Agilent Feature Extraction Software (v9.5)

User Guide

Research Use Only

Agilent Technologies

Notices

Edition

Fourth Edition, February 2007

Printed in USA

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

Agilent Recognized Trademarks

Adobe®; Acrobat®–Adobe, the Adobe Logo, Acrobat and the Acrobat Logo are trademarks of Adobe Systems Incorporated.

Pentium®–Pentium is a U.S. registered trademark of Intel Corporation.

Microsoft®–Microsoft is a U.S. registered trademark of Microsoft Corporation.

Rosetta Luminator™–Rosetta Luminator is a trademark of Rosetta Inpharmatics LLC.

Rosetta Resolver®–Rosetta Resolver is a U.S. registered trademark of Rosetta Inpharmatics LLC.

Windows NT®–Windows NT is a U.S. registered trademark of Microsoft Corporation.

Windows® and MS Windows®–Windows and MS Windows are U.S. registered trademarks of Microsoft Corporation.

Patents

Portions of this product may be covered under US patent 6571005 licensed from the Regents of the University of California.

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 Feature Extraction 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.

Contents

1
Working with Feature Extraction

This chapter gives you an introduction to the process of Feature Extraction and lets you know what you can do with the software to optimize the process.

Feature Extraction process

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 software on Agilent or non-Agilent microarrays scanned on an Agilent or GenePix scanner. The data analysis flow starts with removing outlier pixels and ends with entering Feature Extraction results into third-party data analysis programs to compare the results of different experiments.

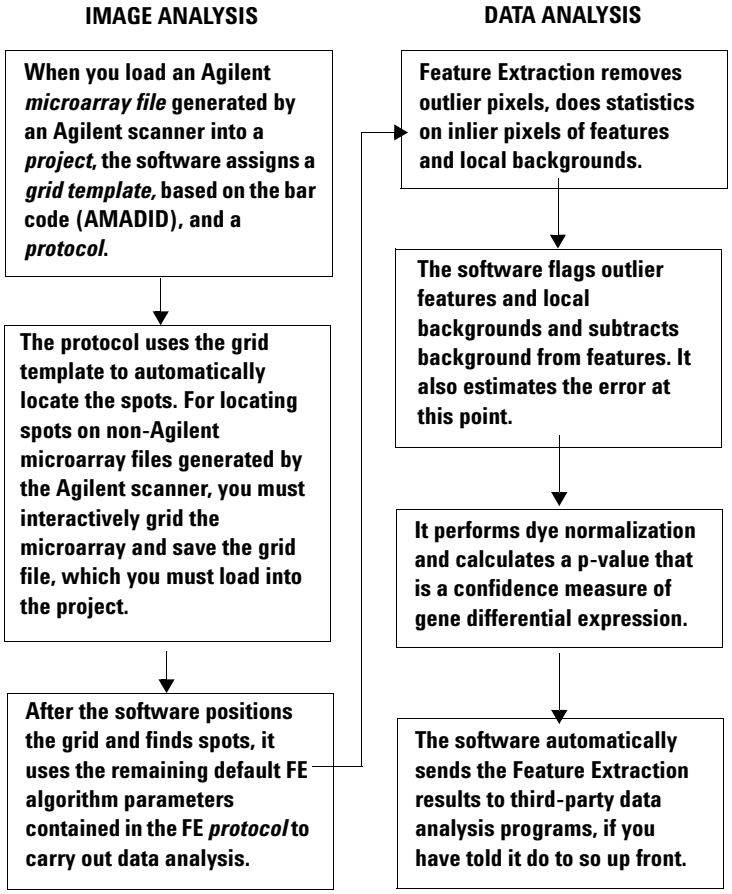


Figure 1 Process flow for Feature Extraction

Overview of what you can do with the software

Feature Extraction 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 FE data analysis.

Data from these microarray images can be extracted by the Agilent Feature Extraction software:

- Agilent microarray images scanned with either an Agilent or a GenePix scanner
  - Gene Expression (GE) 1-color — High Density, Single Pack and Multiplex (formerly known as Multipack) microarrays
  - GE 2-color — High Density, Single Pack and Multiplex microarrays
  - CGH — High Density, Single Pack and Multiplex microarrays
  - microRNA — These are in development; check the Agilent web site for the latest information.
- Non-Agilent microarray images scanned with an Agilent scanner

When you start an extraction, the software automatically extracts all the sets in the project and produces a project summary report, a QC chart for the batch 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 view significant anomalies that may interfere with your extraction, 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 algorithms for data analysis

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

For 2-color and CGH images you also set up the extraction to normalize for any differences in dye signal intensity. After that the software calculates a reliable log ratio, p-value and log ratio error for each feature to give you a confidence measure that a gene has been differentially expressed.

