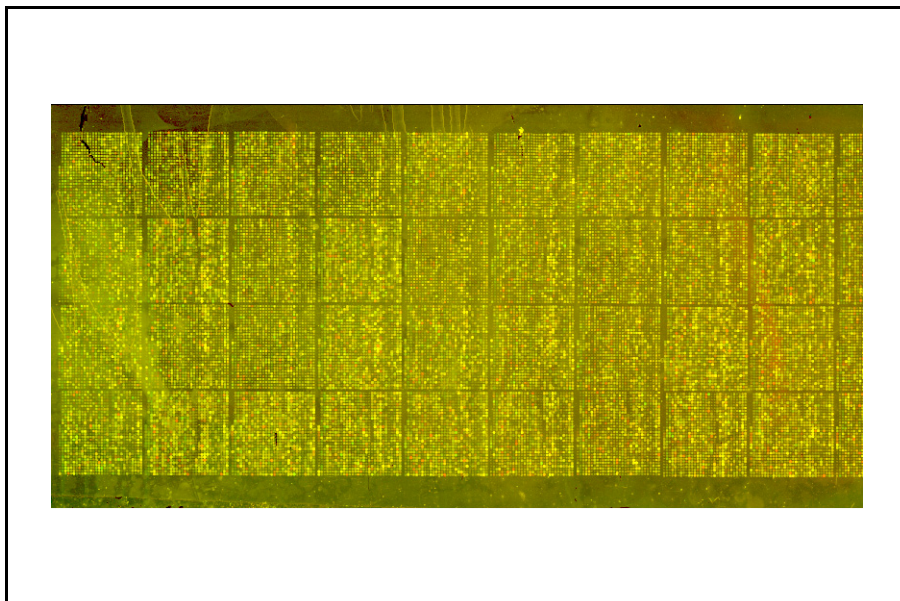


Figure 29 Image view with log scale on and image zoomed

To change the Portrait/Landscape view of the image

You must change the orientation of any opened image if the gene list file that you intend to use to provide grid information organizes its data with the opposite orientation. You typically do not have to do this for Agilent images, but you usually have to do this for non-Agilent images.

The software changes orientation in this way because the Agilent scanner scans the long side of the microarray through the glass. This is an orientation that is flipped and rotated compared with other scanners.

- 1 Select **Tools > Flip Upper Left to Lower Right (Landscape/Portrait)**.

The software takes the upper left corner of the image and flips it to the lower right no matter if you change the orientation from Portrait to Landscape or vice-versa.

- 2 Select **File > Save modified image** to save the image.

To open grid mode for a non-Agilent image

- 1 Open a non-Agilent image file.
- 2 Click the Grid button on the image toolbar.



The Create/Edit Grid dialog box opens.

To open grid mode for a non-Agilent image

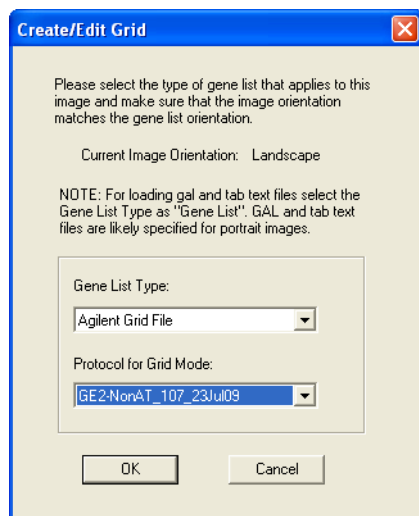


Figure 30 Create/Edit Grid dialog box

- 3 Select the Gene List Type to let you position the grid.

If you select a file whose data orientation is different from that of the image, the Agilent software will not be able to set the grid properly. Cancel out of this dialog box and change the orientation if you have to.

Agilent Grid File (*_grid.csv)

This is a grid file that has been saved from the previous creation of a grid on an image.

If you want to edit a grid or change the location of the spot centroids for an image, select the grid file previously saved for that image.

If you have never set up an initial grid and the image resembles one whose grid you have set up before, select that grid file from the already existing grid files.

Gene List

Select this if the image that you opened has an associated GAL or tab text file.

The layout information and feature annotations must be in tab-delimited text format and must contain at least these columns: Grid index (or SubGridRow/SubGridCol), SpotRow, SpotCol and GeneName or Probename.

The tab text file can be a GAL file, a results file or an Excel file exported as a tab-delimited text file.

Control Type can also be in the tab text file. See the *Feature Extraction for CytoGenomics Reference Guide* to see the definitions of the control types that can appear in the tab text file.

No Gene List

If you have no geometric information associated with the microarray, select this option. You can also select this option if another file you have used does not result in a workable grid. If you select this option, the Feature Extraction output file does not contain any feature annotations.

- 4 Select a protocol.

Select a protocol to help with grid placement and spot finding. Grid placement and spot finding algorithms differ between protocols. You must select a protocol that matches the image for grid mode. If not, the grid and spot placement may not be correct.

- 5 Click **OK**.

What happens next will depend on what you have selected for the gene list type:

To open grid mode for a non-Agilent image

If you select Agilent Grid File as the gene list type

A Browser appears that lets you select the grid file. The grid file must already exist for the image.

- Select a grid file, and click **Open**.

The Main Grid View appears with the grid approximately positioned.

If you select Gene list and then select a gal file as the gene list type

The Geometry Information dialog box appears for the gal file grid.

- 1 Change the default values if you need to.
- 2 Click **Auto Fit** or **Manual Fit**.

For a description of all the fields and buttons in the Geometry Information dialog box, see **“Geometry information”** on page 91.

The Main Grid View for this image appears.

If you select Gene list and then select a tab text file as the gene list type

The Tab Text Grid File Preview window opens.

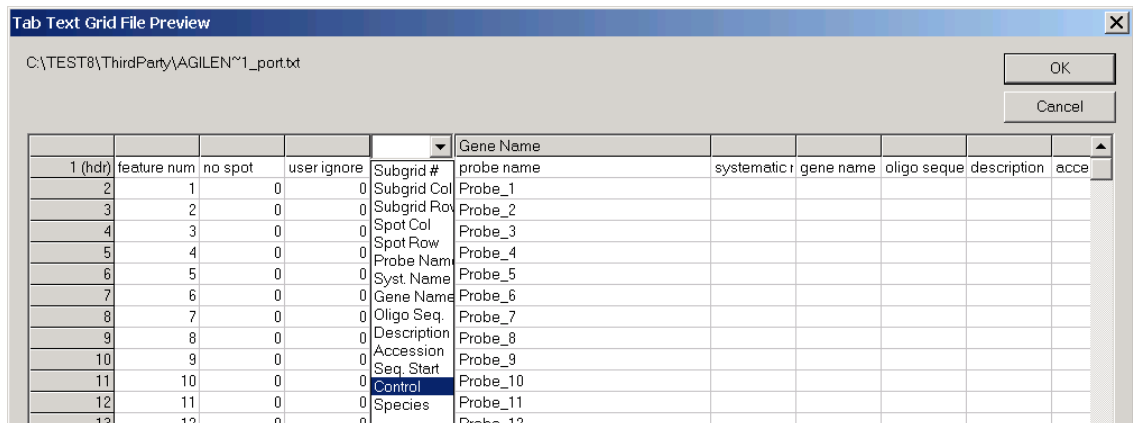


Figure 31 Tab Text Grid File Preview window with header selected

- 1 Click each of the column headers that correspond to the following information fields:
 - Subgrid Col
 - Subgrid Row
 - Spot Col
 - Spot Row
 - Probe Name
- 2 Select the appropriate name from the list.

You must change the names of the columns listed above in order for gridding to work with a tab text file. You can also change other column names, such as Gene Name or Control.

To open grid mode for a non-Agilent image

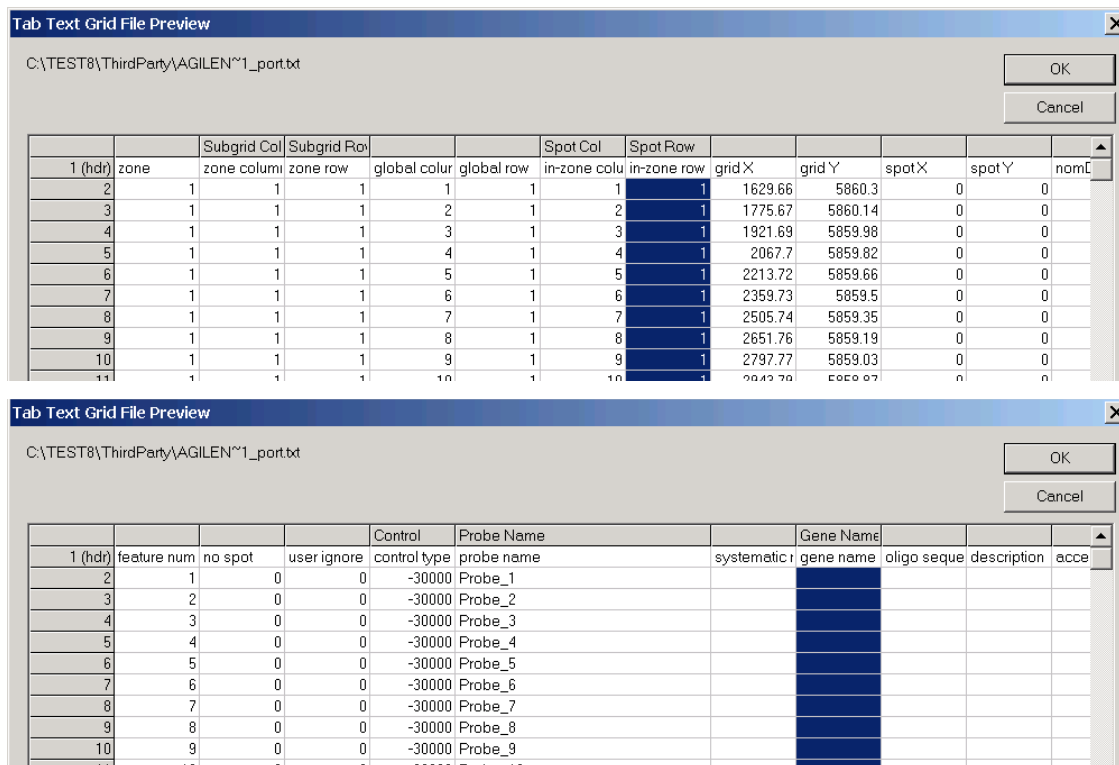


Figure 32 Tab Text Grid File Preview window with headers selected

3 Click **OK**.

The Geometry Information dialog box opens.

4 Click **Auto Fit** or **Manual Fit**.

For a description of all the fields and buttons in the Geometry Information dialog box, see [“Viewing and changing grid geometry”](#) on page 93.

The Main Grid View for the non-Agilent image opens.

If you select **No Gene List**

The Geometry Information dialog box opens.

1 Change the default values if you need to.

2 Click **Auto Fit** or **Manual Fit**.

For a description of all the fields and buttons in the Geometry Information dialog box, see [“Viewing and changing grid geometry”](#) on page 93.

The Main Grid View for the non-Agilent image appears.

Geometry information

This dialog box opens when you click **OK** from the Create/Edit Grid dialog box.

If Gene List or No Gene List are selected as the Gene List Type in the Create/Edit Grid dialog box, the Feature Extraction software attempts to discover the grid and will populate the Geometry Information fields with the information from that attempt.

To open grid mode for a non-Agilent image

For Gene List, if you select a .gal file, the parameters from the .gal file is displayed in the Geometry Information dialog box.

For details on all of the fields and buttons of the Geometry Information dialog box, see **“Viewing and changing grid geometry”** on page 93.

Viewing and changing grid geometry

The following dialog box appears when you click OK from the Create/Edit Grid dialog box when entering grid mode from a non-Agilent image file (see **“Starting grid mode for a non-Agilent image”** on page 87), or when you Refit a grid (see **“Refitting a grid”** on page 112).

Figure 33 Geometry Information dialog box

Main grid/Array geometry

For the main grid, enter dimensions in the form of number of columns and rows, or enter the distance between subgrids in pixels for both columns and rows.

Subgrid geometry

For subgrids, enter dimensions in the form of number of columns and rows, or enter the distance between spots in pixels for both columns and rows.

Double Density Offsets

The double density offset is applied when the array being gridded contains features consisting of two interlaced rectangular grids. The spacing of these rectangular grids is specified in the Double Density row and column offsets.

Nominal spot diameter

Because spots may not be circular, enter the horizontal and vertical dimensions in pixels.

Reset

Click this to reset the displayed parameters to their original values.

Auto Fit

The software uses the entries for Geometry Information as a starting point for placing the grid, then adjusts the subgrids for the best fit and identifies the spot centroids on the image. Choose this unless you know the automatic grid fit is not correct.

Manual Fit

The software places the grid on the image, usually in the upper left-hand corner. You then align the grid with the corners of the image. Based on this grid fit, the software identifies the spot centroids. Choose this if you have tried Auto Fit and you need to fit the grid manually.

Cancel

Click this to close the dialog box and return to the image display.

Table 3 Geometry information that you can change on the Geometry Information (GI) dialog box*

File Type	Initial GI dialog box displays?	Change main grid dimensions?	Change subgrid dimensions?	Change spot spacing?	Change spot size?
Agilent Grid	No [†]	No	No	Yes	No
GAL Format	Yes	Yes	No	Yes	Yes
Tab Text	Yes	Yes	No	Yes	Yes
No gene list	Yes	Yes	Yes	Yes	Yes

* You can change the nominal spot and subgrid spacing in grid mode.

† The Geometry Information dialog box does not appear when you select Agilent Grid File as the gene list type. The columns for the Agilent Grid file type apply to the “refit” Geometry Information dialog box, which does appear when you attempt to refit an Agilent grid (see **“Refitting a grid”** on page 112).

Optimizing grid placement using the help wizard

Although the Help Wizard is available for both Agilent and non-Agilent arrays, some steps apply only to non-Agilent (third-party) arrays.

Grid mode provides a Help Wizard to step you through the process of optimizing the placement of the grids. To view the help steps, in the Help pane on the lower left of the grid mode window, click Next. To go back to the previous step, click Previous. The help wizard provides instructions for the following tasks:

Step 1. Get nominal spot spacing manually (for non-Agilent arrays only)

Step 2. Get nominal subgrid spacing manually (for non-Agilent arrays only)

Step 3. Adjust the position of the main grid

Step 4. Adjust the subgrid

Step 5. Adjusting rotation and skew of subgrids

Step 6. Save grid file

To zoom in or out of the views

Before you optimize the grid placement, it may be necessary to zoom the image in order to see the spots more easily. There are two ways to zoom the image:

- From the Feature Extraction menu bar, select **View > Zoom** and choose a % zoom from 10% to 1000%. A 10% zoom is defined and you can select a zoom level in 10% increments. You can also change the default zoom settings. See **“Changing default settings for Grid Editing”** on page 114 for details.
- On the image toolbar, click one of the zoom icons.

To undo or redo grid position movements

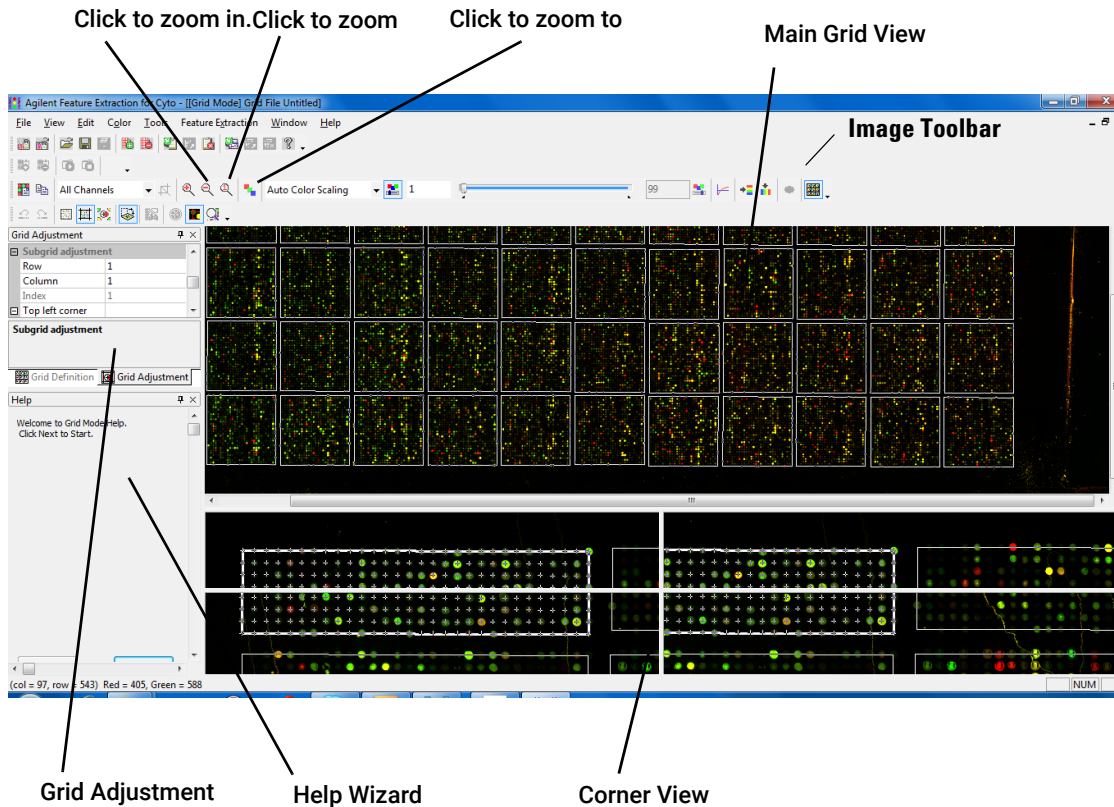





Figure 34 Grid mode view

To undo or redo grid position movements

- Click the Undo button, , to undo an action.
- Click the Redo button, , to redo an action.

Step 1. Get nominal spot spacing manually

This step is optional, and is only used for third-party arrays. Perform this step if spacing information for spots are unknown.

- 1 Make sure that the color log scale  has been turned on to view the corner spots on the main grid more easily.
- 2 On the gridding toolbar, click the View and Measure Image icon .
- 3 If necessary, zoom the image. See **“To zoom in or out of the views”** on page 95.
- 4 On this image view (which is appended with [CROPPED] in the image name), click on the center of the top left spot and drag the mouse horizontally to the center of the spot in the next column. (Do not release the mouse button.) Read and note the column distance from the status bar. This is the left-hand number shown in parentheses; for example, (16x0). (You will need this value later.)

Step 2. Get nominal subgrid spacing manually

- 5 On the image view, click on the center of the top left spot and drag the mouse vertically to the spot center in the first column of the second row. (Do not release the mouse button.) Read and note the row distance from the status bar. This is the right-hand number shown in parentheses; for example, (0x17). (You will need this value later.)
- 6 Close the cropped image view.
- 7 In the grid mode view, click the Grid Definition tab.
- 8 In the Grid Definition panel, double-click on the space next to Col under Nominal Spot spacing, and enter the column distance value from **step 4**.
- 9 In the Grid Definition panel, double-click on the space next to Row under Nominal Spot spacing, and enter the row distance value from **step 5**.

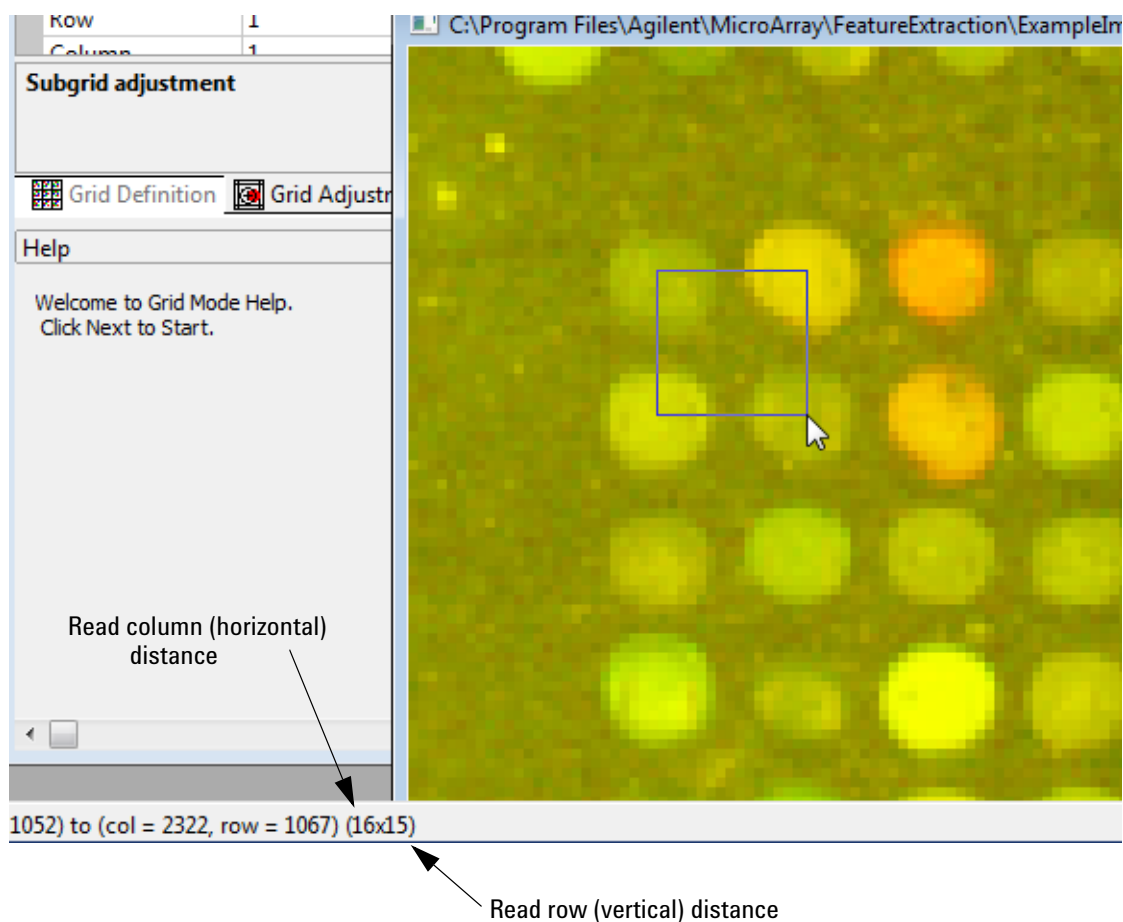




Figure 35 Reading column or row distance from image view

Step 2. Get nominal subgrid spacing manually




This step is optional and only for third-party arrays. Perform this step if spacing for the subgrids is unknown.

- 1 Make sure that the color log scale  has been turned on to view the corner spots on the main grid more easily.
- 2 On the gridding toolbar, click the View and Measure Image icon .

Step 3. Adjust the position of the main grid

- 3 If necessary, zoom the image. See **“To zoom in or out of the views”** on page 95.
- 4 On the newly-opened image, click on the center of the top right spot for the first subgrid and drag the mouse horizontally to the center of the first spot in the adjacent subgrid. (Do not release the mouse button.) Read and note the column distance from the status bar. This is the left-hand number shown in parentheses; for example, (43x0). (You will need this value later.)
- 5 On the opened image view, click on the center of the bottom left spot of the first subgrid and drag the mouse vertically to the center of the top left spot of the subgrid below. (Do not release the mouse button.) Read and note the row distance from the status bar. This is the right-hand number shown in parentheses; for example, (0x42). (You will need this value later.)
- 6 Close the opened image view.
- 7 In the grid mode view, click the Grid Definition tab.
- 8 In the Grid Definition panel, double-click on the space next to Col under Nominal Subgrid spacing, and enter the column value from **step 4**.
- 9 In the Grid Definition panel, double-click on the space next to Row under Nominal Subgrid spacing, and enter the row value from **step 5**.

Step 3. Adjust the position of the main grid

- 1 Make sure that the color log scale  has been turned on to view the corner spots on the main grid more easily.
- 2 On the toolbar, click the Adjust Main Grid icon .
- 3 Move the mouse into the upper Main Grid View until the “move” cursor appears. .
- 4 Click and drag the “move” cursor so that the top left corner of the grid is positioned on the top left spot of the microarray image.
- 5 To do fine adjustment of the main grid, click on the top left spot in the top left pane of the lower Corner View pane.

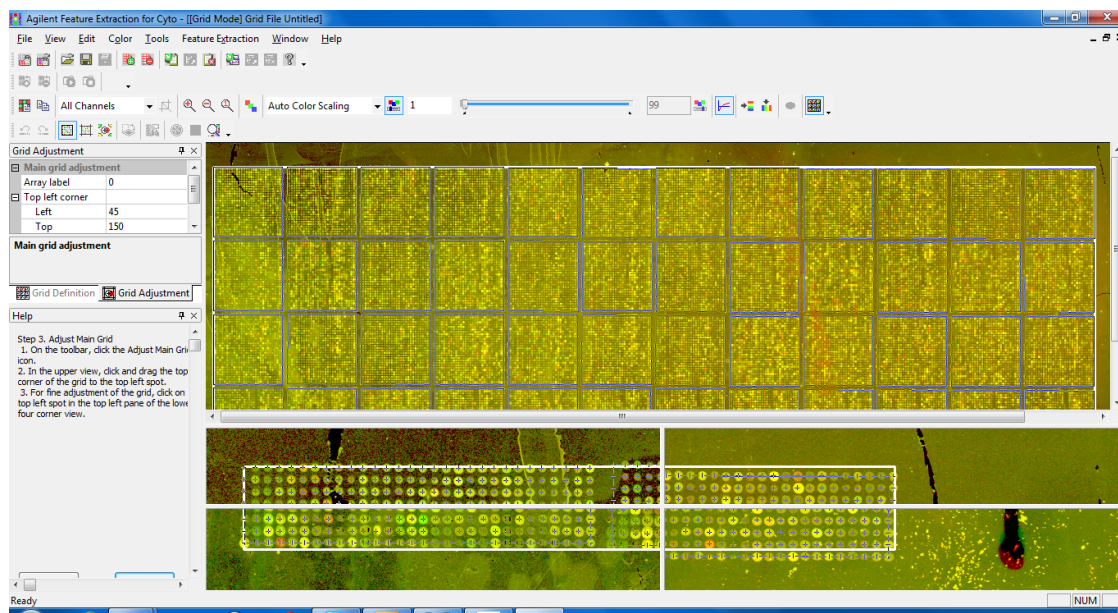




Figure 36 Adjust main grid view - *Note that the thick line represents the main grid and the thin line represents the subgrids.*

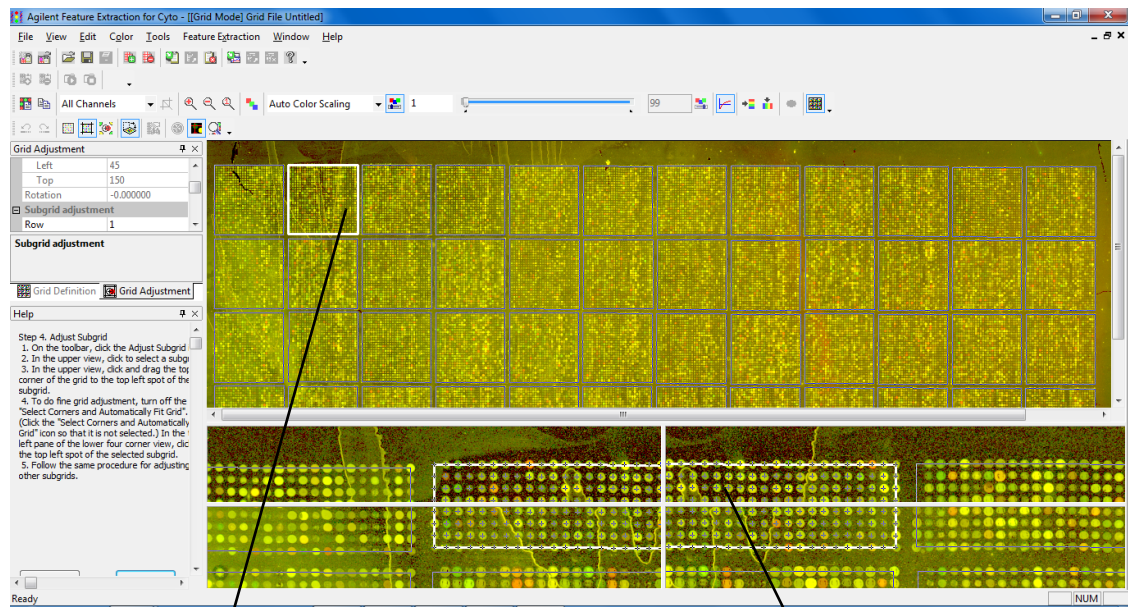
Step 4. Adjust the subgrid

Step 4. Adjust the subgrid

For a microarray with only one grid, if the main grid is well-positioned, there is no reason to move the subgrid. This is the case for most Agilent microarrays.

When you position subgrids you are, in effect, positioning nominal spot centroids (+). These nominal spot centroids are still subject to adjustment by the spot finding algorithm in Feature Extraction. You can position the subgrids such that the nominal spot centroids are just inside the spots. These positions are good enough for the spot finding algorithm to calculate more accurate spot centroids.

- 1 On the toolbar, click the Adjust Subgrid icon. 
- 2 In the Main Grid View, click on a subgrid.
- 3 In the Main Grid View, click and drag the top left corner of the subgrid to the top left spot of the subgrid.
- 4 For fine adjustment of the grid, click the Select Corners and Automatically Fit Grid icon on the toolbar to turn it off. 
- 5 In lower Corner View, click the center of the top left spot of the selected subgrid.
- 6 Repeat **step 2** through **step 5** for the remaining subgrids.





Subgrid Main Image View

Subgrid Corner View


Figure 37 Adjust subgrid view

Step 5. Adjusting rotation and skew of subgrids

*When **Select Corners and Automatically Fit Grid** is enabled, the software will automatically calculate a fit based on three user-selected corner points.*

- 1 In the Main Grid View, click on a subgrid to select it.
- 2 On the toolbar, click to select the Select Corners and Automatically Fit Grid icon, if it is not on. 
- 3 In the lower view, if the corners of the grid are not near to the corner spots of the image, adjust the grid location by clicking and dragging the grid in the bottom right pane of the lower four corner view.
- 4 Move the cursor to the top left pane of the lower Corner View. Notice the cursor turns to a red cross.  Click on the top left spot of the subgrid.
- 5 In the top right pane of the lower Corner View, click on the top right spot (in the first row) of the subgrid.
- 6 In the bottom left pane of the lower Corner View, click on the lower left spot (in the first column) of the subgrid.
- 7 Repeat **step 3** through **step 6** for the remaining subgrids.

Step 6. Save grid file


- 1 Click the Save Grid button on the main toolbar. 
- 2 The grid is saved to a ImageName_grid.csv file where ImageName is the name of the image associated with the grid, and the directory containing the grid file is the same as the directory of the original image. A features file is saved at the same time as a ImageName_feat.csv file.

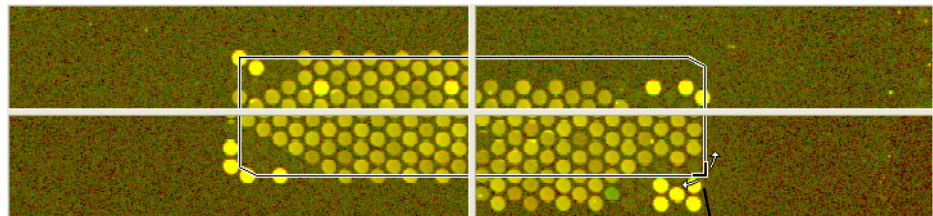
Manual grid adjustments

For most images, grids are adjusted adequately if you follow the steps provided by the Help Wizard. However, if the placement of the grid is not adequate after performing the steps of the Help Wizard, you can correct the grid placement using the following procedures.

To rotate the position of the main grid

To move or rotate the main grid another way, see "To move or rotate the main grid with Grid Adjustment pane" on page 102.

- 1 Make sure that the color log scale has been turned on to view the corner spots on the main grid more easily.
- 2 Click the Adjust Main Grid button on the toolbar. 
- 3 Move the mouse to a corner of the main grid in the lower Corner View.
The "rotate" cursor appears.
- 4 Drag the "rotate" cursor to rotate and optimize the position of the grid on the four corners of the microarray image.



Drag "rotate" cursor.

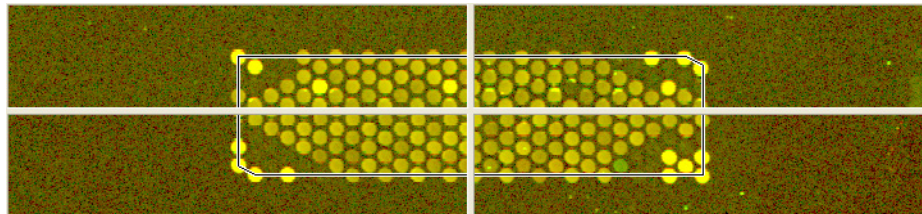


Figure 38 Final position of main grid on microarray

To move or rotate the main grid with Grid Adjustment pane

When the Main Grid View appears, the Grid Adjustment tab is the default selection. You can move or rotate the main grid by entering values for fields in the Grid Adjustment tab.

To learn all the ways that you can move grids and spot centroids, see [“Shortcuts to help you work with grids and find spots”](#) on page 115.

- 1 Click the Grid Adjustment tab, if necessary.

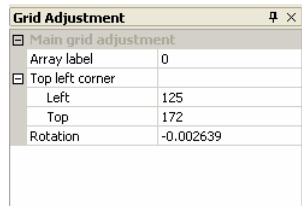



Figure 39 Grid Adjustment pane for Main Grid View

- 2 Click the Adjust Main Grid button on the toolbar. 
- 3 Enter a value for any field that is available to you.



Array Label This number identifies which of the Agilent microarrays on the slide you are looking at.

Top Left Corner Enter pixel locations for the top left corner.

Rotation You can change the number here to rotate this corner of the main grid either in the positive or negative direction.

To rotate the subgrid

If moving the position of the subgrid does not work, try rotating the subgrid.

- 1 Click to select the Adjust Subgrid button on the toolbar. 
- 2 Click to deselect the Select Corners and Automatically Fit Grid button on the toolbar (if it is turned on). 
- 3 Click on a subgrid in the upper Main Grid View.
- 4 In the lower subgrid Corner View, move the mouse to a corner of the subgrid until the “rotate” cursor appears.
- 5 Drag the “rotate” cursor on the subgrid Corner View to optimize the position of the grid on the four corners of that portion of the microarray image.

To skew the subgrid corner

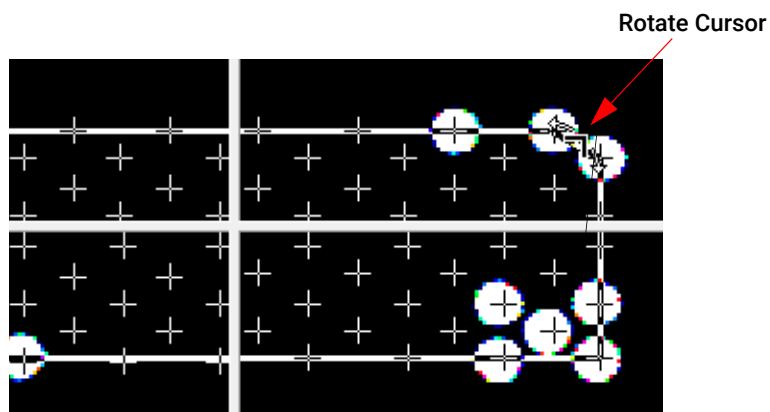



Figure 40 Subgrid with rotate cursor

To skew the subgrid corner

If neither of the above methods work, try skewing the subgrid corner.

- 1 Click the Skew Subgrid button, .
- 2 Move the mouse to a corner of the subgrid Corner View.
The "skew" cursor appears.
- 3 Drag the "skew" cursor on the subgrid Corner View to optimize the position of the grid on the four corners of that portion of the microarray image.

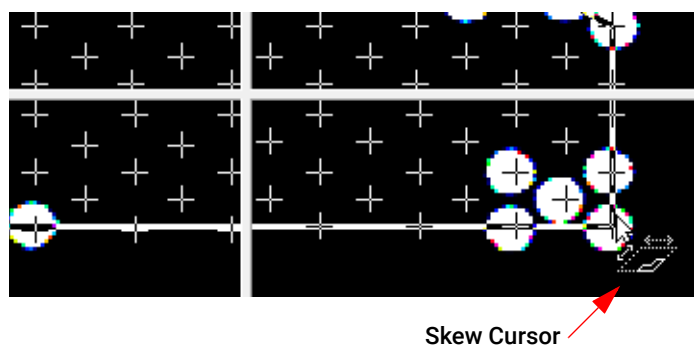



Figure 41 Subgrid with skew cursor

To skew the subgrid corner

To position subgrids with Grid Adjustment pane

When the Adjust Subgrid button is selected , the Grid Adjustment tab is the default selection. You can position subgrids by entering values for fields in the Grid Adjustment tab.

To learn all the ways that you can move grids and spot centroids, see [“Shortcuts to help you work with grids and find spots”](#) on page 115.

- 1 Click the Grid Adjustment tab, if necessary.

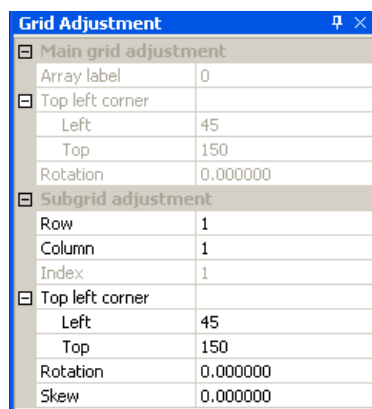


Figure 42 Grid Adjustment pane for subgrids

- 2 Enter a value for any field that is available to you.

Col and Row Numbers

Click the down arrows for Col and Row to select the column and row numbers for the subgrid you want to position.

Corner of Subgrid

Enter pixel locations for the top left corner of a subgrid.

Rotation


Enter a number to rotate the top left corner of the subgrid.

Skew

Enter a number to skew the top left corner of the subgrid.

Saving a grid file



Whenever you have made manual adjustments to the grid or spacing geometry, you must save the grid file in order to keep the changes you have made. There is only one grid file for any given image.

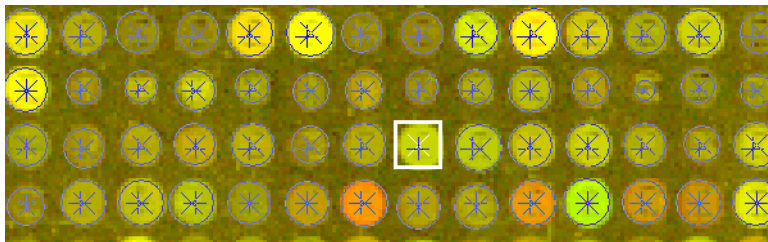
- 1 Click the Save Grid button on the main toolbar. 
- 2 The grid is saved to a ImageName_grid.csv file where ImageName is the name of the image associated with the grid, and the directory containing the grid file is the same as the directory of the original image. For non-Agilent images, a features file is saved at the same time, with the name ImageName_feat.csv.

Calculating spot size and spot centroids

This option is available only for non-Agilent images.

This is the same algorithm used for calculating spot size and finding spot centroids by the software during an automatic extraction. For third-party arrays, running this will show you how well the algorithm will find the spots on the array and let you perform corrections, if required (see **“Locating spot centroids manually”** on page 107).

- 1 Click the Adjust Spot button  to enable spot mode.
- 2 Click the Calculate Spot Size and Centroids button  to let you see the difference between the nominal spot centroid, positioned from the grid (+), and the estimated spot centroid based on the spot size (X).



After you have calculated the spot size and spot centroids, you can make adjustments manually. See **“Locating spot centroids manually”** on page 107. When you move a grid or subgrid after calculating the spot size and spot centroids, the found centroids for the moved grid or subgrid are deleted.

Locating spot centroids manually


If you want to save the manually-placed spots, you must save the grid file.

Once you have calculated the spot sizes and centroids (see **“Calculating spot size and spot centroids”** on page 106) you can manually adjust those. In this section you learn how to move a nominal spot centroid, placed on a spot by the subgrid. When you move the nominal spot centroid (+), it becomes an estimated spot centroid (X). Once the spot centroid is estimated, the software uses this estimated value for Feature Extraction.

You cannot move spot centroids for Agilent images. This function works only for non-Agilent images.

If a spot centroid has been manually moved, the result called IsManualFlag is set to true. See the *Feature Extraction for CytoGenomics Reference Guide* for the result.

To find a spot to adjust

- 1 Click the Adjust Spot button, , to bring up the Main Spot View and the Zoomed Spot View.

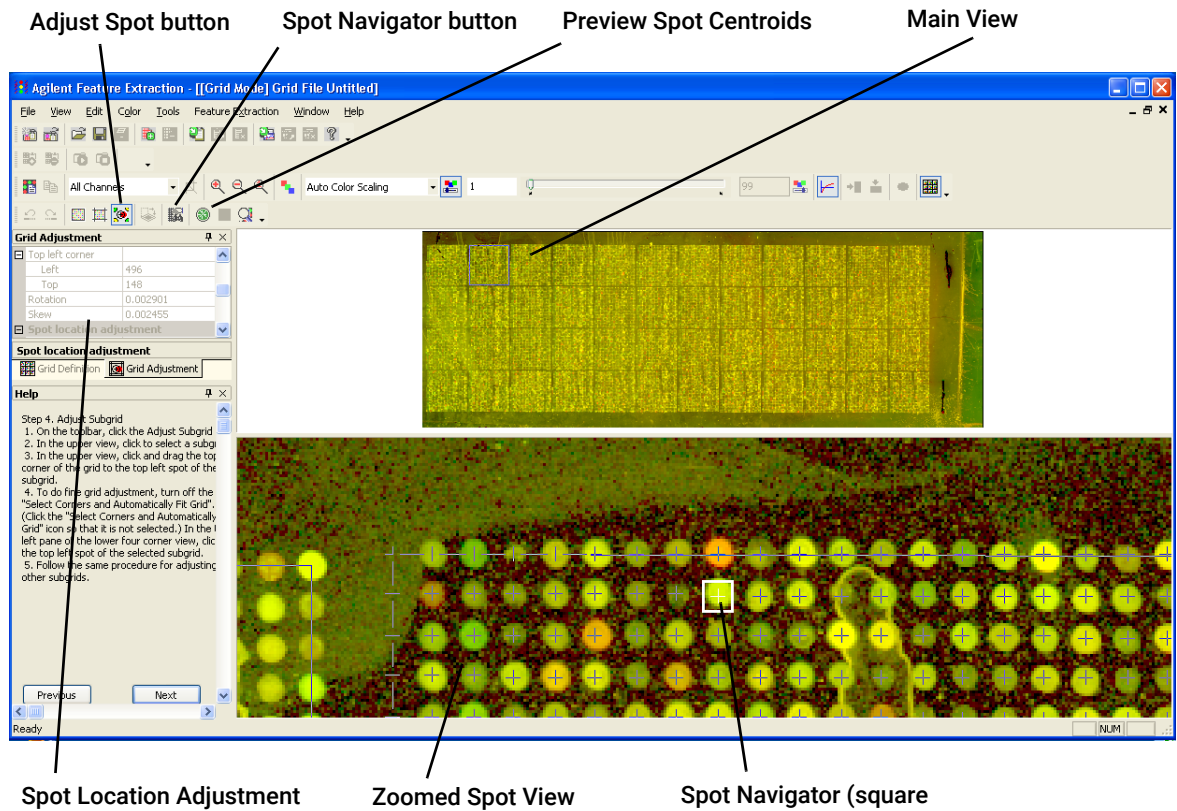



Figure 43 Adjust spot view

To find a spot to adjust

The calculated spot centroids appear when you enter Adjust Spot mode.

- 2 Select **Color > Use Log Color Scale** if this has not already been selected.
- 3 Click a spot on the Main Spot View or on the Zoomed Spot View.
Or
Click the Spot Navigator button, , to bring up the Grid Spot Navigator dialog box.

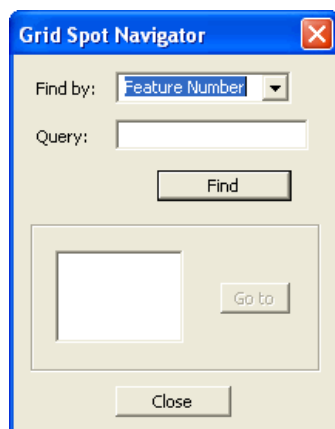


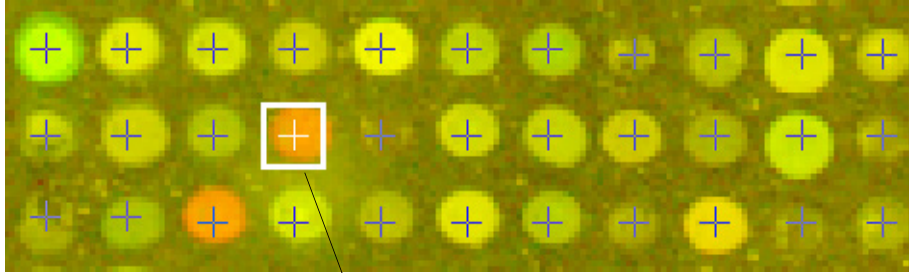
Figure 44 Grid Spot Navigator dialog box

- 4 Click the down arrow in the Find by list, and select the number or name of the spot to navigate to.
- 5 In the Query field, type the feature number or probe name that you want to find.
Case is important.
- 6 Click **Find**.
If only one feature in the microarray matches the query, the software moves the spot square (Spot Navigator) to that feature.
If more than one feature matches, a list of features is displayed in the bottom box.
- 7 Select a feature from the list, and click **Go to**.
The software moves the Spot Navigator to that feature.
- 8 Click **Close**.
The software positions the Spot Navigator (square box) in the center of the Zoomed Spot View.

To ignore spots

To position the spot centroid

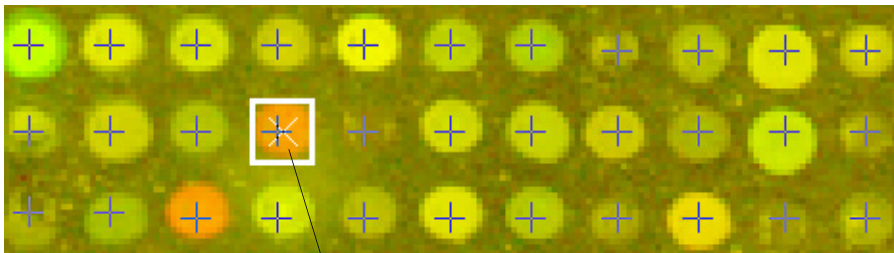
Before you position the Spot Navigator to find the centroid, the position of the nominal spot centroid could look like this:



Nominal spot centroid (+) to be moved

- Move the Spot Navigator square containing the nominal spot centroid position to the assumed center of the spot.

The resulting position might look like this:



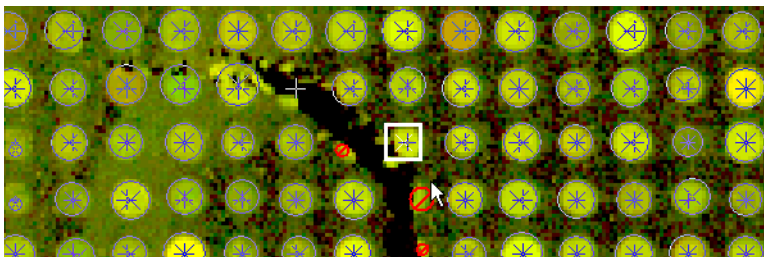
New estimated spot centroid (X) position;

These images represent non-Agilent microarrays. You will more likely need to position the spot centroid for non-Agilent microarrays.

When this altered grid file is used in Feature Extraction, the default spot finding algorithm will not attempt to move spots that have been manually located in the grid file.

To ignore spots

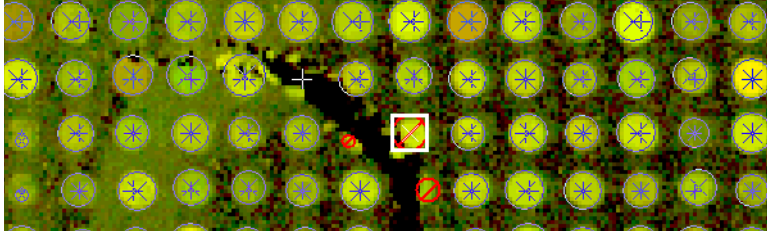
- 1 Position the Spot Navigator where you want to ignore a spot.





- 2 In the Grid Adjustment pane, click the field next to Ignore and then select **True**.

To undo or redo spot position movements

The software will not use this spot for any step in Feature Extraction.



To undo or redo spot position movements

- Click the Undo button, , to undo an action.
- Click the Redo button, , to redo an action.

To position spot centroids with the Grid Adjustment pane

When the Spot View appears, the Grid Adjustment tab is the default selection. You can position spot centroids by entering values for fields in the Grid Adjustment tab.

- 1 Click the Grid Adjustment tab, if necessary.

Grid Adjustment	
Main grid adjustment	
Array label	0
Top left corner	
Left	125
Top	172
Rotation	-0.002639
Subgrid adjustment	
Row	2
Column	2
Index	6
Top left corner	
Left	575
Top	623
Rotation	-0.000220
Skew	0.003636
Spot location adjustment	
Row	1
Column	1
Feature Number	4501
Center	
X	575
Y	623
Diameters (in pixels)	
Horizontal	11
Vertical	11
Probe name	Probe_4501
Gene name	
Ignore	False
Spot location adjustment	

Figure 45 Spot Location Adjustment pane

- 2 Enter a value into any field that is available to you.
- 3 Press **Enter**.

To learn all the ways that you can move grids and spot centroids, see [“Shortcuts to help you work with grids and find spots”](#) on page 115

To position spot centroids with the Grid Adjustment pane

Col and Row Numbers	Click the down arrows for Col and Row to select the column and row numbers for the spot centroid you want to position.
Center	Enter pixel numbers on X and Y axes for spot centroid location.
SpotDiam	Enter distance along X and Y axes in pixels for spot diameter.
ProbeName	Name of probe associated with selected spot; found in gene list
GeneName	Name of gene associated with selected spot; found in gene list
Ignore	Select true to ignore selected spot, or false to include it.

Refitting a grid

You may want to refit a grid under the circumstances mentioned in the table below. Alternate solutions to the problems are also described:

Table 4 Conditions for refitting grids

Grid Issue	Things to try before refitting the grid
Main grid is not properly placed.	Adjust the placement of the main grid. See “Step 3. Adjust the position of the main grid” on page 98.
Subgrids are not correctly spaced.	Adjust the spacing between subgrids. Assumption is that the subgrids are evenly spaced. See “Step 2. Get nominal subgrid spacing manually” on page 97 and “Step 4. Adjust the subgrid” on page 99.
Spot spacing appears off.	Adjust the spacing between the spots. “Step 1. Get nominal spot spacing manually” on page 96.
Average spot size is not correct.	Adjust the spot diameter. This diameter is used by the Calculate Spot Size and Centroids algorithm and the WholeSpot algorithm. See “Calculating spot size and spot centroids” on page 106 and “Locating spot centroids manually” on page 107.

CAUTION

When you refit a grid, any changes you made during manual gridding, if not saved as a grid file, are lost.

To refit a grid, follow these steps.

1 Select **Edit > Refit Grid**.

The Geometry Information dialog box opens.

To learn about Geometry settings, see **“Geometry information”** on page 91.

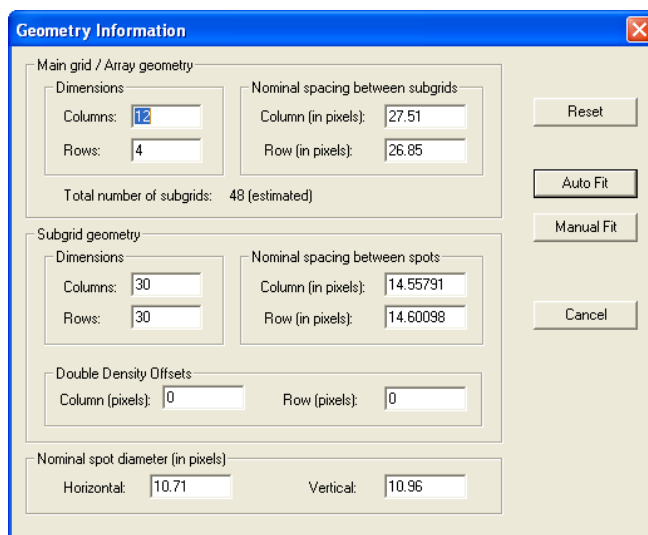


Figure 46 Geometry Information dialog box to refit Agilent grid

Refitting a grid

- 2 Change entries, and click **Auto Fit** or **Manual Fit**. A message will appear reminding you that refitting the grid will cause you to lose any changes you made that have not been saved in the grid file.
- 3 Click **Yes** to continue, or **No** to return to the Geometry Information dialog.
Click **Reset** to re-enter the original number before you changed it.
- 4 Once the grid has been refit, save the grid file to keep the changes made by the refitting procedure.

See "**Saving a grid file**" on page 105.

Changing default settings for Grid Editing

Before you begin to set up a grid, you can change the default settings that appear in the Grid Definition and Grid Adjustment panels on the main window when the grid mode is turned on.

- 1 Select **Tools > Preferences**.
- 2 Click the Grid Editing folder.

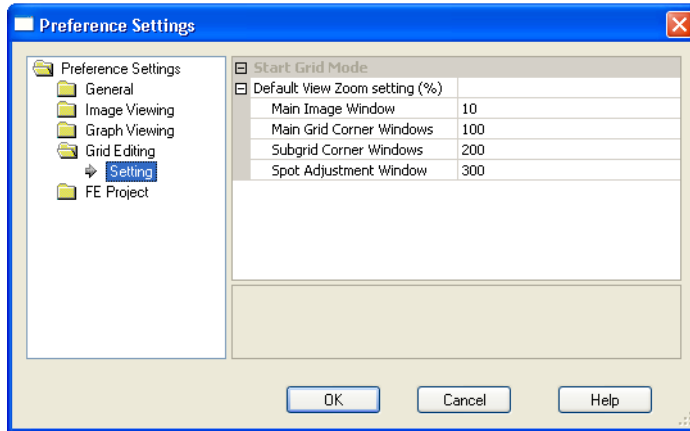


Figure 47 Preference Settings—Grid Editing selected

Default View Zoom setting

Click the cell next to an image window, then click the down arrow to select the zoom setting for each of the image windows listed.

Shortcuts to help you work with grids and find spots

There are several ways that you can select and move grids and spots. These are presented in [Table 5](#). In addition, there are a number of Ctrl and Alt key operations that help you work with grids. These are presented in [Table 6](#).

Table 5 Grid Actions

Location of actions	If you want to do this:	Do this on the Grid Adjustment dialog box:	Or, do this with the keyboard:	Or, do this with the mouse:
Main Grid	Set array label	Edit value, then press Enter .	NA	NA
	Move grid	Edit value, then press Enter .	Hold down the Shift key and select an arrow key.	Click the left mouse button and drag the cursor across the full window. Or, click the left mouse button in the corner window.
	Rotate grid	Edit value, then press Enter .	NA	Drag on corner in corner window.
Subgrids	Select subgrid*	Select column and row number from a choice list.	Use arrow keys or Page Up/Page Down (requires two rows or more) or Home/End	Left-click mouse in full view or corner view.
	Move subgrid	Edit location, then press Enter .	Hold down the Shift key and select an arrow key.	Drag the cursor across the full window. Or, click the left mouse button on the top left spot of the subgrid in the corner window.
	Rotate subgrid	Edit value, then press Enter .	NA	Drag on corner in corner window. Or, turn on automatic skew and rotate and then click the left mouse button on the upper left, upper right, and lower left spot of the subgrid in the corner window.
	Skew subgrid	Edit value, then press Enter .	NA	Click Skew Subgrid button, then drag on corner in corner window. Or, turn on automatic skew and rotate and then click the left mouse button on the upper left, upper right, and lower left spot of the subgrid in the corner window.
Spots	Select spot*	Select column and row number from a choice list.	Use arrow keys or Page Up/Page Down or Home/End	Left-click mouse in full view or corner view.
	Move spot centroid	Edit location, then press Enter .	Hold down the Shift key and select an arrow key.	Drag Spot Locator to position spot centroid.
	Flag spot	Mark the Ignore this spot check box, then press Enter .	NA	Double-click the right mouse button to toggle this flag.

* These are the only grid actions that you cannot undo, redo or reset.

Table 6 Hot Key Actions for Grid Mode

If you want to do this:	Press this "hot key":
Turn Grid Mode On/Off	Alt-G
Select for Main Grid adjustment	Alt-M
Select for Subgrid adjustment	Alt-S
Select for Spot adjustment	Alt-P
Select to ignore a spot	Ctrl-G
Undo	Ctrl-Z
Undo All	Ctrl-A
Redo	Ctrl-Y
Redo All	Ctrl-D
Refit grid	Ctrl-T

4 Changing Protocol Settings

Changing protocol settings	118
Creating a new protocol from an existing protocol	119
Changing Protocol Step 1: Place Grid	124
Changing Protocol Step 2: Optimize Grid Fit	128
Changing Protocol Step 3: Find Spots	130
Changing Protocol Step 4: Flag Outliers	136
Changing Protocol Step 5: Compute Bkgd, Bias and Error	140
Changing Protocol Step 6: Correct Dye Biases	149
Changing Protocol Step 7: Compute Ratios	153
Changing Protocol Step & Calculate Metrics	157
Changing Protocol Step 9: Generate Results	159

This chapter gives you background information on changing settings for the protocols used to perform Feature Extraction. The Feature Extraction protocols contain the parameter values for the algorithms that you specify for each step of the Feature Extraction process.

See Chapter 1 in the *Feature Extraction for CytoGenomics Reference Guide* for a listing of the default settings and parameter values for the protocols shipped with the software.

Changing protocol settings

Protocol templates

A protocol template for CGH extractions is provided with Agilent Feature Extraction for CytoGenomics. The name of protocol templates follow the format:

Agilent_microarray_type_software_version_date

For a complete list of the default protocol settings, see the *Feature Extraction for CytoGenomics Reference Guide*.

Both the protocol parameters and default values may change over time. To ensure you have the current protocol, from the Feature Extraction main window, click **Help > Feature Extraction on the Web > Feature Extraction Protocols**.

To create a new protocol from a protocol template

See Chapter 5, "How Algorithms Calculate Results" in the Feature Extraction for CytoGenomics Reference Guide for detailed information on how the protocol steps and algorithms work.

Each protocol contains protocol steps that represent Feature Extraction algorithms. In the FE Protocol Editor, you make changes to the parameter values for each of the steps.

After you open the FE Protocol Editor for a specific protocol, you can change settings, but you must then save the protocol to a new name. The process is as follows:

- 1 To open a protocol.**
- 2 To view or change the protocol properties.**
- 3 To change the protocol steps to run.**
- 4 To change the settings for each protocol step.**
- 5 To save a protocol with a new name.**

The following sections present instructions for creating protocols from templates and tasks for changing settings.

Creating a new protocol from an existing protocol

To create a protocol for a specific type of microarray, you *must* use an existing protocol for the same type of microarray as a template.

For example, if you want to create a new protocol to run a gene expression 2-color microarray, start by editing an Agilent-created protocol for gene expression 2-color microarrays.

You can track what the original Agilent-created protocol is by looking at the Derived From field of the general section of the new protocol. See **“To change the protocol steps to run”** on page 121.

CAUTION

Existing protocols can be used as templates for new protocols to provide both visible and hidden settings whose values are specific to the type of microarray. Although you can change the visible settings of any two protocols of different types to *appear* identical, you cannot change the hidden settings that distinguish these protocols from one another.

To open a protocol

You can open the FE Protocol Editor in one of three ways:

- In the Protocol Browser double-click the protocol that you intend to edit, or right-click the protocol and select **Properties**.
- In Project Explorer, either double-click the protocol assigned to an extraction set, or right-click the protocol and select **Properties**.
- Click the Extraction Configuration tab, and in the Protocol column double-click the protocol that you intend to edit.

Regardless of the protocol, the first screen that appears is by default the screen that was last displayed before a protocol was closed or saved.

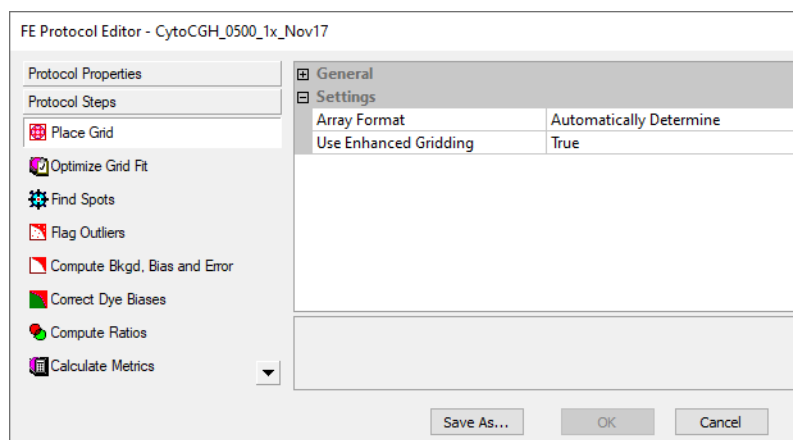


Figure 48 FE Protocol Editor – Protocol Steps

To open a protocol

To view or change the protocol properties

The General section of the Protocol Properties page lists fields to identify the protocol and fields to let you remove the protocol or make the protocol read-only.

- If necessary, in the FE Protocol Editor click the Protocol Properties button to see the General section.

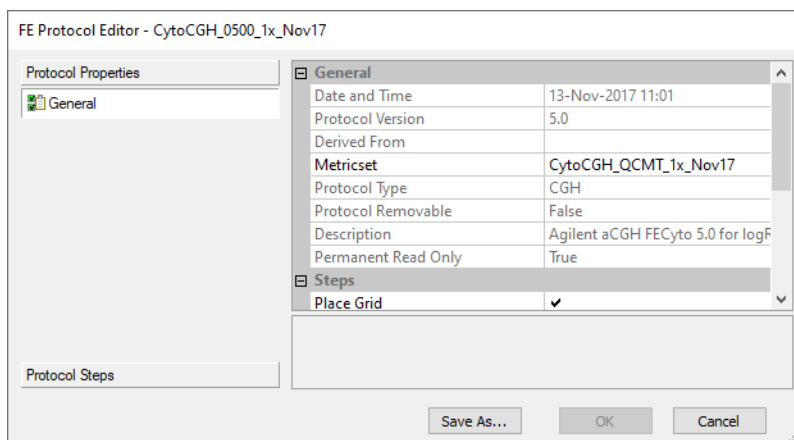


Figure 49 Protocol Editor - Protocol Properties tab

- Date and Time** Timestamp for the creation of the protocol.
- Protocol Version** Version of the original Agilent-created protocol, corresponding to software version with which it was shipped.
- Derived From** Name of original Agilent-created protocol from which the current protocol is derived; blank until you exit and reopen modified protocol.
- Metricset** Name of the QC metric set that is to be used with this protocol to generate metrics (statistics) shown in the QC Report and Stats table.
- Protocol Type** Corresponds to type of microarray for which the protocol is to be used for Feature Extraction: 2-color gene expression, 1-color gene expression, CGH, ChIP, or miRNA data.
- Protocol Removable** Describes if the protocol is removable from the database. User-created protocols are removable. Agilent-created protocols are not removable.
- Description** Description of the type of microarray and format for which the protocol was designed
- Permanent Read Only** True indicates that the protocol cannot be edited. Any changes made to the protocol must be saved with a different name.
False indicates that the protocol can be edited. Any changes made to the protocol can be saved with the existing name.

To change the protocol steps to run

If there is no protocol on your system with the steps you want Feature Extraction to use, you can open a protocol of the appropriate type and change the protocol steps.

- 1 If necessary, in the FE Protocol Editor click the Protocol Properties button to see a list of the protocol steps on the right.
- 2 Mark each check box to enable a step that you want to run. Clear each check box to disable a step that you do not want to run.

Refer to **Figure 48** to see the FE Protocol Editor.

Protocol steps for CGH images

Below is a description of each of the protocol steps for CGH, images, as seen in **Figure 48**, in the order that the software runs them.

- | | |
|--------------------------------------|--|
| Place Grid | <ul style="list-style-type: none">• For Agilent microarrays using a grid template, automatically places a grid on the microarray.• For any microarray loaded with a grid file that you saved during Grid Mode (i.e. interactive gridding), automatically calculates centroid positions. Any spots that were manually adjusted are left unchanged.• For Agilent microarrays, the software performs a background peak shift that helps find peaks of features that are close to the background. See the <i>Feature Extraction for CytoGenomics Reference Guide</i> for details.• For any microarray, user can select to use regions near to the center of packs for slope and skew calculations. This option is of use in cases where pack edges have dim spots and are failing to grid. When this is selected, the user can select to have the software to perform one extra step of correlation following the projection data analysis to get the origin. |
| Optimize Grid Fit | <ul style="list-style-type: none">• For microarrays with 30- or 65-micron feature size, iteratively adjusts the grid by looking at spots in the corners of the microarray to improve overall image registration. |
| Find Spots | <ul style="list-style-type: none">• The spot centers are located and/or optimized starting with the spot nominal positions on the grid.• From the located spots, the features are defined with either a CookieCutter method or a WholeSpot method.• The local background is also defined with a radius method.• Outlier pixels (too low or high intensity) are then removed from the pixel populations that defined the features and backgrounds.• Statistics, such as the mean, median and standard deviation (SD), are calculated on the remaining inlier pixels within the feature and background.• Features with a large number of saturated pixels are flagged.• The protocol determines whether the mean or median signals are used for further analysis. |
| Flag Outliers | <ul style="list-style-type: none">• Features are flagged as non-uniformity outliers if their pixel distributions are outside the threshold limits.• Features that have replicates of the same probe sequence and background regions are flagged as population outliers if their signal distributions are outside the set range. |
| Compute Bkgd, Bias, and Error | <ul style="list-style-type: none">• If background subtraction is selected, background signal is subtracted from raw signal for each feature using a local or global background method. |

To open a protocol

- A statistical test is performed to assess if the feature signal is significantly different from the background signal. This test uses either the error model or pixel statistics, as indicated in the protocol choice "Significance".
- Raw signals or background-subtracted signals can be corrected for systematic gradients, that affect low signals especially, by using additive detrending (i.e. spatial detrending).
- The error is calculated at this point, using the choice between the Universal error model or the larger of the error estimates (i.e. most conservative) from the Universal error model and the propagated error.
- A multiplicative detrending algorithm is also used to correct for gradients proportional to the signal.
- Surrogates, based upon background noise, can be used as processed signals for features that have low-intensities, near the background noise level.

Correct Dye Biases

- The method for how a dye normalization gene list is selected is set.
- Differences in the red and green channel signals caused by different efficiencies in labeling, emission and detection are estimated and corrected.
- Features to be used for measuring dye bias are selected based on a probe selection method chosen.
- A dye normalization curve fit method is selected and used to fit the data for measuring dye bias. (SNP probes are not included in dye normalization calculations.)

Compute Ratios

- Log ratios of the red and green processed signals (i.e. background-corrected and dye-normalized) are calculated, as well as a log ratio error and a p -value for each feature. For gene expression 2-color arrays, these calculations assess the confidence you can have that the gene is or is not differentially expressed.

Calculate Metrics

- In this step, the Stats table is generated and QC metrics are calculated for the QC Report. You must leave this option turned on to generate a complete Stats table and QC report. Some values will be missing if this option is turned off.
- The type of protocol you select determines the type of QC report that is generated. For example, a CGH protocol generates a CGH QC Report, and a 1-color protocol with spike-ins selected generates a 1-color QC Report with spike-ins.

Generate Results

- Results selected for the project are generated. In the Protocol Editor you can select to generate one file for the three tables (FEPARAMS, STATS, FEATURES) of text results or three separate files. You can set the compression for the jpeg file.

To save a protocol with a new name

To change the settings for each protocol step

After you change the protocol steps that you want to run, you may want to change the default settings and parameter values for a specific protocol step.

- 1 Click the Protocol Steps tab to see a list of the protocol steps, or Feature Extraction algorithms, on the left.
- 2 Click the protocol step button whose settings you intend to change.

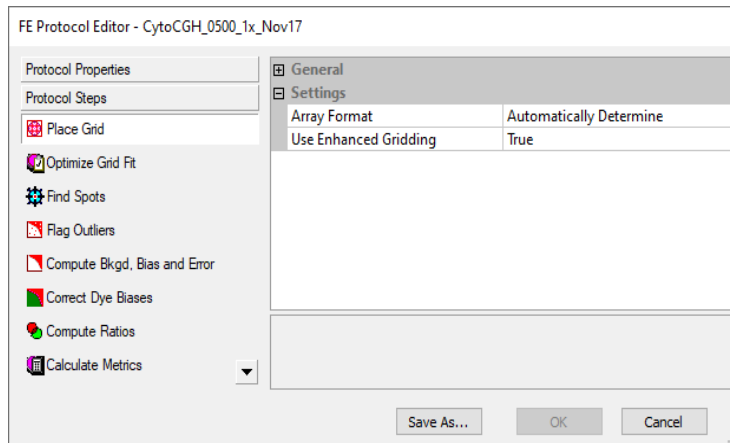


Figure 50 FE Protocol Editor with Protocol Step on left

See the remaining sections in this chapter to learn how and why you change parameters for each protocol step.

To save a protocol with a new name

After you make any change to a protocol, you must save the protocol to a new name.

- 1 Make the changes you intend.
- 2 Click **Save As**.

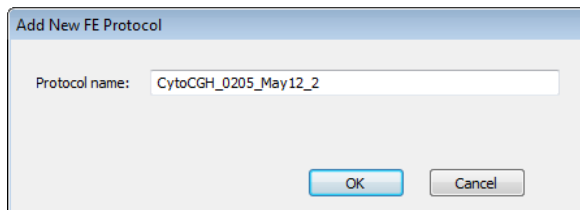


Figure 51 Add new protocol dialog box

- 3 Enter a Protocol name.
- 4 Click **OK**.

Changing Protocol Step 1: Place Grid

The Place Grid algorithm initiates Feature Extraction by placing a grid on the microarray image. The grid comes from either the grid template or the grid file associated with the image file in the project.

To access the settings to place a grid

- Click the Place Grid tab in the FE Protocol Editor, if not already there.

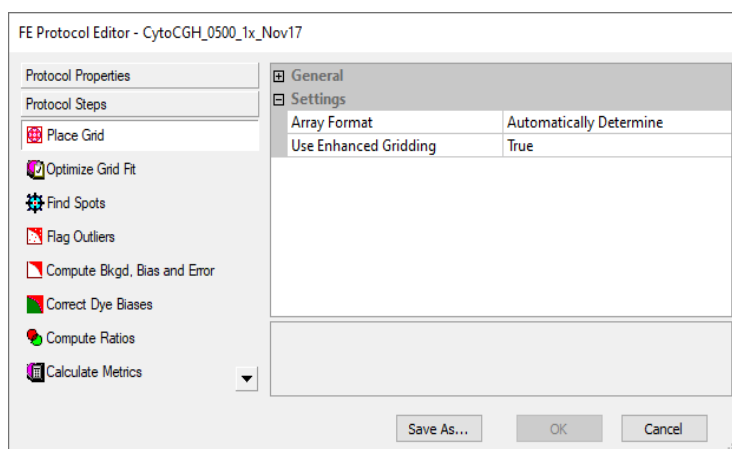


Figure 52 Protocol Editor - Place Grid step

To select the microarray format

This topic describes how to select an array format if needed. You typically do not need to do this step, as the software automatically determines which array format and scan type is used.

Once you change to a specific array format, the protocol loses its ability to be general and will only work for that specific format.

- 1 Click the cell to the right of **Array Format**.
- 2 Click the down arrow, and select a microarray format from the list.

To select the microarray format

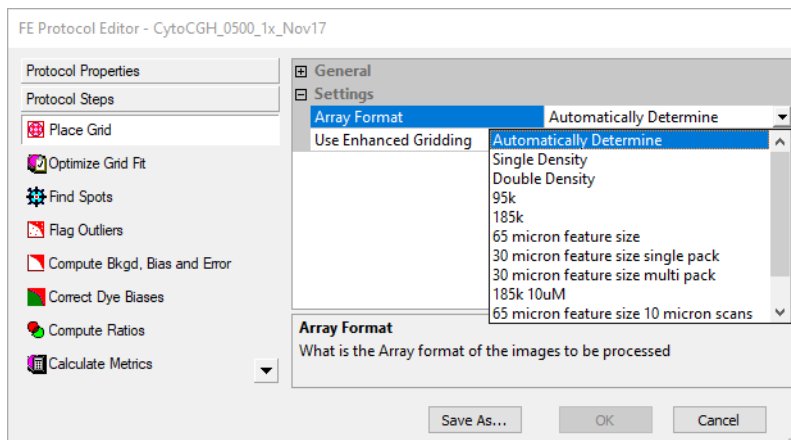


Figure 53 Place Grid - Array Format list

The Placement Method option now appears. See [Figure 54](#) on page 125.

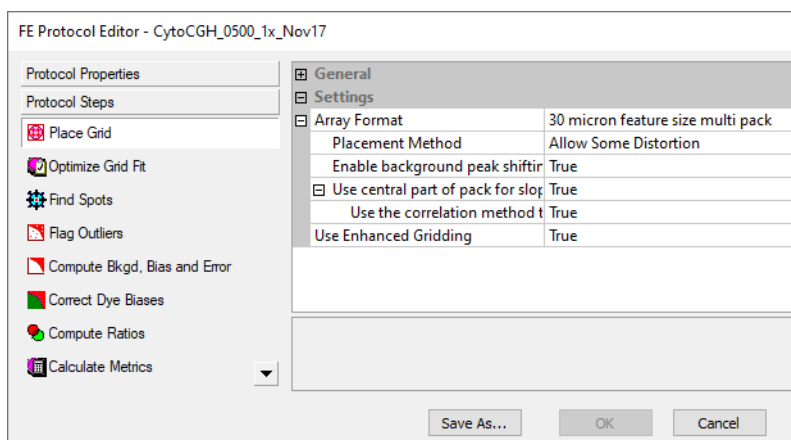


Figure 54 Place Grid - 30 micron multi pack Array Format selected, Placement Method

After you select the format, you can select a method for placing the grid.

Protocols can automatically detect the array format (65- micron feature size or 30-micron feature size). The protocol automatically makes different parameter choices depending on the format and type. These choices are transparent to you and are primarily in the image processing (Place Grid and Find Spots).

With this feature, you do not need different protocols for different array formats. These protocols contain the correct parameter choices for each and use the set that matches the format.

Information about which parameters are affected and which parameters are chosen for each format type can be found beneath any option with a choice "automatically determine". If you believe an incorrect automatic choice was made, you can select a specific format. However, this is not a normal or necessary step.

To select a method for placing the grid

To select a method for placing the grid

- 1 Click the cell next to **Placement Method**, and click the down arrow.
- 2 Select one of the two options listed below.

Allow Some Distortion Allows some variance in the feature spacing, subgrid spacing and the skew. All of the protocol templates have this default setting.

Place and Rotate Only Only allows rotation of the grid and no changes to feature spacing, subgrid spacing or the skew.

To select to use enhanced gridding

As of version 5.0, Feature Extraction for CytoGenomics offers an enhanced gridding algorithm used in the preloaded protocols that was not available in protocols that were created prior to version 5.0. These enhancements were designed to improve automatic gridding in array images of lower quality, including those with non-linear features. Array images that may have previously received a “failed grid” error and require manual gridding can now be processed successfully using the enhanced algorithm.

Importantly, use of the enhanced algorithm may yield changes in some of the QC metrics results and aberration calls compared to the previous algorithm. Validation is recommended.

If you are editing an imported protocol of any type, the enhanced gridding setting is off by default.

- 1 Click the cell next to **Use Enhanced Gridding**, and click the down arrow.
- 2 Select **True** or **False**.

When set to True, the enhanced gridding algorithm that was added in Feature Extraction for CytoGenomics 5.0 is used in the protocol.

When set to False, the enhanced gridding algorithm that was added in Feature Extraction for CytoGenomics 5.0 is NOT used in the protocol. The previous version of the gridding algorithm is used instead.

To select to enable background peak shifting

- 1 Click the cell next to **Enable background peak shifting?** and click the down arrow.
- 2 Select **True** or **False**.

If you select True, the background pixel value is obtained using the histogram of the image. All the pixels in the image that belong to background are reassigned a value of zero. Note that this assignment is temporary and does not change the actual image data for subsequent steps.

To select to use central part of pack for slope and skew calculations

This option lets you choose the region to use for slope and skew calculation during subgrid placement.

- 1 Click the cell next to **Use central part of pack for slope and skew calculations?** and then click the down arrow.
- 2 Select **True** or **False**.

When set to True, regions near to the center of packs are used. This option is of use particularly in cases where pack edges have dim spots and are failing to grid. (This is the default setting for 30 micron single- and multi-pack array formats.)

When set to False, the regions near to the edges of packs are used. Selecting this option uses the same parameters as Feature Extraction version 10.7/10.9 or earlier.

To select to use the correlation method to obtain origin x of subgrids

This option lets you choose the technique to use to determine X location of origin of subgrids.

- 1 Click the cell next to **Use central part of pack for slope and skew calculations?** and click the down arrow.
- 2 Select **True**.
- 3 Click the cell next to **Use the correlation method to obtain origin x of subgrids?** and then click the down arrow.
- 4 Select **True** or **False**.

When set to True, the software performs one extra step of correlation following the projection data analysis to get the origin. This option is of use particularly in cases where pack edges have dim spots and are failing to grid.

When set to False, the results obtained from the projection data analysis are used to estimate the origin. This is how the origin is estimated using Feature Extraction version 10.7/10.9 or earlier.

Changing Protocol Step 2: Optimize Grid Fit

The Optimize Grid Fit step is a post grid placement process to improve the grid registration. Leveraging from the spot finder algorithm, this step examines the spots in the four corners of the microarray to iteratively adjust the grid for best fit.

To access the settings to optimize the grid fit

- Click the Optimize Grid Fit tab in the FE Protocol Editor.

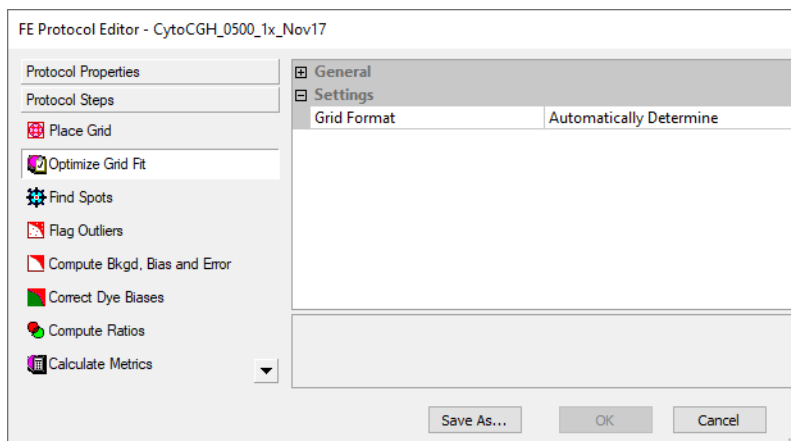


Figure 55 FE Protocol Editor - Optimize the Grid Fit selected

To select the grid format

- 1 Click on the cell next to **Grid Format**.
- 2 Click the down arrow, and select from the list of grid formats.

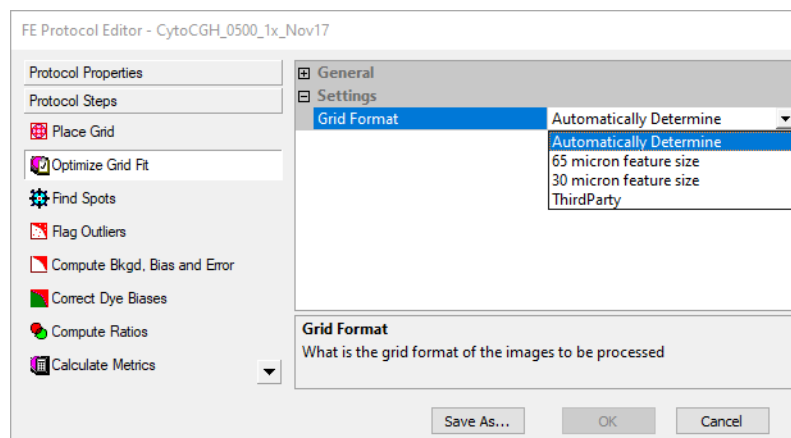


Figure 56 Optimize Grid Fit - Grid Format list

To change the settings to iteratively adjust corners when grid format is selected

If you select either 65-micron or 30-micron feature size, the **Iteratively Adjust Corners?** option appears. See [Figure 57](#).

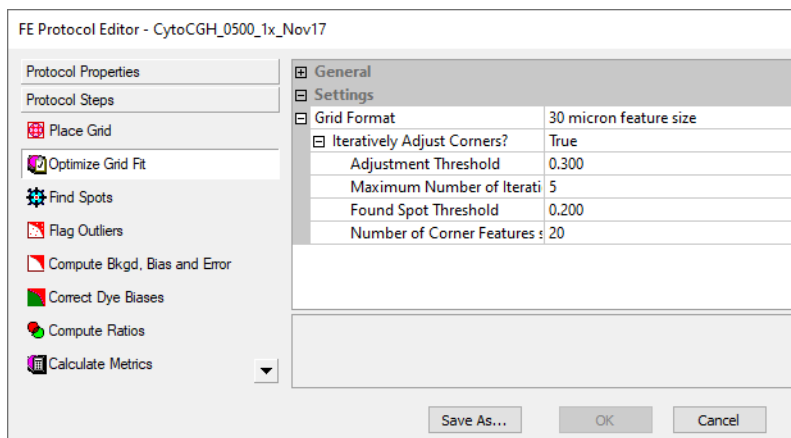


Figure 57 Optimize Grid Fit - 30-micron Feature Size selected, Iteratively Adjust Corners

To change the settings to iteratively adjust corners when grid format is selected

- 1 Click the cell next to **Iteratively Adjust Corners?**.
- 2 Either select **True** or **False**.
- 3 If you select True, you can change any of the default parameters that you need to.

When set to True, the Interactively Adjust Corners method runs the spot finder algorithm in the corners to see how well the grid is placed on the microarray and iteratively adjusts the grid based upon the results.

Adjustment Threshold The grid is adjusted based upon the difference between the grid position and the spot position if the absolute average difference is greater than this value.

Maximum Number of Iterations This is the maximum number of times the spot finder algorithm runs to adjust the grid for better grid fitting in the corner spots.

Found Spot Threshold For each corner, the grid is adjusted only if this fraction or more of the features in that corner are considered found by the spot finder algorithm.

Number of Corner Features Side Dimension? This specifies the number of features (i.e. spots) on each side of the square in each corner of the microarray to be used to calculate the average difference.

The default number is 20 for the 30- and 65-micron feature size microarrays. This means each corner of the microarray has a square area of 400 (i.e. 20 x 20) features that are calculated.

Changing Protocol Step 3: Find Spots

After Feature Extraction places the grid, it is ready to locate the spot centers and define the features and local background regions for each spot. The options available and default values for each setting depend on the microarray format selected in **Changing Protocol Step 1: Place Grid**.

See Chapter 1, “Default Protocol Settings”, of the Feature Extraction for Cyto Reference Guide for a setting comparison between the various microarray formats.

To access the settings to find spots when spot format is automatically determined

- Click the Find Spots tab in the FE Protocol Editor.

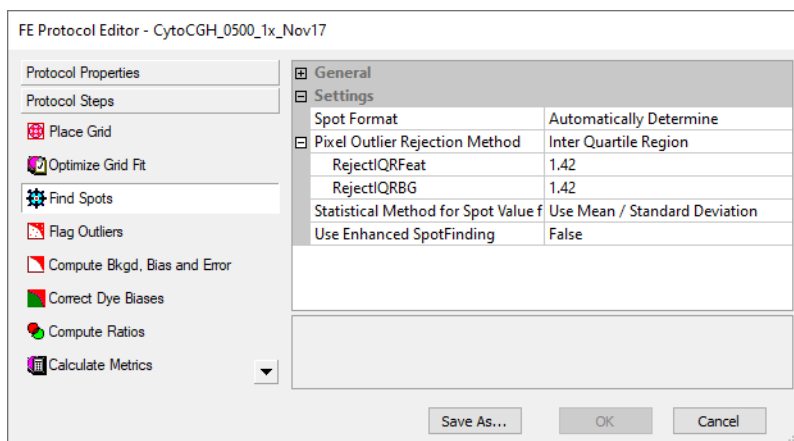


Figure 58 FE Protocol Editor - Find Spots selected

When the spot format is automatically determined (default setting), these are the only settings available for change:

- Select the statistical method to reject outlier pixels from features and local background before Feature Extraction calculates feature and background statistics.
- Select whether to use mean or median for data analysis

A *spot* is a circular or amorphous area on the microarray containing a hybridized sample.

A *feature* is defined as the pixels in either the center portion of a spot or in the whole spot, depending on the method you choose to define the feature.

To select the statistical method to reject pixel outliers

- 1 Click the cell next to **Pixel Outlier Rejection Method**, and click the down arrow.
- 2 Select one of these three choices:

Standard Deviation This calculation rejects outliers outside of a specified number of standard deviations from the mean of the intensity for all pixels. If selected, you can enter the number of standard deviations (Feature SD and Background SD).

Inter Quartile Region This calculation rejects outliers based on the Interquartile Range (IQR). See the *Feature Extraction for Cyto Reference Guide* for how this calculation works. If selected, you can enter the IQR factor for both the feature and background (RejectIQRFeat and RejectIQRBG).

None If selected, no pixels are rejected.

Feature Extraction rejects outlier pixels after calculating statistics on the intensities of all the pixels of a feature or local background region for each color. If a pixel is rejected in one color, it is rejected in both colors.

To select whether to use mean or median for data analysis

- 1 Click the cell next to **Statistical Method for Spot Value from Pixels**, and click the down arrow.
- 2 Select one of these choices to be used in the next step of data analysis:

Mean This selection uses the mean and standard deviation.

Median This selection uses the median and normalized IQR for the next steps of data analysis.

During the background correction stage the background can be subtracted from either the mean signal or the median signal. This option lets you decide which signal type to use for downstream analysis.

To enable the Enhance Spot Finding option

- 1 Click the cell next to **Use Enhanced SpotFinding**, and click the down arrow.
- 2 Either select **True** or **False**.
 - If you select True, then the enhanced spot finding method for is used during the spot finding step of the protocol. This method allows for more accurate identification of the center of each spot and of the pixels that comprise each spot. The result is that fewer features are called as non-uniform.
 - If you select False, then the protocol does not use the enhanced spot finding method. The Use Enhanced SpotFinding setting is disabled when the Use Enhanced Gridding setting (in the Place Grid menu) is set to False. See **"To select to use enhanced gridding"** on page 126.

Note that enabling the Use Enhanced SpotFinding setting can slightly increase the array processing time.

To change values to find and size the spots.

To access the settings to find spots when spot format is selected

- 1 Click the cell to the right of **Spot Format**.
- 2 Click the down arrow, and select from the list of spot formats.

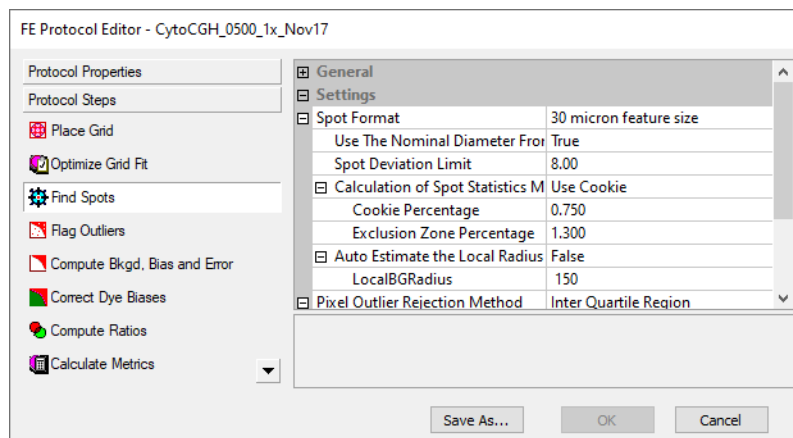


Figure 59 Find Spots - 30 micron Spot Format selected

In this protocol step you can also change the values for 3 additional sets of parameters associated with Feature Extraction processes to find spots once you have selected a spot format:

- Change values to find and size spots.
- Select a spot statistics method to define features.
- Select to autoestimate the local radius.

To change values to find and size the spots.

- 1 Click the cell next to the names of the following parameters.
- 2 Either select **True** or **False**, or enter a value, whichever is applicable.

Use the Nominal Diameter from the Grid Template

If you select True, the nominal diameter of each spot is obtained from the grid template. If you select False, the nominal diameter that is the starting spot size assumption for the Spot Finding algorithm is obtained from the grid placement algorithm.

Spot Deviation Limit

Enter a value that defines the distance the actual centroid may move relative to its nominal grid position (stated as a multiple of nominal radius). This parameter appears in the FEPARAMS table as SpotAnalysis_kmean_cen_reject.

To change values to find and size the spots.

To select a spot statistics method to define features

- 1 Select either **Use Cookie** or **Use Whole Spot**.

Cookie Method

The Cookie method locates an ellipse (usually a circle) over the centroid of each spot with a fixed radius entered in the protocol and defines the following regions represented in the figure below:

The Cookie method is usually used for Agilent microarray spots because of their uniformity.

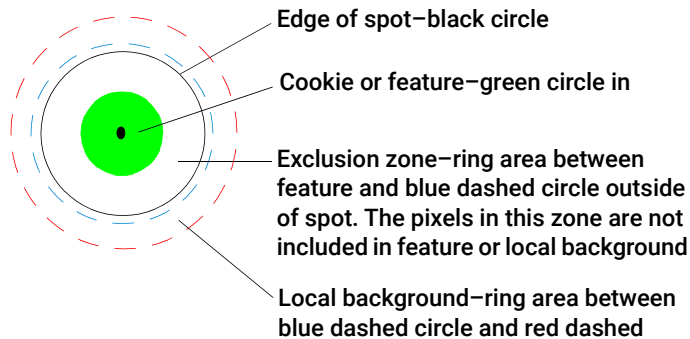


Figure 60 Feature and background definitions – Cookie method

Whole Spot Method

This method uses a detection technique based on a noise model to define the feature and exclusion zone. This method performs the spot size calculation first and defines the feature as the whole spot from the x and y diameters.

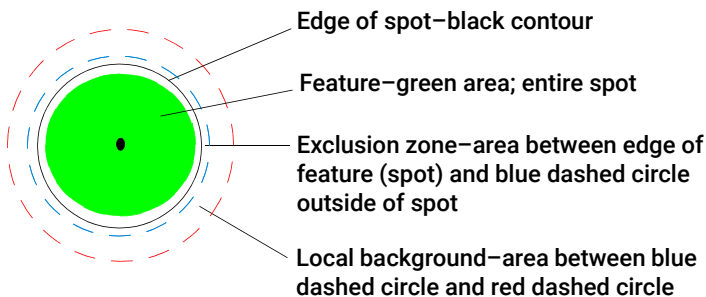


Figure 61 Feature and background definition – Whole Spot method

- 2 If you selected **Use Cookie**, change the **Cookie Percentage** and/or the **Exclusion Zone Percentage**, if necessary.

If you selected **Use Whole Spot**, change the **Exclusion Zone Percentage**, if necessary.

Cookie Percentage

Percentage of the nominal radius that is the cookie radius.

Exclusion Zone Percentage

Percentage of the nominal radius that is the exclusion zone radius.

To select to autoestimate the local radius to define local background

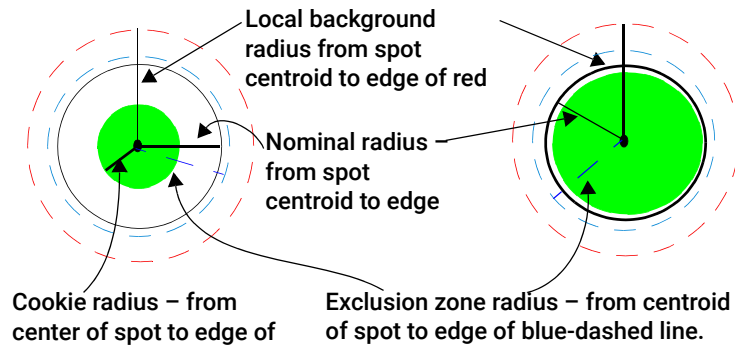


Figure 62 Definitions of different radii

The radii depicted in **Figure 62**, “Definitions of different radii,” on page 134 are calculated differently for the two methods.

See the *Feature Extraction for CytoGenomics Reference Guide* for more information.

To select to autoestimate the local radius to define local background

This radius method estimates background that is local to a feature of interest. See the *Feature Extraction for CytoGenomics Reference Guide* for details on this calculation.

- 1 Click the cell next to **Autoestimate the local radius**.
- 2 Select **True** or **False**.
- 3 If you select False, enter a larger radius if you want to include more background pixels or a smaller one if you want to include fewer.

For non-Agilent microarrays when there is a big difference in spot sizes, change this setting to False and make sure that the number you enter is big enough to accommodate the largest spot on the microarray.

If you select True, the software automatically estimates the radius to use to estimate the local background.

This approach eliminates any dependence of the local background computation on distance between spot-centroids along both axes. This ensures elimination of cross-contamination between feature and background signals, especially in closely packed microarrays.

The default radius produces a local background calculation around only one spot (Self-radius). If you want to include more background, you can increase the radius to include the background around the spot’s nearest neighbors (NN radius).

Final calculations for protocol step to find and measure spots

Final statistics

The final statistics calculated for all colors of each feature and local background, based on the remaining inlier pixels, are the average, standard deviation, median, IQR, and number of pixels. In addition, the pixel correlation between the red and green channels is calculated.

Pixel saturation

Saturation is defined as the maximum allowed intensity value for a pixel, its value depends on the make and model of the scanner.

Pixel saturation is also tested during calculation of the feature statistics. For the Agilent Microarray Scanner system, saturation is defined as an intensity value of 65535 minus the dark offset from the scanner. If more than 50% of the pixels remaining after outlier rejection are saturated, Feature Extraction sets a saturation flag.

With XDR extractions, the saturation limit could be as much as 20x higher than 65535. The array will saturate before it gets to 20x of 65535. The calculated saturation value based upon the XDR fit could be as high as 1,310,000 counts, but no biological signals are that high.

20-bit scans also have a high saturation value, one close to 1,048,576 counts.

Propagation of Find Spot errors to other algorithms

The measurement errors from the calculation of these statistics are propagated to the next steps in Feature Extraction according to standard error propagation techniques. For example, the errors for the signal average for the feature and local background regions are propagated to the background subtraction algorithm.

Changing Protocol Step 4: Flag Outliers

Not every feature is created, hybridized and scanned perfectly. Some backgrounds contain artifacts. Therefore, you can flag those features or local backgrounds that may represent misleading or erroneous data.

Although the grid tool of the Feature Extraction software lets you manually examine scanned microarray images and flag anomalous features, this task becomes exceedingly tedious and subjective for thousands of features on a microarray. This algorithm automatically finds all the anomalous features on the microarray.

To access the settings to flag outliers

- Click the Flag Outliers tab to change settings for flagging abnormal features and backgrounds.

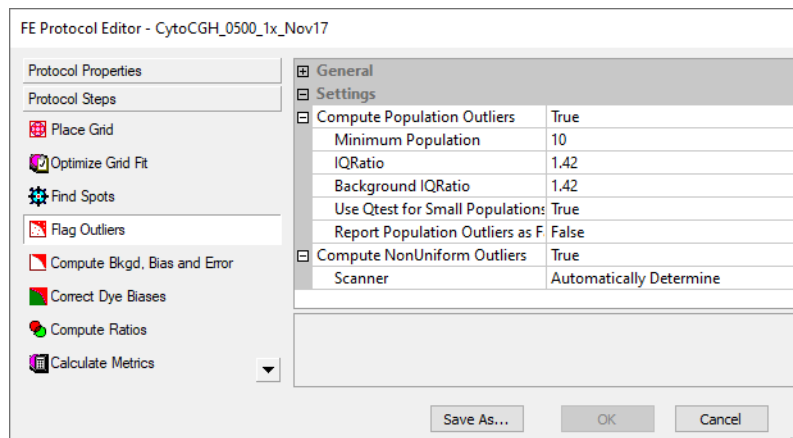


Figure 63 FE Protocol Editor - Flag Outliers selected

You can change two types of settings for this algorithm:

- Change settings to flag population outliers.
With this option you can flag features and local background regions that have intensities beyond the confidence limit for the distribution of the intensities for the replicate features and local background regions on the microarray.
- Change settings to flag non-uniform features and background region (outliers).

To change settings to flag population outliers

A type of outlier, called a “population outlier,” can be detected if there are enough replicates to perform population statistics on the signal intensities of the features or the signal intensities of the background regions.

For example, there are a number of control features on a typical Agilent microarray (typically, over 100 features) and thousands of local background regions. Out of this population, outliers can be rejected using box plot analysis.

1 Click the cell next to **Compute Population Outliers**, and click the down arrow.

2 Select **True** or **False**.

3 If you select True, change any of the default parameters that you need to.

The Flag Outliers tab contains values for the minimum population size to be considered for population outlier flagging, as well as a cutoff range. If a feature or local background region signal lies outside of this range, Feature Extraction flags that feature or local background region as a population outlier.

Minimum population This is the minimum number of replicates needed to perform the population outlier analysis. Although 10 is the default number of replicates, you can enter as low as 5 replicates.

To take advantage of the replicated probes in statistical calculations, Agilent uses the default threshold of 10 because its catalog microarrays have 100 probes that are replicated 10 times on the microarray. The default value for non-Agilent microarrays is 15.

IQRatio This formula determines the cutoff for box plot analysis. The $1.42 \times \text{IQ Range}$ (Interquartile Range) encompasses 99% of normal distribution. The Interquartile Range is the 75th percentile - 25th percentile. This calculation is the same as that for rejecting pixels that are of too low or too high intensity. In this case, however, the outlier features are flagged, not removed.

Background IQRatio This option uses the IQRatio formula for estimating the background region outliers.

Use Qtest for Small Populations? When the number of replicates falls below the minimum number specified in the protocol, a Q-test can be used to estimate the number of population outliers instead of the IQRatio. See Chapter 5 of the *Feature Extraction for Cyto Reference Guide* to learn more about Q-test calculations.

Report Population Outliers as Failed in MAGEML file When this option is set to *True*, any feature that is flagged as a population outlier is reported as *Failed* in the MAGE-ML output file.

To change settings to flag non-uniform outliers

The Agilent Microarray system includes Agilent microarrays, label/hyb/wash protocols and reagents, scanner and Feature Extraction algorithms.

To learn more about how the algorithm calculates the terms in the polynomial equation below, see the Feature Extraction for Cyto Reference Guide.

The non-uniformity outlier algorithm flags anomalous features or local backgrounds based upon statistically significant deviations from a noise model that makes use of the known noise characteristics of the Agilent Microarray system.

You do not typically need to do this step, as the software automatically determines the scan type and sets the parameters appropriately. To view or change the parameters that are automatically set, select a specific scan type. However, if you change the parameters, the protocol is no longer general and can only be used with that scan type.

- 1 Click the cell next to **Compute NonUniform Outliers**, and click the down arrow.
- 2 Select **True** or **False**.

“True” is the default setting for all Agilent-created protocols. If True is set, move on to the next step.

- 3 Click the cell next to **Scanner**, and click the down arrow.
- 4 Select either **Automatically Determine** or **Agilent Scanner**.

The settings change depending on your selection.

- 5 Click the cell next to **Automatically Compute OL Polynomial Terms**.
- 6 Select **True** or **False**.

True is default setting. A different set of parameters appears for a True or a False setting.

- 7 Enter the values for the parameters.

Parameters for a “True” setting

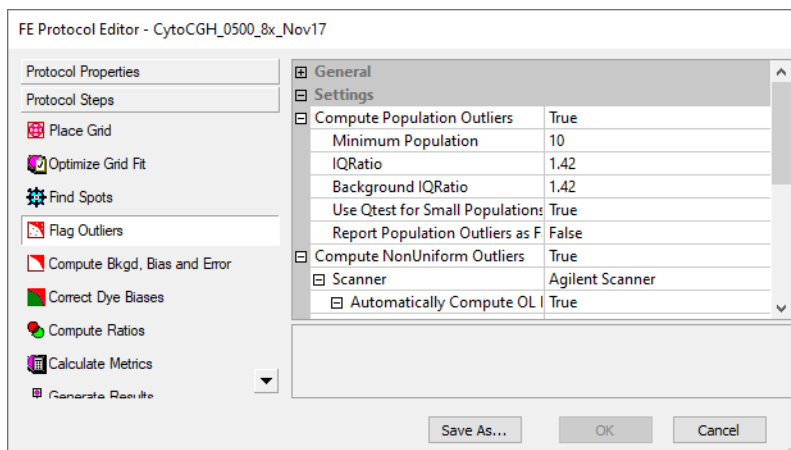


Figure 64 Flag Outliers – Agilent Scanner selected, Automatically Compute OL Polynomial Terms: True

The polynomial equation for the noise model is still used, but rather than entering values for the A, B and C terms as in the “False” case, you enter the values for the A (%CV)² term and multipliers for the B and C terms in both color channels (2-color) or one channel (1-color).

(%CV)² Term (A)

This term is the square of the CV (the constant coefficient of variation), or the intensity-dependent variance.

To change settings to flag non-uniform outliers

Red Poissonian Noise Term (B) Multiplier

The B term for the red channel is equal to this multiplier times the net signal at the 25% percentile of a histogram plot of number of negative control features or background vs. net signal.

Red Signal Constant Term (C) Multiplier

The C term for the red channel is equal to this multiplier times the variance at the 25% percentile of a histogram plot of number of negative control features or background vs. variance.

Green Poissonian Noise Term (B) Multiplier

The B term for the green channel is equal to this multiplier times the net signal at the 25% percentile of a histogram plot of number of negative control features or background vs. net signal.

Green Signal Constant Term (C) Multiplier

The C term for the green channel is equal to this multiplier times the variance at the 25% percentile of a histogram plot of number of negative control features or background vs. variance.

CAUTION

The default values for these terms have been optimized for Agilent microarrays, with an Agilent scanner and recommended lab protocols. You must perform optimization experiments to determine the optimum values for your microarrays and lab protocols.

Parameters for a “False” setting

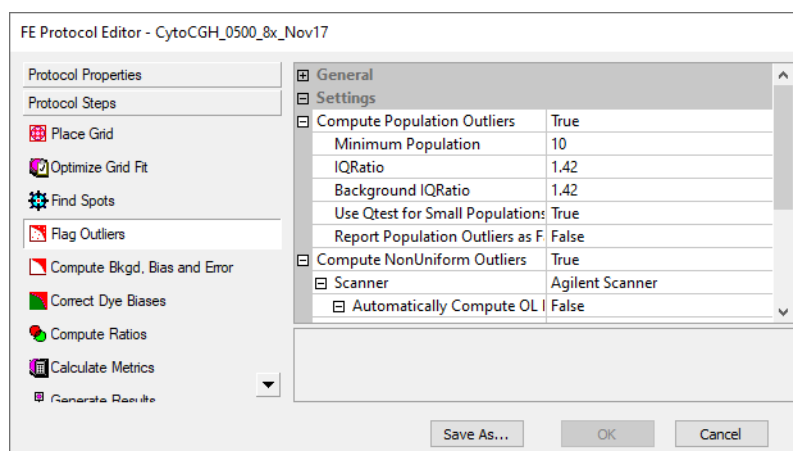


Figure 65 Flag Outliers – Agilent Scanner selected, Automatically Compute OL Polynomial Terms: False

A noise model, based upon a polynomial equation, is made up of three variables whose values are shown in the Automatically Compute OL Polynomial Terms folder:

(%CV)² Term (A)

Same as for the “True” case above.

Poissonian Noise Term (B)

This term specifies the variance due to Poisson distributed noise, which accounts for error in counting statistics.

Background Term (C)

This term specifies the variance due to signal-independent background noise (e.g., scanner and glass substrate noise).

Changing Protocol Step 5: Compute Bkgd, Bias and Error

At this stage of the Feature Extraction process, several systematic sources may still remain that contribute to the background signal component of the feature signal:

- Scanner offset, which can vary from 20 counts to 300 counts, depending upon the version of Agilent scanner [B (20) or A (300) scanner]. The value for each microarray offset is shown in the Stats table in the fields: gDarkOffsetAverage and rDarkOffsetAverage (See the *Feature Extraction for CytoGenomics Reference Guide*.)
- Glass and any contaminants or non-specifically-bound fluorescent signal on the glass. Artifacts from labeling, hybridization and washing steps can contribute to this source.
- Fluorescent signal that is non-specifically associated with the DNA probes themselves.
- A systematic gradient across a microarray.

Background region levels can also vary depending on the dye label and color channel used.

To access the settings to compute background, bias and error

- Click the Compute Bkgd, Bias and Error tab to select a method and change settings for background subtraction.

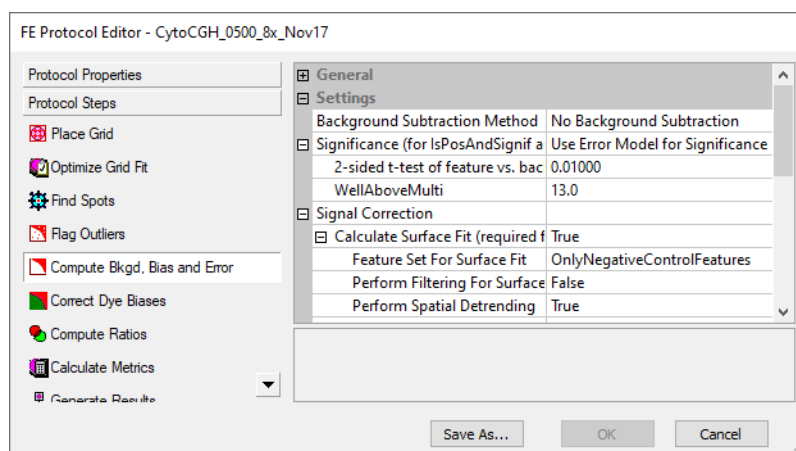


Figure 66 FE Protocol Editor - Compute Bkgd, Bias and Error

The selections for this protocol step let you decrease the contribution of the background signal to the feature signal.

- Select the background subtraction method.
- Choose to decrease the signal contribution due to an additive gradient at low signal across the microarray – spatial detrend.
- Choose to adjust the signal contribution due to a systematic multiplicative gradient across the microarray – multiplicative detrend. SNP probes are not included in multiplicative detrending.
- Choose to globally adjust the initial background-subtracted signals.
- Choose an error method to estimate the error in the calculations for the processed signal.

To select the background subtraction method

- Change settings to assess if the background-subtracted signal is significant compared to the background.
- Choose to use or not use surrogates to improve the reliability of low-signal error calculations.

To view the values used for background subtraction depending on the subtraction and correction methods selected, see the *Feature Extraction for CytoGenomics Reference Guide*.

To select the background subtraction method

There are several methods for subtracting the background signal from the raw signal, both local and global, available in the Correct Bkgd and Signal Biases sheet.

- 1 Click the cell next to **Background Subtraction Method**, and click the down arrow.
- 2 Select the background subtraction method most relevant to your microarray and processing conditions.

The default method for Agilent microarrays is *No background subtraction*. The default method for non-Agilent microarrays loaded with a grid file is Minimum signal (feature) on the microarray. This option makes the fewest assumptions about the microarray system. If you are analyzing non-Agilent microarrays or procedures, you have to determine the optimum default method.

No Background Subtraction

Even though no background is subtracted, you can make corrections to the signal using spatial detrend and the global background adjustment tools. The spatial detrend algorithm will correct for sources of background that are available in this protocol step, in addition to correcting for spatial gradients.

Local background method

This method subtracts the local background estimated for each spot using the radius method in the Find Spots protocol step. For example, local background for spot 300 is subtracted from the feature raw signal of spot 300.

The local background method may be a better background subtraction method than a global background subtraction method if your microarray has large regions of local backgrounds with much higher signals than the rest of the microarray.

These regions are often due to hybridization or wash processing artifacts. The local background method subtracts these local variations. This yields a better estimate of true feature signal if the local background signal is correlated with the feature signal, as determined by replicate statistics.

Global methods The Feature Extraction software supports a number of global background subtraction methods in case local background subtraction is not appropriate. With a global method, one number is calculated as background for the entire array. That number is subtracted from every feature. The number is reported in the STATS table for later queries. Global methods require that both non-uniformity outlier and population outlier methods be turned on. See the topic **“Changing Protocol Step 4: Flag Outliers”** on page 136 for more information.

Average of All Background Areas

This global method subtracts the average of the local background inliers estimated for all the spots using the radius method in the Find Spots protocol step. Inlier background regions have passed both the non-uniformity and population algorithms.

For example, the average background for all spots is subtracted from the feature raw signal of spot 300 and every other spot on that microarray.

To choose to remove the surface trend found in the data

Average of Negative Control Features

If the local background is not additive and your microarray has validated *negative controls*, Agilent recommends using the Average of negative control features option. If the local background is not additive and your microarray does not have negative controls, Agilent recommends the Minimum signal (feature) option.

This global method subtracts the average of the inliers of all the negative control features from each feature on the microarray. Inlier negative control features have passed both the non-uniformity and population algorithms. For example, the average of the inlier negative control features is subtracted from spot 300 and every other spot on that microarray.

Negative controls are probes that consistently show very little target binding across many types of targets and have no specific target bound to them by design.

If there is more than one probe sequence on the microarray annotated in the grid template or grid file as a negative control, this method analyzes the population of feature signals of each sequence to determine if there are significant differences among the different sequences at a fixed significance level of 0.01. The method then chooses the probe sequences that have the lowest mean signals and that are not significantly different from each other as the population of features to calculate negative control statistics.

Agilent provides validated negative controls for catalog microarrays. These probes have been shown to reflect the non-specific signal level that is present on probes.

Minimum signal (feature or background)

This global method subtracts the minimum signal on a microarray from each spot, whether the signal is a feature or local background region. The feature or background region signal chosen for the minimum signal must pass a low pixel standard deviation requirement. For example, the minimum feature or background signal on the microarray is subtracted from spot 300 and every other spot on that microarray.

Minimum signal (feature) on the array

This global method subtracts the minimum signal of a microarray feature that has passed a low pixel standard deviation requirement. This is the default setting for microarrays loaded with a grid file, but not necessarily the recommended setting. For example, the minimum feature signal on the microarray is subtracted from spot 300 and every other spot on that microarray.

To choose to remove the surface trend found in the data

What is detrending?

To learn more about the detrending calculations, see the Feature Extraction for CytoGenomics Reference Guide.

Removing the surface trend (detrending) decreases the contribution to the background of any systematic signal gradient on the microarray. For both 2-color and 1-color data, Feature Extraction removes the additive surface trend from the data, a process called *spatial detrending* (background detrending). Feature Extraction also removes any “domed” surface trend, a process called *multiplicative detrending*. SNP probes are not included in multiplicative detrending calculations.

This option is the default choice with protocols for Agilent microarrays (except the miRNA protocol). The detrending is done individually for each color channel and corrected separately.

Background detrending consists of two steps: calculating a surface fit and subtracting the surface fit from the data. This process is done independently for the red and green channels after background subtraction and before the global background adjustment.

To choose to remove the surface trend found in the data

Background Spatial Detrending

For spatial detrending, Agilent separates the surface fit and subtraction calculations that must exist together to remove the surface trend found in the data.

The surface fit calculation estimates the “foreground” of the data by measuring the intensity of the feature after background subtraction (or intensity of the MeanSignal of the feature if the “No background subtraction” option is selected) and performs a surface fit to the selected data. Even if the surface fit is not subtracted from the data (perform spatial detrending), the calculation appears in the QC Report to help evaluate your data.

The advantage of this method is that it can remove variation in foreground intensity for different regions of the microarray. By performing spatial detrending and subtracting the surface fit through this foreground, the data becomes more consistent and reproducible.

The *foreground* is the portion of the signal in a feature that is not related to the intended signal from the dye-labeled target complementary to the probes on that feature.

Feature Set for Surface Fit

AllFeatureTypes Select if microarray has a set of very low intensity features evenly distributed on the slide.

OnlyNegativeControlFeatures Select this if you have enough negative controls that are evenly spread around the microarray and ideally have different sequences represented – default for CGH protocols.

FeaturesInNegativeControlRange Select for low intensity features that are within 3 standard deviations of the spatial interpolated values of the negative controls.

Perform Filtering for Surface Fit

If True, allows a subset of the features into the set to be used to calculate the surface fits. You may want to do this to reduce the time needed to calculate the surface without reducing the accuracy of the fit.

If the Feature Set is AllFeatureTypes, then the filter chooses 1 to 2 percent of the dimmest features to be selected across the array. Select True for AllFeatureTypes. Select False for OnlyNegativeControlFeatures.

If the feature Set is FeaturesInNegativeControlRange, the filter chooses 1 to 2 percent of the features at random to go into the surface fitting set. Random x% rather than dimmest x% is used because the features in this set are already known to be dim.

Perform Spatial Detrending

If true, subtracts any additive trend found from the surface fit from the data.

To choose to adjust the background globally (2-color)

To correct the errors at the low end of the signal range, Agilent lets you choose to globally adjust the background after initial background subtraction. The global adjust method was developed to be used by those customers who decide to turn off all background and detrending methods.

This adjustment helps estimate and remove background signal due to fluorescent signal that is non-specifically associated with the DNA probes and that may be under-estimated by upstream background subtraction methods. You can also select this correction if you also chose *No background subtraction*.

- 1 Click the cell next to **Adjust Background Globally**, and click the down arrow.
- 2 Select **True** or **False**.
- 3 If you select True, enter a "pad" value between 0 and 500 to force the log ratio towards zero when both red and green signals are small.

The constant supplied to the *Adjust background globally* algorithm "pads" all the feature signals by that value. This will have the effect of compressing log ratios, but will decrease the variability (standard deviation) in the log ratio among inter- and intra-array replicates. You can try small pad values, typically in the few tens, and monitor the trade-off between log ratio compression and log ratio variability. The log ratio versus signal plot is a useful preliminary tool to assess the effect of the pad value.

By setting the *Adjust background globally* algorithm to the same constant with multiple microarrays, the residual background can now be defined by you and is therefore more consistent across multiple microarrays and multiple experiments than if a variable residual background signal is left for each feature, if no background is subtracted.

To choose to perform multiplicative detrending

Multiplicative detrending is an algorithm designed to compensate for slight linear variations in intensity if the processing is not homogeneous across the slide, resulting in different chemical reaction concentrations, for example, between the sides and the center, a "dome" effect. The multiplicative gradient is removed by dividing the signal by the observed gradient. SNP probes are not included in multiplicative detrending calculations.

- Make the selections described below.

Perform Multiplicative Detrending

Choose whether to perform multiplicative detrending (can be chosen for one- or 2-color arrays).

Detrend on Replicates Only

Choose whether to detrend on all of the data or to only use those probes that have replicates. If you choose to use replicates only, the data are normalized to the respective replicate averages before going into the fitting set.

Filter Low signal probes from Fit?

You can choose to filter out dim signals from the fit of the multiplicative trend surface. This is useful because the dimmer signals have a less obvious trend when that trend is multiplicative. You also choose how dim the signals must be in order to be removed from the set. This is stated as a multiple of the Negative Control Spread.

Perform filtering for fit

You must decide whether to use all of the data points or use local averages or medians.

To choose to use robust negative control statistics

This algorithm was developed to be used with microarrays that have many different probes classified as negative controls, such as CGH. The population outlier analysis, described on the previous pages, is performed on these negative controls, one probe sequence at a time.

- Select a **Robust Neg Ctrl Stats?** option.

True Performs a second round of population outlier analysis on the entire population of negative control features

Once outliers are identified, they are excluded from the calculation of the robust negative control statistics (average and standard deviation).

False Does not perform a second round of population outlier analysis

See the *Feature Extraction for Cyto Reference Guide* for more information on how this algorithm works.

To select the method to estimate the error (Error Model)

1 Click the cell next to Choose universal error, or the most conservative (propagated or universal), and click the down arrow.

2 Select an error method from one of these options:

Universal Error The Universal Error Model represents the expected error of the difference between the red channel and the green channel. This error is represented as the square root of the sum of two squared components from each channel: a constant noise term (additive error) for each channel, and an intensity scaled term (multiplicative error).

Using this error model, a p -value is calculated that corresponds to the probability that the difference between the red and green channels is zero. The Universal Error Model is a good estimator of error in high intensity features. It may underestimate the error of low signal features or features that have low signal-to-noise (e.g., non-uniform features).

Most Conservative This is the recommended method. This is a hybrid method which evaluates both the propagated error model and the Universal Error Model, and reports the higher p -value, and therefore the more conservative estimate of error, of the two models. This method captures the advantages that the propagated error model has for low intensity features and features with low signal to noise and the advantages the Universal Error Model has for high intensity features.

Propagated Error This method measures the error on the log ratio by propagating the pixel-level error of that feature through all the calculations on the raw signal, including raw pixel statistics, background subtraction, dye normalization and the log ratio calculation. Agilent follows standard propagation of error techniques and calculates a p -value that corresponds to the probability that the log ratio = 0.

The propagated error is a good representation of noise that predominates at low signal ranges, such as instrumentation noise and the noise due to non-uniform features.

The propagated error cannot capture other sources of noise that dominate at higher signal ranges, such as biological and chemical variability (e.g., writing and labeling, hybridization and wash processes). The propagated error method usually underestimates the noise, especially at higher signal levels.

To select the method to estimate the error (Error Model)

See “[Parameters for Universal and Most Conservative error models](#)” on page 146 for more instructions on changing the parameters for these two error models.

Parameters for Universal and Most Conservative error models

Parameters to change With this error model and the most conservative error model, you can enter a change to the following settings and parameters:

MultErrorGreen(Red)

Enter the multiplicative error component in the channel selected

Autoestimate Additive Error Green(Red)

You can select to enter your own constant for the additive error or turn on auto-estimation of the additive portion of the error model. If you select True, then Feature Extraction will auto-estimate the additive error. If you select False, then you must enter the constant.

The constant setting of the additive error approximates the error of feature signals that are near background. If the overall signal of the microarrays is increased or decreased (e.g. due to “hotter” or “cooler” labeling or variable scanner PMT settings), the additive error needs to be adjusted. The auto-estimation method of determining the additive error adjusts the scale accordingly.

CAUTION

The default additive and multiplicative parameters have been optimized using the most current Agilent microarray-based system (e.g., microarrays, protocols, scanner and most recent Feature Extraction algorithms). If you use non-Agilent microarrays, Agilent microarrays with non-Agilent protocols, or Agilent microarrays with non-Agilent scanners, you **MUST** re-evaluate these values to reflect the noise in your system.

Auto-estimation of additive noise Provided that the microarray, platform and protocols remain the same, the additive error can vary from one microarray to the next. Therefore, it is useful to correctly estimate the additive noise and set the additive constant equal to this noise.

To do this the software looks at microarray statistics: the width of the negative control distribution used in the background surface fit and the spatial variability of the spatial detrend surface (RMS of the difference between each point on the surface and the mean of the surface). Because the algorithm uses these surface statistics in the calculation, this option is only usable for microarrays for which surface calculation works.

You do not have to turn on “Spatial detrending” for the additive error to be automatically estimated. You do have to turn on the surface fit calculation for the additive error to be estimated.

Reliability of auto-estimation of additive error Agilent has found that the auto-estimation of the additive error results in error values that are more reliable than those generated if a constant is used. If the microarray quality or processing is less than optimal, the additive error is more reliable with auto-estimation. This will make the significance calls more correct and less susceptible to errors induced by microarray and processing variability.

To choose to use or not use surrogates

Agilent Feature Extraction software uses surrogate signals to calculate a more accurate and reproducible log ratio for signals of low intensity.

You choose to use surrogates in this protocol step so that the signals to be used for log ratios are subjected to dye normalization to remove this kind of bias.

Agilent Feature Extraction software uses the error estimate option that you select for all features, no matter if surrogates are used for the log ratio calculations for some features.

1 Click the cell next to **Use Surrogates**, and click the down arrow.

2 Select **True** or **False**.

Conditions under which surrogates are used

If you select True, the software uses surrogates under two conditions:

- If either red or green MeanSignal fails the IsPosAndSignif test (i.e., IsPosAndSignif = 0)
- If BGSubSignal is less than the Error

A log ratio cannot be calculated from a negative ratio. For any measurement system there is increased confidence in using a signal only if it is greater than a lower limit that is particular to that system. This ensures that the system lower limit is enforced as a lower limit on the signal.

Feature Extraction calculates the surrogate signal from the Error. For all default Agilent protocols that use Error Model significance, this is simply the Additive Error calculated during background subtraction.

Propagation of errors from the Compute Bkgd, Bias and Error algorithm to other algorithms

The measurement errors from the background subtraction are propagated to the next steps in Feature Extraction according to standard error propagation techniques. For example, the errors for the feature and background signals are propagated with the error on the background-subtracted signal to the DyeNorm algorithm.

Estimation of error is performed in this step. See **“To select the method to estimate the error (Error Model)”** on page 145 to see the choices show for this step. In this case, you cannot select the propagated error calculation by itself. If you select Most Conservative, Feature Extraction uses the larger of the two estimates from the Universal Error model or the propagated error.

To change settings to test the significance of a feature

Determining if a signal is significant, or just in the noise of the background, is an important step. The following two tests are performed to establish the significance of a feature: a two-sided t-test and a method to flag features that are “well above” background.

2-sided t-test of feature vs. background maximum p-value

The t-statistic is calculated based on the feature and background signals and their corresponding errors.

The background value used for this test depends upon the background method used. The values are reported in the Feature Extraction text output column “g(r)BGUsed”. If you have chosen not to subtract the background and Spatial detrend = Off, the local background for each feature

To change settings to test the significance of a feature

{g(r)BGMeanSignal} is used for the t-test but is not subtracted from the feature signal. If Spatial detrend = On when the background subtraction is off, the surface value for each feature is used for the t-test and is subtracted from the feature signal.

To view the values used for the t-test depending on the subtraction and correction methods selected, see the Feature Extraction for Cyto Reference Guide.

Estimating background error for significance testing uses one of two methods. The method used is dependent upon the protocol parameter, "Significance (for IsPosAndSignif and IsWellAboveBG)".

Pixel significance

For this method, the SD of the background that is used for the t-test is the same value used for surrogates (See "[Use of surrogates for log ratio and p-value calculations](#)" on page 154 of this guide). This is a population level SD for global background methods and is a pixel-level SD for both the local background and No background methods. These SD's are reported in the Feature Extraction text output columns "g(r)BGSDUsed". (Not used for error model significance, which is the favored method.)

The number of degrees of freedom is also based on the error calculations and the number of pixels, which depends on the background subtraction method. The p -value is then calculated based on a two-sided t-test. If this p -value is less than the maximum p -value defined in the protocol, the feature is considered to be significantly different from the background.

The features that "fail the null hypothesis", that is, have a signal that is significantly different from the background and also that have background-subtracted signals > 0 will have a boolean set = 1 in the Feature Extraction text output column "g(r)IsPosAndSignif". The features that are not significantly different from background or that have background-subtracted signals < 0 will have the boolean set = 0 for the same Feature Extraction text output column.

Note that entering a positive pad value for global background adjust may increase the number of features considered positive and significant in the t-test, in some cases, all features.

WellAboveMulti

Features that have background-subtracted signals that are greater than "WellAboveMulti*BackgroundSignal Standard Deviation" are flagged as well above background. A default value of 2.6 SD for the multiplier is consistent with a 99% confidence interval. The text output for the result of this test appears as g(r)IsWellAboveBG in the Features table of the output text file.

Error model significance

The other method to determine feature significance is based on the error model. The protocol parameter "Significance (for IsPosAndSignif and IsWellAboveBG)" is set to "Use Error Model for Significance" for this method. The additive error is used as the estimate of background noise and is used for surrogates.

The significance test for positive and significant then becomes the probability of the feature actually having zero signal for a feature with a given background-subtracted signal with a known additive error (or noise of features with no true signal) assumed to be gaussian distributed about zero.

You choose a p -value to decide how far along the Gaussian distribution the feature needs to be before it should be considered Significant. When this choice is made, the Well Above Background is decided by a multiple of the additive error.

Changing Protocol Step 6: Correct Dye Biases

Systematic differences in dye bias must be removed from the CGH data. To remove dye bias, the Feature Extraction software uses a dye normalization algorithm.

A *dye bias* is the difference in the red and green signals caused by different efficiencies in labeling, emission or detection.

To access the settings to correct dye biases

- Click the Correct Dye Biases button to change settings to remove difference in dye bias.

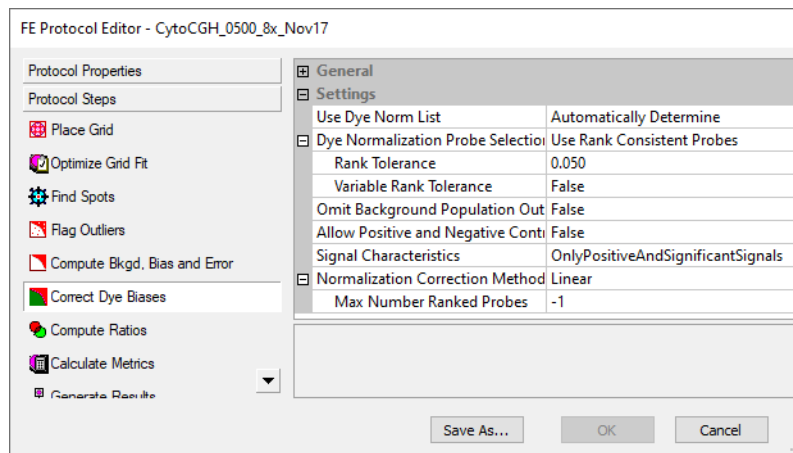


Figure 67 FE Protocol Editor - Correct for Dye Biases selected

You can change the following types of settings for this algorithm.

- Select how dye normalization gene list is determined
- Select features that are included in the normalization set
- Select to omit background population outliers
- Select to allow positive and negative controls
- Select the signal characteristics
- Select a curve to determine the bias from that normalization set of features

The software normalizes the entire microarray from this measured bias.

If you use a self versus self experiment, in which the feature signals should be the same in both channels across the dynamic range of signal intensities, differences in dye behavior can result in the signals in one channel increasing across the signal range at a different rate than in the other channel.

This behavior can be visualized by plotting the background-subtracted signals of one channel versus the other and noting that the slope of the line fit through the data points is not equal to one. Dye normalization corrects for this by multiplying the signals by appropriate dye normalization factors.

To choose how to select dye normalization list

Three options are available for choice of the dye normalization gene list for correction of dye biases.

1 Click the cell next to **Use Dye Normalization Gene List**, and click the down arrow.

2 Select one of the following choices.

Automatically Determine

The DyeNormList is used, if present in the grid template. Otherwise, the list is generated internally, based on the Dye Normalization Probe Selection Method. If no DyeNormList is selected in the grid template, extraction will not fail.

True When this is selected, the DyeNormList from the grid template is used. Extraction fails if there is no DyeNormList selected in the grid template.

False When selected, a DyeNormList will not be used even if it is specified in the grid template. During the dye normalization step, an internal dye normalization list is generated, based on the Dye Normalization Probe Selection Method.

To choose a method to select a set of dye normalization probes

Feature Extraction supplies two methods for selecting representative features that constitute the set used for bias measurement.

1 Click the cell next to **Dye Normalization Probe Selection Method**, and click the down arrow.

2 Select one of the following methods.

Use Rank Consistent Probes

This is the recommended method of selecting normalization features. The rank consistency filter selects features that fall within the central tendency of the data by observing consistent trends between the red and green channels. This method assumes that the central tendency of the data is unregulated. This method requires that the features pass the three criteria described below:

- They have not been flagged as non-uniform in either channel.
- They are not saturated.

The condition for saturation is more stringent in this algorithm (> 5% of pixels saturated) than in the algorithm for features as saturated (> 50% of pixels saturated).

- They are not population outliers in either channel.

Rank Tolerance

You can enter a *Rank Tolerance* value if you select the Use Rank Consistent Probes choice. The *rank tolerance* is the red rank value minus the green rank value divided by the number of features, considered for rank consistency, times 100, and expressed as a percentage. If this value is 5, for example, then a rank consistent feature would be one whose red rank did not vary above or below the green rank by more than 5%.

To assign a red rank and a green rank, order all of the probes based on the intensity of their background-subtracted signal. A rank 1 is the lowest intensity signal feature. The highest rank is assigned to the most intense signal. A perfectly "rank consistent" feature would be one exactly at the 75% percentile of the distribution in both the red and green signal.

Variable Rank Tolerance

You can choose to allow variable rank tolerance. This does rank tolerance after grouping the data into signal groups. In each group the same designated percentage of most rank-consistent data are included.

To select to omit background population outliers

Use All Probes This method assumes that the average log ratio over all inlier non-control features should be zero. All non-control features that pass the four criteria listed for the Rank Consistent Probe method are candidates for this method set.

To select to omit background population outliers

- 1 Click the cell next to **Omit Background Population Outliers**, and click the down arrow.
- 2 Select **True** or **False**.

Selecting True excludes any features from the dye normalization set if the local backgrounds associated with those features have been flagged as population outliers (in either channel). The default recommendation is False.

To select to allow positive and negative controls

If you turn this function on, Feature Extraction uses positive and negative controls in the dye normalization calculation.

- 1 Click the cell next to **Allow Positive and Negative Controls?**, and click the down arrow.
- 2 Select **True** or **False**.

The default recommendation is False.

To select signal characteristics

You can select to include features with the following signal characteristics as part of the set of rank consistent features.

- 1 Click the cell next to Signal Characteristics, and click the down arrow.
- 2 Select one of the following options.

OnlyPositive andSignificant Signals Use only positive and significant signals for ranking in the rank consistency filter.

AllPositive Signals Use all positive signals for ranking in the rank consistency filter.

AllNegative AndPositive Signals Use all negative and positive signals for ranking in the rank consistency filter.

To select a normalization correction method

After you select a set of normalization features, you select a normalization method to apply to that set. The results from the calculation are applied to all features on the microarray.

- 1 Click the cell next to **Normalization Correction Method**, and click the down arrow.
- 2 Select from one of the following methods:

To select a normalization correction method

Max Number of Ranked Probes The maximum number of ranked probes allows only a certain number of probes into the dye normalization set. If the number of rank-consistent probes exceed this number, a random set is chosen to meet the maximum number. If the number entered is -1, this filter is not used. This option can speed up the data processing without adversely affecting accuracy for data sets with large numbers of probes, especially for 65-micron spot size and higher density arrays.

Linear This method multiplies the background-subtracted signal by a global constant. The global constant is determined separately for the red channel and the green channel using the selected normalization features. It is calculated using the geometric mean of these features as a value of 1000.

Linear and LOWESS This is the default method for normalizing Agilent microarrays. Dye biases tend to be signal intensity-dependent. A dye normalization method that uses a constant scaling factor cannot correct for dye biases across the entire range of signal intensities. This method does a linear normalization across the entire range of data, then applies a non-linear normalization (see LOWESS only description below) to the linearized data set.

LOWESS only This method assumes the dye bias is intensity-dependent, and therefore the background-subtracted signal is multiplied by a value that is a function of the feature intensity. The selected normalization features are used to calculate a normalization curve that measures the potential log ratio bias across the entire range of feature intensities.

A well-known curve fitting method called LOWESS (a locally weighted linear regression curve fit) is used to fit this curve. Given any feature and its intensity, this curve determines the potential dye bias that should be present in the log ratio measurement of the feature. Features are then corrected by multiplying the red and green channels by an adjustment that will remove the bias from the log ratio.

Changing Protocol Step 7: Compute Ratios

To determine if a feature is differentially expressed, three measures are important: *log ratio*, *log ratio error* and the *p-value*.

Log ratio is the log (base 10) of the ratio of the red channel signal to the green channel signal. *Log ratio error* is the error about the mean of the log ratio. *P-value* is a measure of confidence that a feature is, or is not, differentially expressed

The log ratio error is determined by the error model and therefore used to help assess the significance of the computed log ratio.

The *p-value* is calculated using the log ratio and the log ratio error. Genes with lower *p-values* have higher confidence of being differentially expressed. For example, if the *p-value* is ≤ 0.01 , there is a 99% confidence level that the gene is differentially expressed.

To access the settings to compute ratios

- Click the Compute Ratios tab in the FE Protocol Editor to change the settings to calculate these measures.

See **Figure 68**.

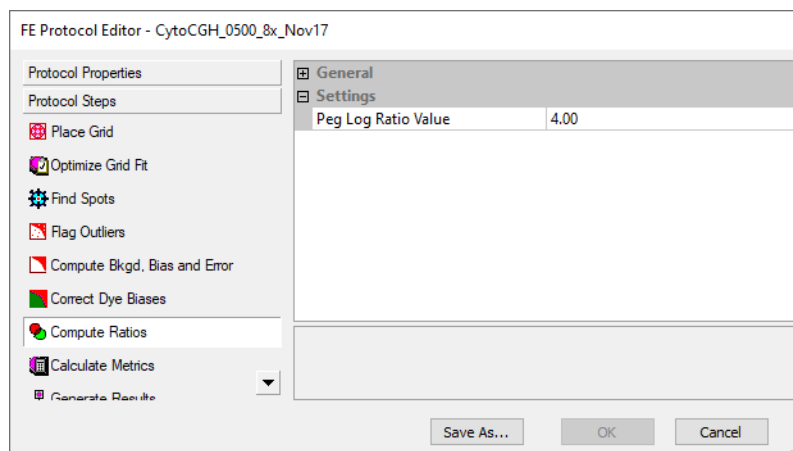


Figure 68 FE Protocol Editor - Compute Ratios selection

You can change only one parameter for this protocol step:

- Enter the value for the log ratio limit.

To enter the value for the log ratio limit

When you set a limit, the features whose log ratio exceeds that limit, either in a positive or negative direction, will not be used in any calculations. The recommended limit for this value is 4.

- 1 Click the down arrow in cell next to **Peg Log Ratio Value**.
- 2 Enter a positive or negative value up to 4.

Use of surrogates for log ratio and p-value calculations

Surrogates are used in this step to calculate certain log ratios, p-values and log ratio errors.

The table below shows the conditions for use of surrogates to calculate log ratios, log ratio errors and p-values.

Table 8 Use of surrogates for log ratio and p-value calculations

If red channel signal is	And green channel signal is	Then ratio is	And Feature Extraction does this*
Positive and significant vs. background and greater than red background SD (R)	Positive and significant vs. background and greater than green background SD (G)	R/G	Uses the dye-normalized signals for both channels for the calculations.
Not positive and significant vs. background or less than red background SD (r)	Positive and significant vs. background and greater than green background SD (G)	r/G	Uses the dye-normalized green signal and the dye-normalized red background SD signal (surrogate) for the calculations.
Not positive and significant vs. background or less than red background SD (r)	Not positive and significant vs. background or less than green background SD (g)	r/g	Uses surrogates for both channels. Sets the log ratio = 0. Sets the p-value = 1.
Positive and significant vs. background and greater than red background SD (R)	Not positive and significant vs. background or less than green background SD (g)	R/g	Uses the dye-normalized red signal and the dye-normalized green background SD signal (surrogate) for the calculations.

*For log ratios equal to or between -4 (ratio = .0001) and +4 (ratio = 10000). If (r/G) > 1 or (R/g) < 1, log ratio = 0 and p-value = 1.

Table 9 Summary – Use of surrogates for calculations

<p>Case 1: R/G Both channels use DyeNorm Signals. P-value and log ratio are calculated as usual. For signals not using surrogates, $g(r)DyeNormSignal = g(r)ProcessedSignal$, which is then used to calculate log ratio.</p>	<p>Case 2: r/G $r = rSurrogateUsed$ $G = gDyeNormSignal$ P-value and log ratio are calculated as usual. If $r/G > 1$, then Feature Extraction automatically sets $LogRatio = 0$ and $PvalueLogRatio = 1$</p>
<p>Case 3: R/g $R = DyeNormSignal$ $g = gSurrogateUsed$ P-value and log ratio are calculated as usual. If $R/g < 1$, then Feature Extraction automatically sets $LogRatio = 0$ and $pValueLogRatio = 1$</p>	<p>Case 4: r/g Both channels use surrogates. Feature Extraction automatically sets $LogRatio = 0$ and $pValueLogRatio = 1$ For signals using surrogates, $g(r)ProcessedSignal =$ $g(r)SurrogateUsed \times g(r)DyeNormFactors$.</p>

Effect of using surrogates on log ratios

Below are plots that show the difference between using surrogates and not using surrogates for a set of self vs. self microarray data.

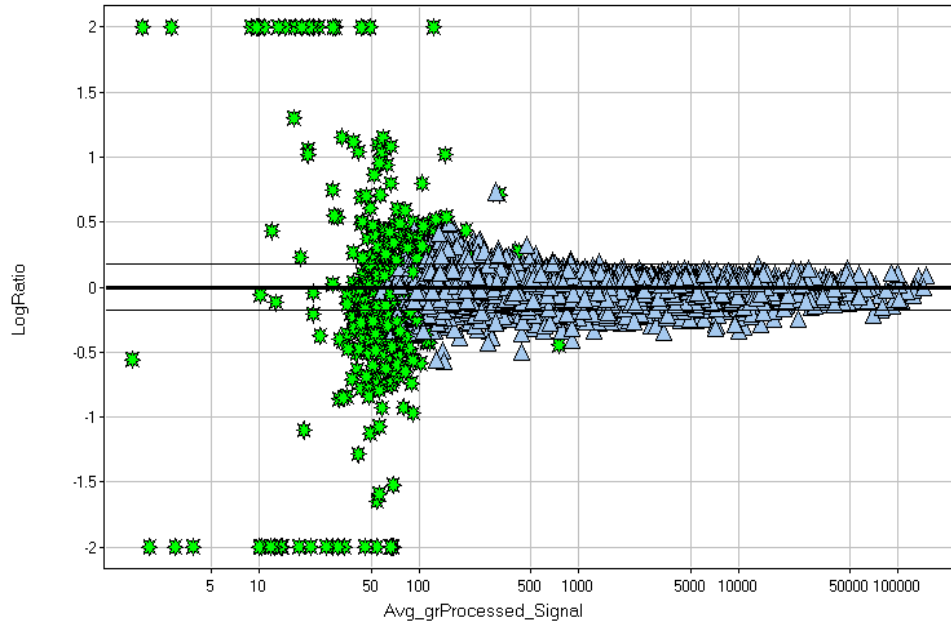


Figure 69 Plot of LogRatio vs. Avg_grProcessedSignal–No surrogates

Green stars are data points affected by use of surrogates in **Figure 70**.

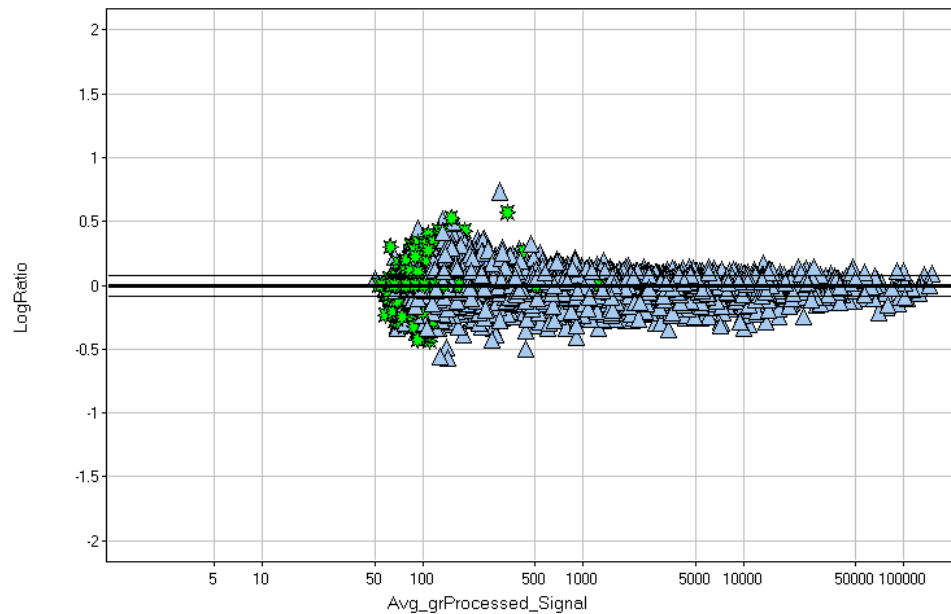


Figure 70 Plot of LogRatio vs. Avg_grProcessedSignal–With surrogates

Use of surrogates for log ratio and p-value calculations

Green stars are data points affected by use of surrogates. See [Figure 69](#) for their positions without surrogates.

Effect of using surrogates on log ratio errors

If "Use Surrogates" is True, the software calculates log ratio errors using the dye-normalized and/or surrogate signals, according to the error model selected for the calculation.

If "Use Surrogates" is False, the software calculates log ratio errors using the dye-normalized signals, according to the error model selected for the calculations. If either or both channels have negative dye-normalized signals, the p -value is set to 1 and the log ratio error is set to 1000. If red channel < 0 , then Feature Extraction sets the log ratio = -4. If green channel < 0 , then Feature Extraction sets the log ratio = +4. If both red and green channels < 0 , then Feature Extraction sets the log ratio = 0.

Changing Protocol Step & Calculate Metrics

In this step you can change the settings for the metrics that appear in the QC report.

To access the settings to calculate metrics

- Click the Calculate Metrics tab in the FE Protocol Editor.

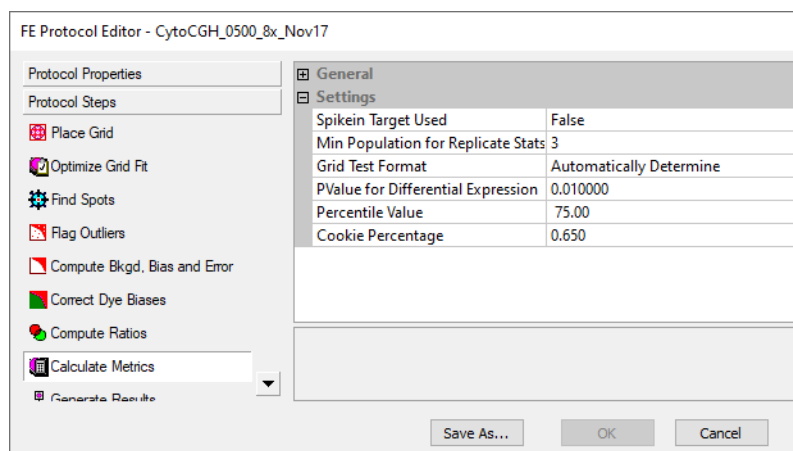


Figure 71 FE Protocol Editor - Calculate Metrics selected

You can change the following parameters in this step:

- Select **True** for **Spikein Target Used** if Agilent spike-in targets were used during lab processing of the microarray. The appropriate metrics are generated in the QC report.
- Enter the minimum population for replicate statistics.
- Select the grid test format.
- Enter the p -value for differential expression.
- Enter the percentile value.

To select to use spikein targets

- 1 Click the cell next to **Spikein Target Used?**, and click the down arrow.
- 2 Select **True** or **False**.

If you select True, Feature Extraction will do all the calculations necessary to report the statistics for both the non-control features and the spike-in controls.

To change the minimum population for replicate statistics

The QC report includes statistics based on replicate probes (%CV, log ratio uniformity). Feature Extraction must know the minimum population of replicates in order to calculate these statistics.

- 1 Click the cell next to **Minimum Population for Replicate Statistics**.
- 2 Enter a number of 3 or greater.

To select the grid test format

- 1 Click the cell next to **Grid Test Format**.
- 2 Select the grid format of the images to be tested, or select **Automatically Determine** to let the software decide for you.

To change the p -value for differential expression

The QC report contains a plot of the log ratio vs. the magnitude for each feature to show which features are up-regulated, which are down-regulated and which show no differential expression. This assessment is relative to the p -value chosen for this setting.

- 1 Click the cell next to **PValue for Differential Expression**.
- 2 Enter a p -value against which you intend to measure differential expression on the microarray.

To change the percentile value

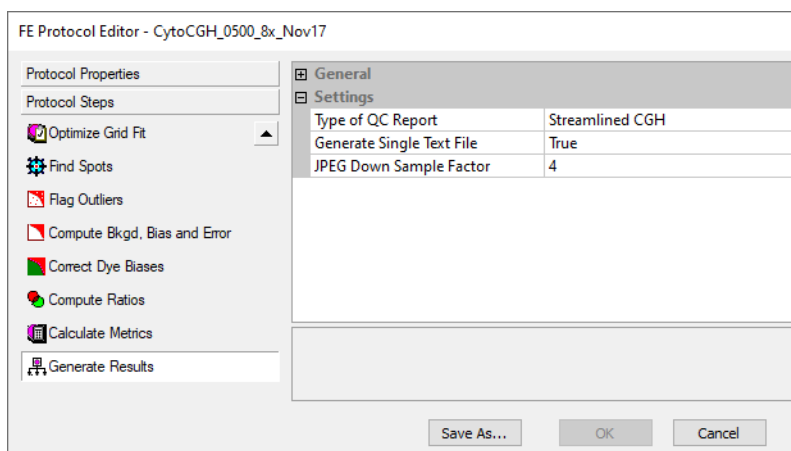
This value is an indicator of the distribution of signals of the non-control features and is useful in downstream normalization of 1-color data.

- 1 Click the cell next to **Percentile Value**.
- 2 Enter a value to set the distribution of signals of the non-control features.

Changing Protocol Step 9: Generate Results

To access the settings to generate results

- Click the Generate Results tab in FE Protocol Editor to change the values for output.



FE Protocol Editor - Generate Results selected

You can change the following settings for this protocol step:

- Select the QC Report type.
- Select to generate a single file for the text output.
- Change the compression factor for the JPEG output.

To select the type of QC Report

- 1 Click the cell next to **Type of QC Report**, and click the down arrow.
- 2 From the list, select the QC report type to be generated.

Gene Expression	Generate a Gene Expression QC report.
Streamlined CGH	Generate a Streamlined CGH QC report. For more information, see the <i>Feature Extraction for CytoGenomics Reference Guide</i> .
CGH_ChIP	Generate a standard CGH_ChIP QC report.
miRNA	Generate an miRNA QC report.
Test Chip	Generate a Test Chip report. This report is used by Agilent service personnel for scanner performance verification.

To select to generate a single file for the text output

To select to generate a single file for the text output

- 1 Click the cell next to **Generate Single Text File?**, and click the down arrow.
- 2 Select **True** or **False**.

If you select true, the text output appears in a single file. If you select False, it appears in three separate files, one for FEPARAMS, one for STATS and one for FEATURES (results).

To change the compression factor for the JPEG output

The value for this factor determines the degree to which the image is scaled down in both directions before being converted to JPEG.

- 1 Click the cell next to **JPEG Down Sample Factor**, and click the down arrow.
- 2 Enter a number.

5 Changing Image Displays

- Displaying and saving image files **162**
- Loading Feature Extraction visual results **167**
- Changing image display settings **171**
- Changing Feature Extraction window displays **185**

This chapter presents instructions on changing image displays and working with image files to view the images before and after extraction.

Displaying and saving image files

To open an image file

- 1 Click on the toolbar Open image file button, or click **File > Open > Image**, to display the Open Dialog Box.
- 2 In the Open Dialog Box, double-click on a .tif file from the scanner, or select a file and click **Open**.

An image of the scanned microarray contained in the file appears in the Workspace.

You can also open an image file by dragging the file to the Feature Extraction desktop icon.

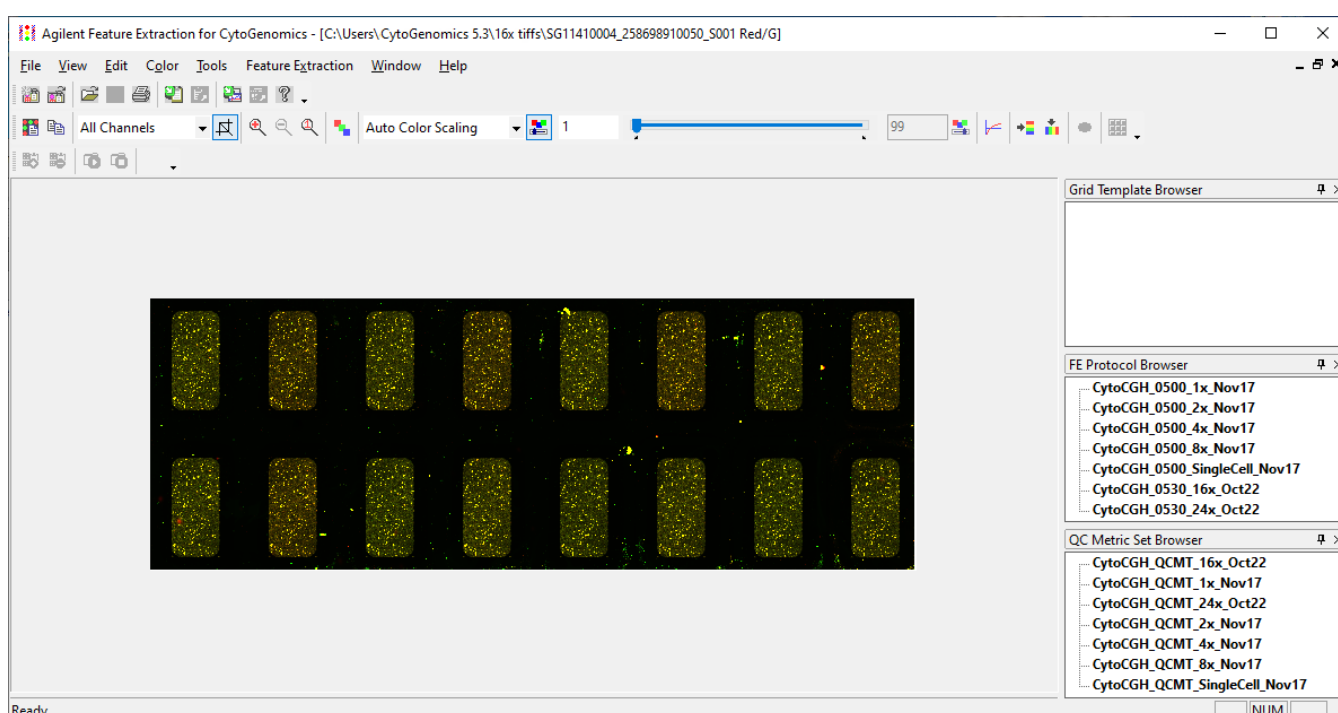


Figure 72 Uncropped TIFF image for Agilent microarray

You can also open an image file double-clicking the image file in a project extraction set. For information on how to do this, see [Chapter 2](#), “Extracting Microarrays Automatically”.

To crop an image

Crop an image using the cursor

You do not have to crop an Agilent image before you load it into a project and extract. You may want to crop images to view a subset of features or to use for illustration purposes.

When you first open an image, the Cropping Mode button is turned ON because this is the default setting. (See **“To change default image viewing settings”** on page 182.)

- 1 If not already on, click the Cropping Mode On/Off button.
- 2 Position the crop cursor at the top left corner of the image that you want to crop, and hold the left mouse button down.
- 3 Drag the crop cursor to the opposite corner.
- 4 Release the left mouse button to display the cropped image.



Figure 73 Image top menu and toolbar in Feature Extraction Tool

If you are viewing high-resolution images, you must crop them to one-tenth or less (i.e., at least 200%) of their original size in order to see the features at high resolution. Any larger cropped image is a preview image.

The color of the crop cursor helps you recognize the proper size for the high fidelity image. If the cursor is black, the cropped image is the high fidelity image. If the cursor is red, the cropped image is still a preview image.

You can cancel the cropping operation by using the (Esc) key, or click the Cropping Mode On/Off button to turn cropping off.

The Status Bar is located at bottom of the Feature Extraction window and contains the pixel information about the viewed image. If the viewed image is the preview image, the pixels are from the preview. If the viewed image is the high-resolution image, the pixels are the high-resolution pixels.

To save a cropped image to a new file

Crop an image using corner locations (manual crop)

This option lets you crop an image with greater precision than you can attain with the cursor option.

NOTE

For manual cropping to work, the image after cropping must be a minimum of 30 x 30.

- 1 Click **Edit > Manual Crop**.

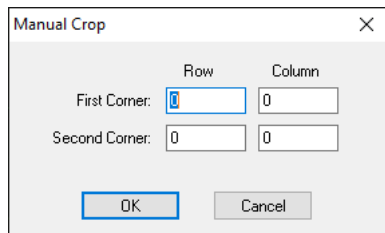


Figure 74 Manual Crop dialog box

- 2 Enter the locations of the corners (diagonal to one another) that you intend to have as the corners of the cropped image.
- 3 Click **OK**.

To save a cropped image to a new file

If the image file is from an Agilent microarray with a barcode:

- 1 Click **File > Save modified image** to save the cropped image to a new file.

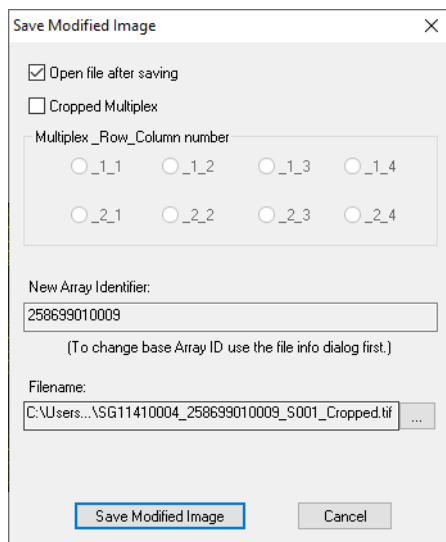


Figure 75 Save Modified Image dialog box for Agilent microarray

- 2 To close the file after saving, clear the Open file after saving check box.

To save a cropped image to a new file

- 3 To specify which pack in the multiplex (formerly known as multipack) microarray was cropped, select a Multiplex_Row_Column number.

This lets you save the cropped single pack with its multiplex coordinates.

If no multiplex was cropped, clear the Cropped Multiplex check box.

- 4 To change the New Array Identifier, see **“To add or change the New Array Identifier”** on page 166.

For Agilent microarrays with a barcode, the New Array Identifier is automatically generated from the barcode and displayed, along with a suggested file name.

The New Array Identifier can be changed. In fact, if for any reason the barcode does not generate a New Array Identifier, you must enter one if you intend to send the MAGE-ML results file to Resolver.

- 5 To change the path for the cropped image:

- Click the ellipsis (...) button.
- Select the new path.
- Click **OK**.

- 6 Click **Save Modified Image**.

If the image file is from a non-Agilent microarray:

- Click **File > Save modified image** to save the cropped image to a new file.

For microarrays without an identifier, the New Array Identifier field is blank. You can add an identifier. See **“To add or change the New Array Identifier”** on page 166.

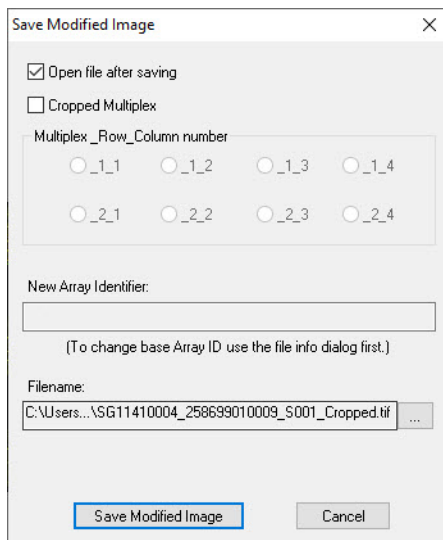


Figure 76 Save Modified Image dialog box for non-Agilent microarray

To save a cropped image to a new file

To add or change the New Array Identifier

When you attempt to save a modified image, the Save Modified Image dialog box appears with a field for New Array Identifier. Agilent barcodes automatically generate this number. For non-Agilent images, this field is blank.

If the New Array Identifier is not generated, for whatever reason, the software will not transfer MAGE-ML or GEML results files to a third-party program such as Resolver.

You can change or add this identifier in the Scan Image Properties dialog box.

- 1 Click **View > Image File Info** to bring up the Scan Image Properties dialog box.
- 2 Click the cell next to the Identifier field, and enter the new number.
- 3 Click **Apply**, and click **Close**.

A message appears that says you must click **File > Save As** before the change is accepted.

- 4 Click **OK**, and click **File > Save As**.
- 5 Enter a new name or select the old name of the .tif file, and click **Save**.

If you were trying to save a modified image, try again. The New Array Identifier you entered appears in the field.

Loading Feature Extraction visual results

You can select to store visual results in the same local folder as the other types of result files or to have it be sent via FTP to a remote directory. In either case the visual results file is saved with the same name as the results file name with an extension of .shp.

By default, the software shows the results for outlier features only when you load the visual result file or shape file (.shp) for the image. To view all results, you must clear the View Outliers Only check box. It is recommended, the you first crop out a region of the image before clearing the outlier check box. For instructions, see [“To change the magnification \(zooming\)”](#) on page 176 and [“To change the extraction results display”](#) on page 178.

- 1 Open the image for which you want to see the visual result.

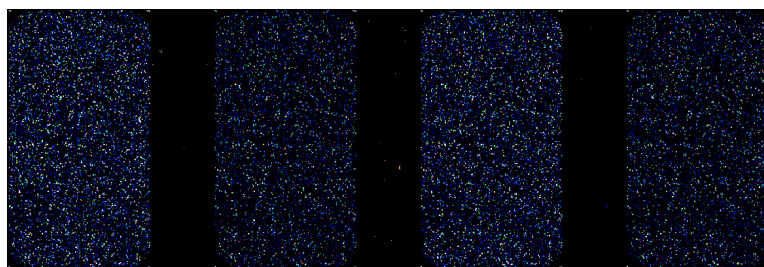


Figure 77 Image of Agilent 1-color, 4x44K microarray

- 2 Click **Feature Extraction > Load Visual Result** to load previously saved visual results for the currently displayed image file.

The software opens the folder containing the image file whose visual results you want to view. If you chose to save the output files for the project to a results folder separate from the image file folder, you must move to that folder.

- 3 Select the shape file, and click **Open**.

Or, you can also open the visual result file that is associated with the extraction set from the Project Explorer.

- 1 From Project Explorer, right-click the image.

- 2 Click **Load Visual Results**.

This opens the image and loads the shape file associated with that extraction set.

To view visual results for low-resolution images:

- 1 After .shp file is loaded to the image, zoom in until you see enough of the spots to help you evaluate the results.
- 2 Pass cursor to feature of interest and hold for a few seconds.
A description of the feature appears.

To view visual results for high-resolution images (2- by 3-micron or 20-bit):

To find a specific feature, use Ctrl-F. Enter the Feature number of the specific feature you want to find.



Figure 78 Visual results for low-resolution microarray image

To view visual results for high-resolution images (2- by 3-micron or 20-bit):

After you first load the image, you, in fact, are viewing a preview image. You can only view the original high-fidelity image if you crop the size of the image to one-tenth or less of the original size.

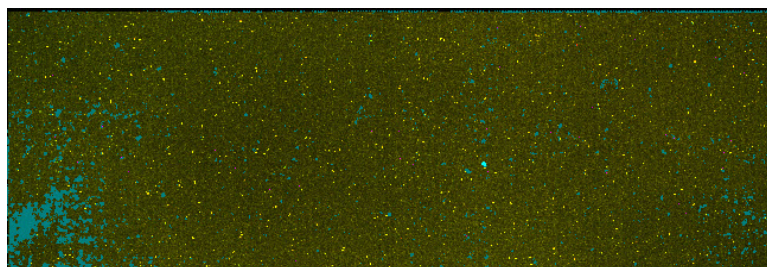


Figure 79 Preview display of high-resolution image with outliers in blue

- 1 After .shp file is loaded to the image, zoom in until you see enough of the spots to help you identify the region of interest.
- 2 Click the left mouse button at the point where you want to start forming a rectangle for cropping.
- 3 With the crop cursor, draw a rectangle equal to or less than one-tenth the size of the image.

Note that the crop cursor is black if the size is within one-tenth (200% to 1000%) of the original size, but it turns red if you are drawing outside the high fidelity range.

- 4 Then release the mouse button.

Note the status bar at the bottom of the window. It should now read "High Fidelity".

High fidelity mode is when you zoom in the image by at least 200%.

- 5 Pass cursor to feature of interest and hold for a few seconds.

A description of the feature appears.

To save an image (original or cropped) for import to another program

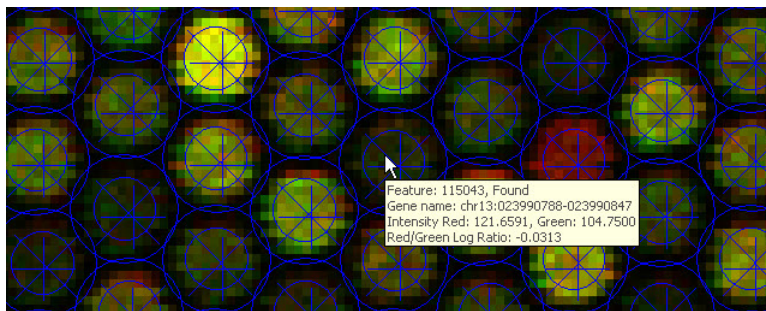


Figure 80 Visual results of high-resolution microarray image in High Fidelity mode, zoomed to 500%

View pixels used

This view option for low-resolution images and for high-resolution images that are in “High Fidelity” mode (i.e., when the image is zoomed in to one-tenth of its original size).

The View Pixels Used option displays the pixels used to quantify the feature signal within the cookie area. Pixels that are determined to be outliers are shown in black.

- 1 If the image is a high resolution image, complete steps 1-4 from **“To view visual results for high-resolution images (2- by 3-micron or 20-bit):”** on page 168.
- 2 Under **View > Extraction Results**, select **View Pixels Used**.

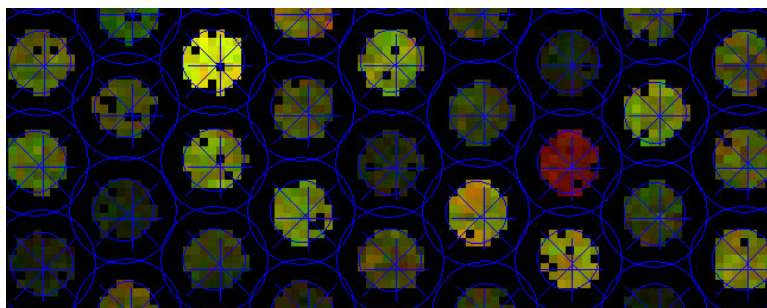


Figure 81 View pixels used for high-resolution microarray image in High Fidelity mode, zoomed to 500%

To save an image (original or cropped) for import to another program

- 1 Click **File > Save As** to produce high resolution images.
The Save As dialog box lets you save .tif files or .jpg files of the current image or cropped image.
- 2 Click the down arrow to the right of the **Save as type** list.
- 3 Select either Tif Files, BMP files Png Files, or Jpeg Files.
- 4 Enter the name of the image.
- 5 Click **Save**.

To display file information

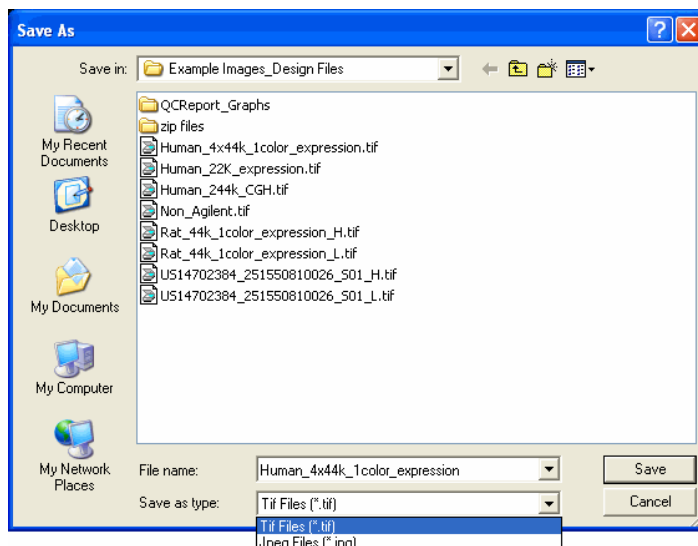


Figure 82 Save As dialog box

To display file information

You can use the File Info dialog box to view valuable information, such as scanner warning messages, scanner version, PMT voltage, dark offset information, the scan resolution and glass thickness

- Click **View > Image File Info** to display a tabbed dialog box that displays all available information about an image file.

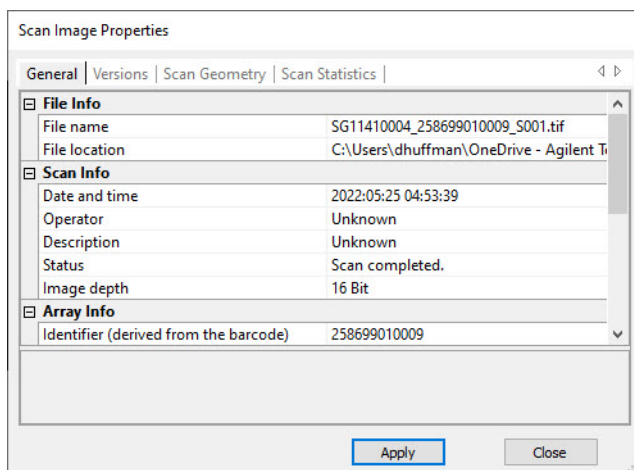


Figure 83 Scan Image Properties dialog box

The only field that you can change is the Identifier field. See **“To add or change the New Array Identifier”** on page 166.

Changing image display settings

Changing the image display does not affect the Feature Extraction results in any way.

When you open an image you can manipulate displayed images from the original scan file, the cropped scan file or the Feature Extraction results in the following ways:

- Change image orientation (landscape/portrait)
- Change the number of channels displayed
- Change the display color
- Change between log and linear color scale
- Change the intensity scale (color display range)
- Change the magnification (zooming)
- Change the fit of the window to the image
- Change the extraction results display

This section also shows you how to change settings for histograms and line graphs, the default image analysis parameters and printing.

To change image orientation (landscape/portrait)

When you first load a microarray image into the Feature Extraction Tool, the image appears in the orientation produced by the scan, either landscape or portrait. If the gene list that you want to use to position a grid and find spots has a different orientation than the image, you must change the image orientation to match that of the gene list.

Due to the size of 2um or 3um scan images, the flip and rotate operation can take several minutes.

- 1 Click **Tools > Flip Upper Left to Lower Right (Landscape/Portrait) and Save** from the top menu.

The Save As dialog box opens.

- 2 Click the down arrow to the right of the **Save as type** list.
- 3 Select either Tif Files, BMP files Png Files, or Jpeg Files.
- 4 Enter the name of the image.

Click **Save**.

If the original orientation was landscape, the software takes the upper left corner and moves it to the lower right corner to create a portrait orientation. The same is true if you go from a portrait to landscape orientation.

To change the channels that display

The default setting for displaying the channels is to display one window with the two color channels combined.

To change the channels for a single window:

- 1 On the image viewing toolbar click the down arrow for the Channel Display menu.

The default selection is All Channels.



- 2 Select **Red Channel** or **Green Channel**.

The window display changes to the channel selected.

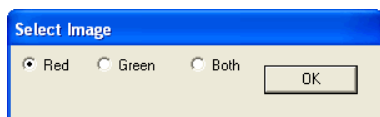
You can also change the number and color of the channels displayed when you set default image settings.

To display separate channel images:

Displaying the images of separate color channels side by side can be useful for comparing background anomalies on the microarray.

- 1 Click **Window > New "Filename" View Window**.

"Filename" represents the path and name of the image that you loaded. The Select Image dialog box opens.



- 2 Click **Red** or **Green**.

- 3 Click **OK**.

If you repeat steps 1-3, all three windows—the red, the green and the combined color—display at the same time.

To change the display color

“Color Mode” choices are False Color, Gray Scale, Reverse Gray Scale and False Color w/White BG.

- From the image viewing top menu, click **Color > “color mode”**.

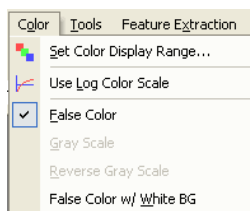


Figure 84 Color menu

The color type that you select depends on the channel image display. For combined channels, there are two choices of color, and for single channels there are three choices.

Combined channel display

The default setting for the color on the TIFF image is called False Color. You also have the option of selecting False Color with a white background.

- Click **Color > False Color w/ White BG**.

To return to False Color, click **Color > False Color**.

False Color This selection displays 1-color images using a color palette that progresses from black intensity (zero) to blue to cyan to green to yellow to red to magenta intensity (65535). For 2-color images, the red channel is displayed as shades of red and the green channel as shades of green that are combined to show shades of yellow.

False Color w/ White BG This selection maps intensities the same as the False Color palette, except that a pixel with zero values in both red and green is turned white instead of black. This feature is useful when printing because it saves ink. Generally, the minimum data range must be manually adjusted to produce the best possible image.

Single channel display

The default setting for the color on the single-channel display is also False Color.

- Click either **Color > Gray Scale** or **Color > Reverse Gray Scale**.

Gray Scale Presents the spot intensity in white and the background in black.

Reverse Gray Presents the spot intensity in black and the background in white.

To return to False Color, click **Color > False Color**.

Examples of images with different color selections

The following images are intentionally not aligned.

To change image orientation (landscape/portrait)

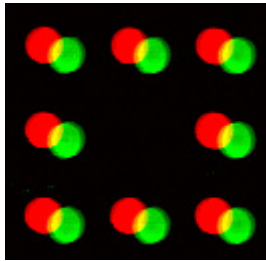


Figure 85 False Color for 2-color image

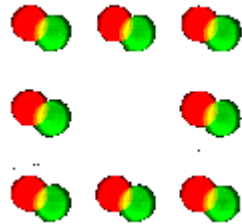


Figure 86 False Color with white BG for 2-color image

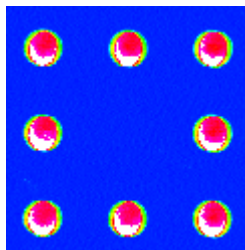


Figure 87 False Color for 1-color image

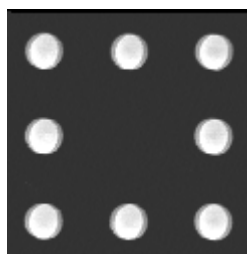


Figure 88 Gray Scale

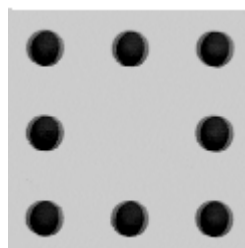


Figure 89 Reverse Gray Scale

To change the intensity scale (color display range – CDR)

To change between log and linear color scale

- Click **Color > Use Log Color Scale** to toggle the color scale from linear to log₁₀ (intensity).
To return to the linear scale, click **Color > Use Log Color Scale** again to remove the check mark.

With the log scale you can see low intensity portions of the microarray at the expense of detail at high intensities.

You can also use the Log/Linear button on the toolbar to toggle the color scale. 

To change the intensity scale (color display range – CDR)

Changing the color display range affects how the image intensity is scaled. Modifying the color display range can make it easier to visualize your microarray image but does not affect the raw intensity.

You can change the color display range in one of two ways:

- Using the color scaling toolbar.
- Using the Set Color Display Range dialog box.

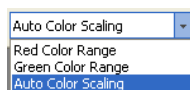
Use the color scaling toolbar

The easiest way to determine the best range is to use the Auto Scale function on the toolbar.



Figure 90 Color scaling segment of the image viewing toolbar

- Click the down arrow next to the color scaling menu, and click Auto Color Scaling if necessary.



- Click the CDR Low Mode icon, and move the lower bar cursor to the minimum percentage for the display range.
- Click the CDR High Mode icon, and move the higher bar cursor to the maximum percentage for the display range.


Notice the change in color of the image display as you increase or decrease the blue bar size.

- To fine tune the color display ranges manually, click either the Red Channel or the Green Channel from the color scaling menu.
- Click the CDR Low Mode icon, and move the lower bar cursor to the minimum pixel number for the display range.
- Click the CDR High Mode icon, and move the higher bar cursor to the maximum pixel number for the display range.

To change the magnification (zooming)

Use the Set Image Color Display Ranges dialog box

Alternatively, you can set the color display range through the Set Image Color Display Ranges dialog box.

- 1 Click **Color > Set Color Display Range**, or click **Set Color Range** on the toolbar. 

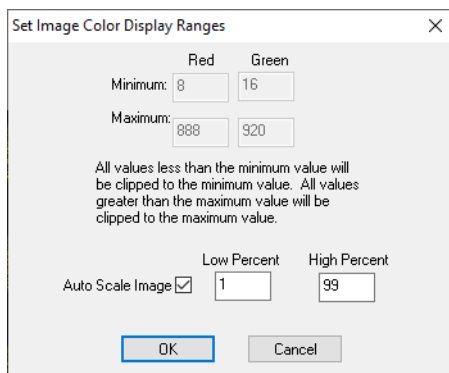


Figure 91 Set Image Color Display Ranges dialog box

- 2 Mark the Auto Scale Image check box to set a low and a high percentage cutoff value.

Setting 0 for a low value and 100 for a high value will compress the entire range of possible values into the 8-bit displayable range. If this produces a black image, try 1 and 99 (the default settings). Generally, this gives you an acceptable data range, even for very poor images.

- 3 To fine tune the color display ranges manually, clear the Auto Scale Image check box.
- 4 Change the Minimum and Maximum Red and Green values.

The data range determines how the 16-bit image data is mapped to the 8-bit data that is displayed on a computer screen. Generally, scanned images do not cover the entire range of possible values: 0 to 65535. Therefore, if you try to display the entire range, all you can see is black.

To change the magnification (zooming)

To set the default % zoom settings for the loaded image and the grid mode images, see "To change default image viewing settings" on page 182 of this guide.

Magnification can only be set to one of the % zoom levels shown in **Figure 92**. Numbers greater than 100% map one scan pixel to multiple pixels on the screen. For example, 600% maps one scan pixel to a 6 x 6 square of identical pixels. A zoom value of 100% maps one scan pixel to one pixel. Values less than 100% display the mean value for several scan pixels to one pixel. For example, 50% averages four scan pixels, a 2 x 2 matrix, and displays the average as one pixel.

You can control zooming in one of three ways:

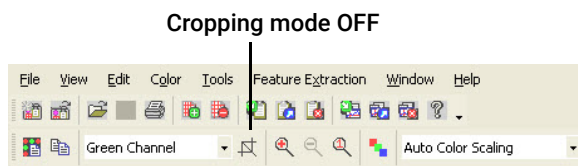
- Dragging the cursor to make a rectangle
- Using buttons on the tool bar and on the mouse
- Using the View menu

Drag the cursor to make a rectangle

You can only zoom the image using this method if cropping is turned off.

To change the magnification (zooming)

- 1 Clear the Cropping Mode On/Off button to turn cropping off.



- 2 Position the cursor at the top left corner of the image section that you want to zoom, and hold the left mouse button down.
- 3 Drag the cursor to the opposite corner of the section you want to zoom.
- 4 Release the left mouse button to display the selected area.

Use the Tool Bar and Mouse Buttons

Three buttons on the tool bar and the right and left mouse buttons control zooming. You can use these buttons with Crop Mode On or Off.

Use the Zoom Menu

You can use the Zoom menu with Crop Mode On or Off.

- Click **View > Zoom** to jump to any zoom level directly.

For low-resolution images, if the image is uncropped, the zoom levels range from 10% to 1000%.

For high-resolution images, if the viewed image is uncropped, the zoom levels range from 10% to 100%. To activate zoom levels from 200% to 1000% (High Fidelity mode), you must crop the image.

Note that the crop cursor is black if the size is within one-tenth (200% to 1000%) of the original image size, but it turns red if you are drawing outside the High Fidelity range.

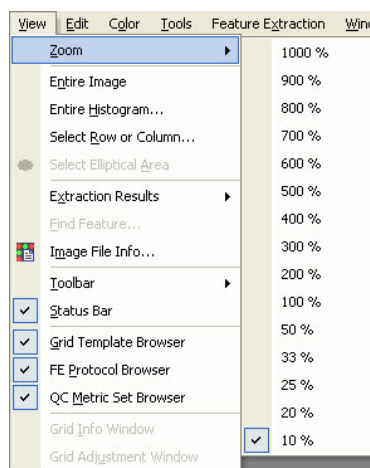


Figure 92 Zoom menu for low-resolution image

To change the window size to fit the image

To change the window size to fit the image

If you have to scroll to see the entire image, use this option to increase the size of the window to fit the image. If the window is too big for the image and contains white space around the perimeter of the image, use this option to decrease the size of the window to fit the image.

- Click **View > Entire Image** to resize the window.

To change the extraction results display

Use this option to control the portions of the feature extraction results that are displayed.

- Click **View > Extraction Results**.

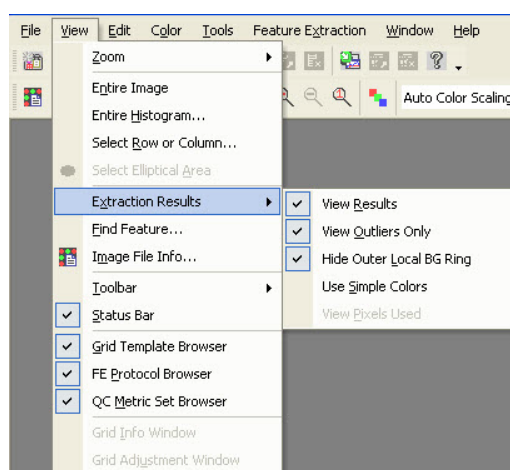


Figure 93 Extraction Results – Default views selected

View Results

Results shown depends on whether View Outlier Only is selected or not.

Shows all visual results when this option is selected and View Outliers Only is not selected.

Shows outlier results only when both this option and View Outlier Only are selected.

Hide all visual results when this option is not selected, even if View Outliers Only is selected.

View Outliers Only

Shows outlier features only when both this option and View Outliers Only are selected.

Hide Outer Local BG Ring

Shows or hides the local background ring. See **“To select a spot statistics method to define features”** on page 133 of this guide for a definition of the local background.

Use Simple Colors

Toggles the color display of outlier features and background.

View Pixels Used

Shows or hide pixels within the cookie that are used to quantify the feature mean or median intensity.

For more details, see **“View pixels used”** on page 169.

To change and save histograms and line graphs

For high-resolution images, histograms and line graphs use the original data, not the data from a low-resolution preview image.

Set up a histogram graph

Histograms can be very useful to initially visualize your results. They can show you if you have an anomalous distribution of intensities across the entire slide or across rectangular or elliptical subregions of the image.

Entire viewing area

- 1 Click **View > Entire Histogram** from the main menu bar.

If the selected file has more than one color, a Channel Selection dialog box appears.

- 2 Select the desired channel, or select both.

Rectangular subregion

- 1 Make sure that cropping is OFF.
- 2 Click the right mouse button to position your cursor in one corner of the region to capture.
- 3 Hold down the right button and drag the pointer to the opposite corner, then release.

If the selected file has more than one color, a Channel Selection dialog box appears.

- 4 Select the desired channel, or select both.

Elliptical subregion

- 1 Click **Tools > Preferences > Graph Viewing**.
- 2 Click the cell next to **Select Elliptical Area for Histogram**, and click the Up arrow.
- 3 Select **True**, and click **OK**.

False is the default setting and results in a rectangular subregion.

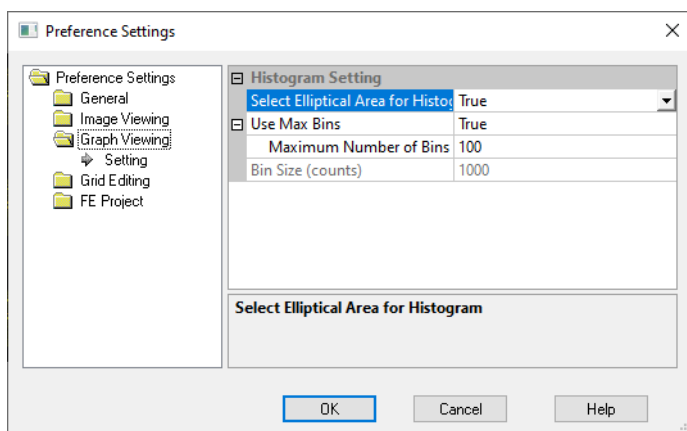



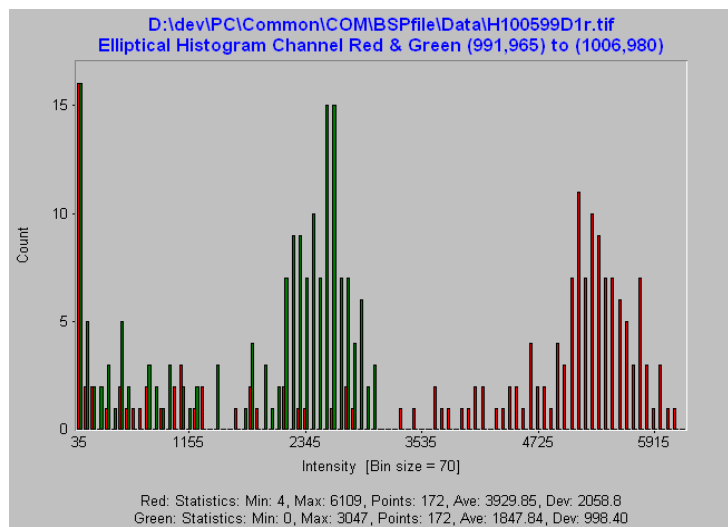
Figure 94 Histogram Preference Settings dialog box

- 4 Make sure that cropping is OFF.
- 5 Click .

To change and save histograms and line graphs

Or, click **View > Select Elliptical Area**.

- 6 Click the right mouse button to position your cursor in one corner of the region to capture.
- 7 Hold down the right button and drag the pointer to the opposite corner, then release.
If the selected file has more than one color, a Channel Selection dialog box appears.
- 8 Select the desired channel, or select both.



Change the bin size for histograms

- 1 To adjust the bin size after creating the initial graph, click **Histogram > Change Bins**.
- 2 Change the bin size or the maximum number of bins.
- 3 Click **OK**.

The maximum bin size is 65535; the minimum is 1.

The maximum number of bins is 32768; the minimum is 2.



Change the default bin size for histograms

- 1 Click **Tools > Preferences**.
- 2 Click the Graph Viewing folder.
- 3 Change the bin size or the maximum number of bins.
- 4 Click **OK**.

Set up a line graph

Line graphs show a vertical or horizontal slice through the image. They can be useful for spotting extreme differences between the red and green signal intensities across the microarray. Line graphs are also useful for displaying average intensity over all rows and columns, minimizing the effect of noise.

To create a line graph more quickly, double-click a selected row within the image with the left mouse button, or double-click a selected column with the right mouse button.

- 1 Click **View > Select Row or Column**, or click  or  .
The Select Row/Col dialog box opens.

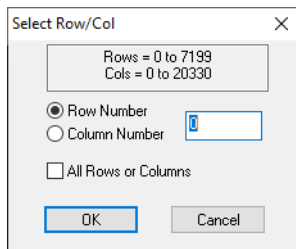


Figure 95 Select Row/Col dialog box

- 2 Click **Row Number** for showing signals in a row, or click **Column Number** for showing signals in a column.
- 3 Enter the number of the row or column.
- 4 Mark the **All Rows or Columns** check box to average the rows or columns in the view depending on the Row Number or Column Number selection.
- 5 Click **OK**.

The Select Image dialog box opens for selecting the channels to include in the graph.

- 6 Select **Red, Green,** or **Both**. Click **OK**.

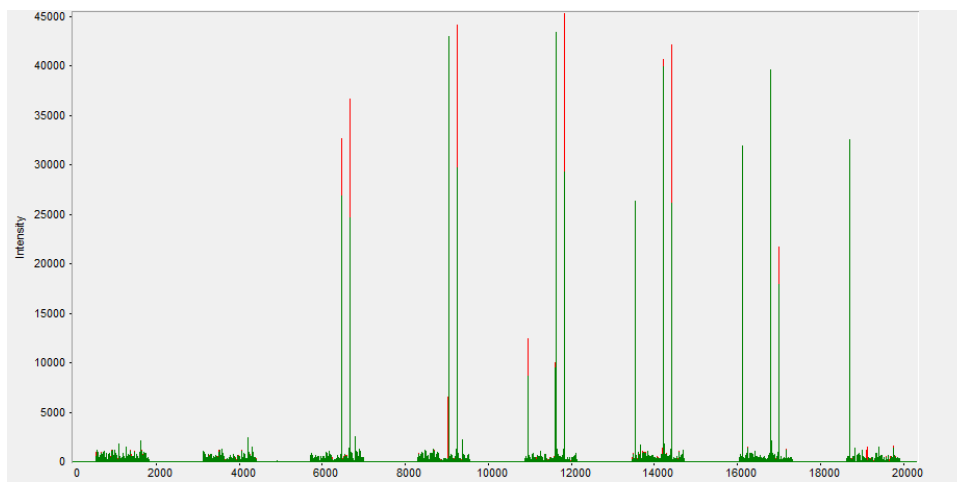


Figure 96 Line graph example

Use the Graph Toolbar

You save graphs produced from the image files to third party programs, such as word processors and spreadsheets, with the buttons in the Graph Toolbar. You can also print and modify the graph display with this toolbar.

- Click the selected button in the Graph Toolbar at the top of the histogram or line graph.

To change default image viewing settings

You can also change the default settings that appear when you open a new file. To activate the settings you must exit the image file and either reopen the same file or open a new one.

See *“To change the display color”* on page 173, *“To change between log and linear color scale”* on page 175, and *“To change the intensity scale (color display range – CDR)”* on page 175 in this chapter for details on the selections in the first two sections of the Image Viewing Preferences window.

- 1 Click **Tools > Preferences**.
- 2 Open the Image Viewing folder, if necessary.
- 3 Change the default selections if you need to.
- 4 Click **OK**.
- 5 Close the image.
- 6 Reopen a file.

The new default settings now take effect.

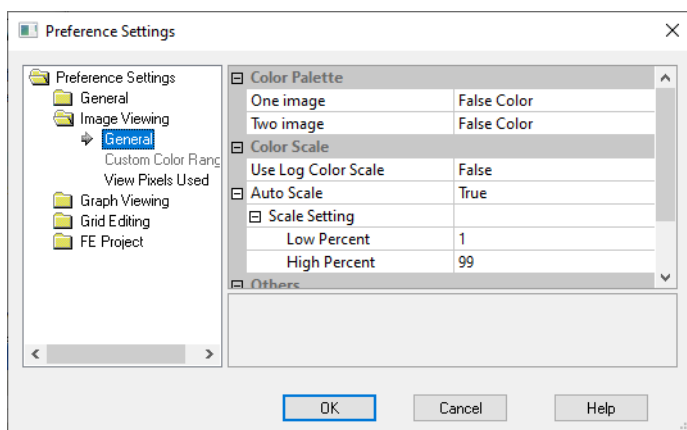


Figure 97 Image View Default Options dialog box

- Color Palette** You can set the color for the images found on a microarray.
- Use log color scale** If set to True, the software uses the log₁₀ intensity scale for the color. If False, it uses a linear scale.
- AutoScale** If set to True, you can change the Low Percent and the High Percent for autoscaling the color display range.
- Start with crop mode** Select True if you want to be able to crop the image with mouse movement right after you load the image. Select False if you want the mouse to activate the zoom function.
- Ratio of window size to feature cross for found feature** When you attempt to find a feature in the visual result image, you can change the ratio of the resulting window size for the image to the size of the feature cross for this found feature.

Change other default settings in the Preferences dialog box

See the tasks in the relevant chapters for changing the default settings for the Grid Mode (Chapter 2), for Graphs (histogram bin settings in this chapter) and for Projects (Chapter 1).

To change the print settings

The Feature Extraction application supports printing to both color and black and white printers.

Change page settings

- Click **File > Page Setup** to set the following printer paper parameters:
 - Paper size
 - Paper source
 - Paper orientation
 - Margins
 - Printer

Change print settings

- 1 Click **File > Print Setup** or **File > Print** to bring up the standard Windows Print Setup dialog box.

Within this dialog you can change the following parameters:

- Paper size
 - Paper source
 - Paper orientation
 - Margins
 - Printer
- 2 Click **OK** on the Print Setup dialog box to save changes to the printer setup without printing, or Click **Print** to save changes to the printer setup and to print.

Preview a file before printing

- Click **File > Print Preview**.

Printing single color images

If you are printing a single color image that is displayed using the false color palette to a black and white printer, the Image Analysis Tool of Feature Extraction automatically converts the image to gray scale. This is necessary because the gray scale intensities that a black and white printer automatically generates for a false color image are incorrect.

Examples The following images are intentionally not aligned.

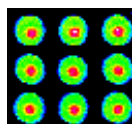


Figure 98 False Color image

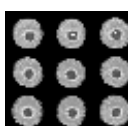


Figure 99 Incorrect image generated by black and white printer

Hot Key actions for image displays

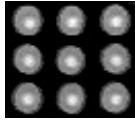


Figure 100 Correct gray scale generated by the Image Analysis Tool

Printing 2-color images

Printing 2-color images on a black and white printer does not require the Image Analysis Tool to convert the image to gray scale, but you cannot distinguish between the colors. See the following example.



Figure 101 2-color image

These images are intentionally not aligned.



Figure 102 2-color image printed on a black and white printer

Images always print at maximum resolution. The zoom level does not affect printing. This resolution setting makes printing the image slow but also generates the highest quality image possible.

Hot Key actions for image displays

Table 10 Hot key actions for image displays

If you want to do this:	Press this "hot key":
Turn Crop Mode On/Off	Alt-C
Call up the manual crop dialog box	Ctrl-M
Call up the File Info dialog box	Ctrl-I
Change image orientation (landscape/portrait)	Shift-U
Zoom in	Ctrl-double-click left mouse
Find feature when viewing extraction results	Ctrl-F
Open new window	Ctrl-N
Does not work for high resolution scans (2um, 3um, or 20-bit)	
Open file	Ctrl-O
Print TIFF image	Ctrl-P
Save the image as a different file	Ctrl-S

Changing Feature Extraction window displays

When you start the Feature Extraction software, you see the default window layout that Agilent configured before shipment. This window layout always appears unless you move windows and panes and change the default settings in the Preferences dialog box.

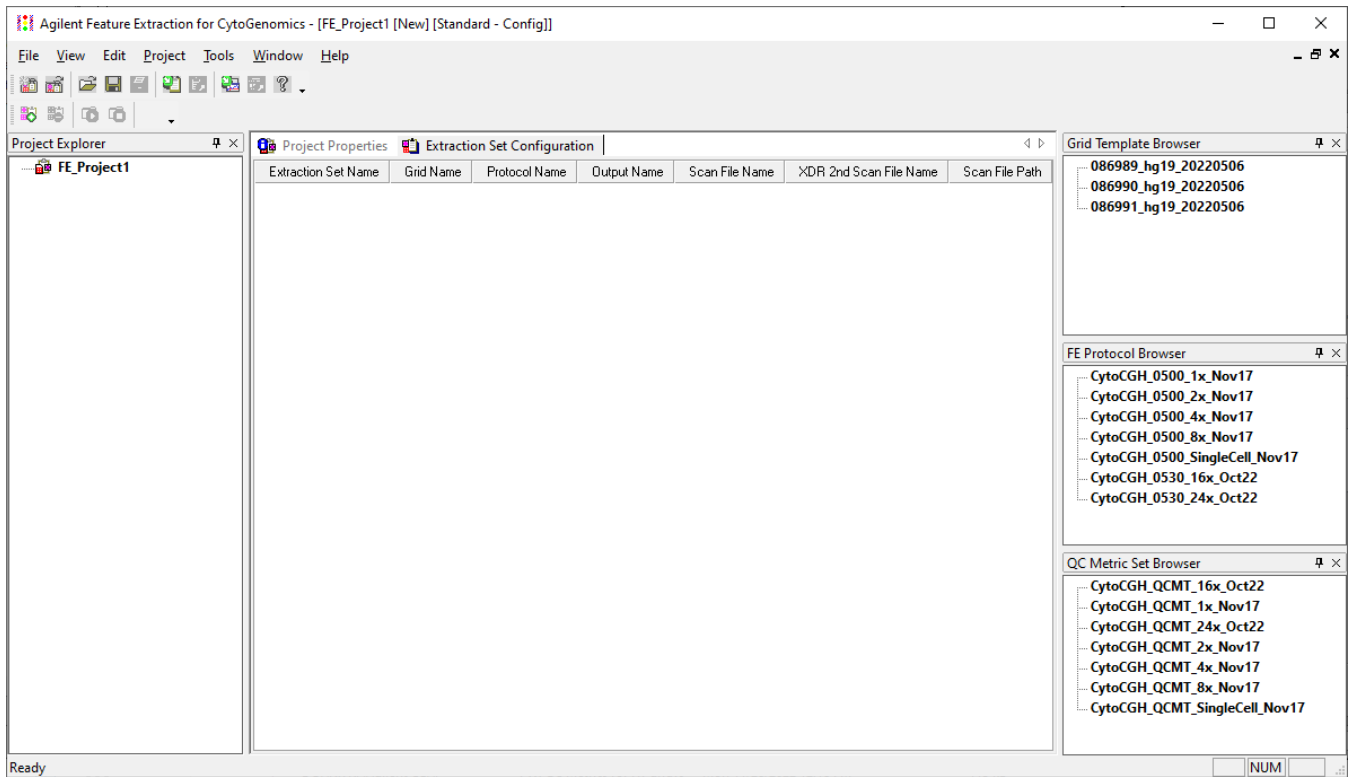


Figure 103 FE Standard Project Configuration window - default

This section shows you how to change from one layout to a new layout and to save the new layout so that it appears when the software starts.

To move windows and panes

Windows can be separated from window frames and blended into frames (“maximized”).

- Click the Restore Window icon for the Project Work Window.

The Project Window becomes separated from the main window frame, exposing the Workspace.

Panes can be unattached from one edge of a window frame and docked to another edge of the frame.

- 1 Click on the title of the Grid Template Browser Pane and drag it from the right edge of the window frame.
- 2 Move the Browser across the Workspace until the cursor hits the left edge of the frame, then release the mouse button.

The Grid Template Browser is now “docked” to the left edge of the frame. You can dock a pane to any edge of the frame in this way.

- 3 Repeat step 1 for the FE Protocol Browser Pane.
- 4 Move the Browser across the Workspace and place the cursor on the title of the Grid Template Browser.

Now both panes are overlapping one another with tabs at the bottom of the panes. To separate the two panes once more, click on a tab and move it to a new position.

If you had moved the FE Protocol Browser Pane to the left edge of the frame, the two Browsers would be next to each other.

Click on the title of the Project Explorer Pane and drag it across the Project window until the cursor hits the right edge of the window frame, then release the mouse button. Two other actions can be taken with panes:

- Close panes. Click on the X in the upper right corner.
- Autohide panes when docked. Click on the pin icon in the upper right corner of the pane.

The pane now appears as a bar button on the edge to which it was docked.

When you click on the resulting bar button, the pane floats from the edge so that you can view its contents. Note that the pin position is now horizontal instead of vertical. If you click on the pin now, the pin becomes vertical and the pane redocks itself to the edge of the frame.

To select the layout to use when software starts

You set the default layout for the Feature Extraction display in the Preferences dialog box.

- 1 Click Tools > Preferences from the top menu.
- 2 Change the following settings to produce the layout you want upon software startup.

To select the layout to use when software starts

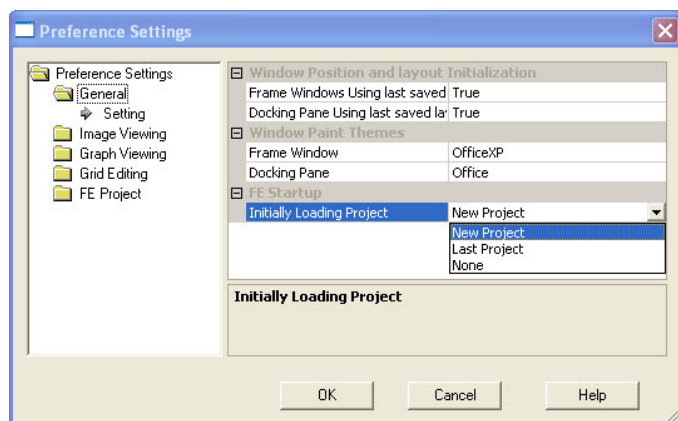


Figure 104 Preference Settings dialog box

Frame Windows Using last saved Position

If True, the window positions you have changed in the last session appear upon startup. If False, the factory default window position appears.

Docking Pane Using last saved Layout

If True, the docking positions from the last session appear upon startup. If False, the factory default docking positions appear.

Initially Loading Project

If New Project, FE Project1 always appears upon startup. If Last Project, the project on the screen before exiting appears at startup. If None, the Workspace is blank upon startup.

If you want to return the windows and panes to their original positions at any time, set "Frame Windows Using last saved Position" and "Docking Pane Using last saved Layout" to FALSE, close Feature Extraction and reopen the software.

In This Book

The *User Guide* presents instructions for using Feature Extraction for CytoGenomics software.

This guide provides instructions for:

- Setting up and running Feature Extraction on batches of image files.
- Creating grid files to assign to image files for automatic extraction.
- Changing the parameter values for each step in a protocol that can be assigned to an image file for automatic extraction.
- Working with image displays.