For microRNA (miRNA) images, Feature Extraction uses a special analysis, described in more detail in Chapter 2.

Extraction results

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 export a portion of these results as MAGE-ML files, as well as TIFF or JPEG image files, into GeneSpring and third-party analysis programs. The text and MAGE-ML results can be saved in both full and reduced formats.

You can load and view the visual 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 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:

- Create projects and save them to new names
- Load image files to create extraction sets. For Agilent image files, the software automatically loads a grid template to match the design file and a protocol specified in the grid template
- Drag and drop scan files on to the project
- Enter or change the protocol or grid template for the project
- Drag and drop grid templates or protocols onto each extraction set
- Associate a grid file with an extraction set
- Choose the type of result file to save
- Set up to transfer the files via FTP after extraction to Rosetta Resolver/Luminator
- Delete extraction sets
- Delete grid templates and protocols from the database
- Add grid templates and protocols to the database
- Edit protocols and grid template properties
- Change the order of extraction sets
- Choose to automatically stop Feature Extraction after a specified time when extraction is in real-time mode
- Decide whether eXtended Dynamic Range (XDR) scans from the Agilent scanner should be treated as a set and extracted together to generate a single combined output covering a wider dynamic range.

Run extraction

- Start and stop extracting
- Monitor up to four different extraction runs
- View a log history of the extractions in real time

What you can do with images before extraction

Find spots interactively for non-Agilent and Agilent images

- Flip and rotate the image from landscape to portrait mode and vice versa
- Interactively position grids to find spots on the microarrays
- Calculate the spot size and centroid positioning
- Compare the position of the nominal spot centroid (+) laid down by the grid with the found centroid position (X) for the spot
- Move the found centroid position (X) to where you want it on the spot
- Feature Extraction uses these manually adjusted centroid positions to define features rather than those calculated by the Place Grid algorithm.
- Select spots to ignore.
- The designated spots will not be used in Feature Extraction.
- Enter and save barcode information with the tiff image.

Change image display

- 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

What you can do with protocol algorithms

Protocols contain the values for the parameters used by Feature Extraction algorithms to calculate intermediate and final results. See “Feature Extraction process” on page 16.

What you can do with the algorithms

- Select the algorithms 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 algorithms do for you

These algorithms operate in the following order.

Place grid

- For Agilent microarrays using a grid template, automatically positions a grid and finds the centroid positions of each spot on the microarray
- For any microarray loaded with a grid file, takes the grid from the saved grid file and 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, provides the option of interactively placing the grid on the microarray

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 • Using a radius method, defines the perimeter of the local background surrounding the feature • Calculates statistics for the pixel populations of the features and local backgrounds, removing outlier pixels • Calculates signal statistics for inlier pixels
Flag outliers	• Flags non-uniform feature and background regions and population outliers
Compute background, signal biases and errors	• 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 a 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 than the background and significantly different from background using the error model calculation • Can calculate signals and background for low-intensity features using surrogates
Correct dye biases	• 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
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 calcualtes QC metrics for the QC Report.

MicroRNA Analysis

- Calculates the TotalGeneSignal for microRNA (miRNA) microarray images only.

  This analysis includes all the previous Feature Extraction steps but does not correct for dye biases or compute ratios.

  See Chapter 4, “Changing Protocol Settings” of this guide to learn how to change protocol settings for the miRNA microarray image extraction.

  See Chapter 5, “How Algorithms Calculate Results” of the *Reference Guide* to learn more about the algorithm in this step.

Generate Results

- Generates tab-text files, MAGE-ML files and QC Reports for each extracted image.
- Also generates a new file for the microRNA image in addition to the tab-text file, MAGE-ML file and QC Report.

  The new file is a GeneView.txt file and contains the SystematicName, the ControlType, the gTotalGeneSignal and the gTotalGeneError for the miRNA microarray.

  See Chapter 3, “Text File Parameters and Results” of the *Reference Guide* to learn more about this new text file.

What you can do with extraction images and .shp files

- Show and hide annotations on the results
- Search and display one feature at a time using feature number, probe name or gene name
- Get help for the results annotations from the visual results
- Print the images

What you can do with extraction results

- View a QC report (one of seven: GE 2-color, GE 1-color or CGH, each one with or without Agilent spike-ins; and miRNA) 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 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 or Luminator (FTP) or manually load the files into a database of your choice.

2 Extracting Microarrays Automatically

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).

Setting up/editing projects

Start the software to set up a batch extraction

You can start the Feature Extraction (FE) software in one of several ways.

- Double-click the **Feature Extraction** shortcut on your desktop.

  This shortcut was created when you installed the Feature Extraction software application.

- Select **Start > (All) Programs > Agilent > Feature Extraction > Feature Extraction 9.5**.

  The FE Standard Project Configuration window appears with FE Project1 displayed in Project Explorer.

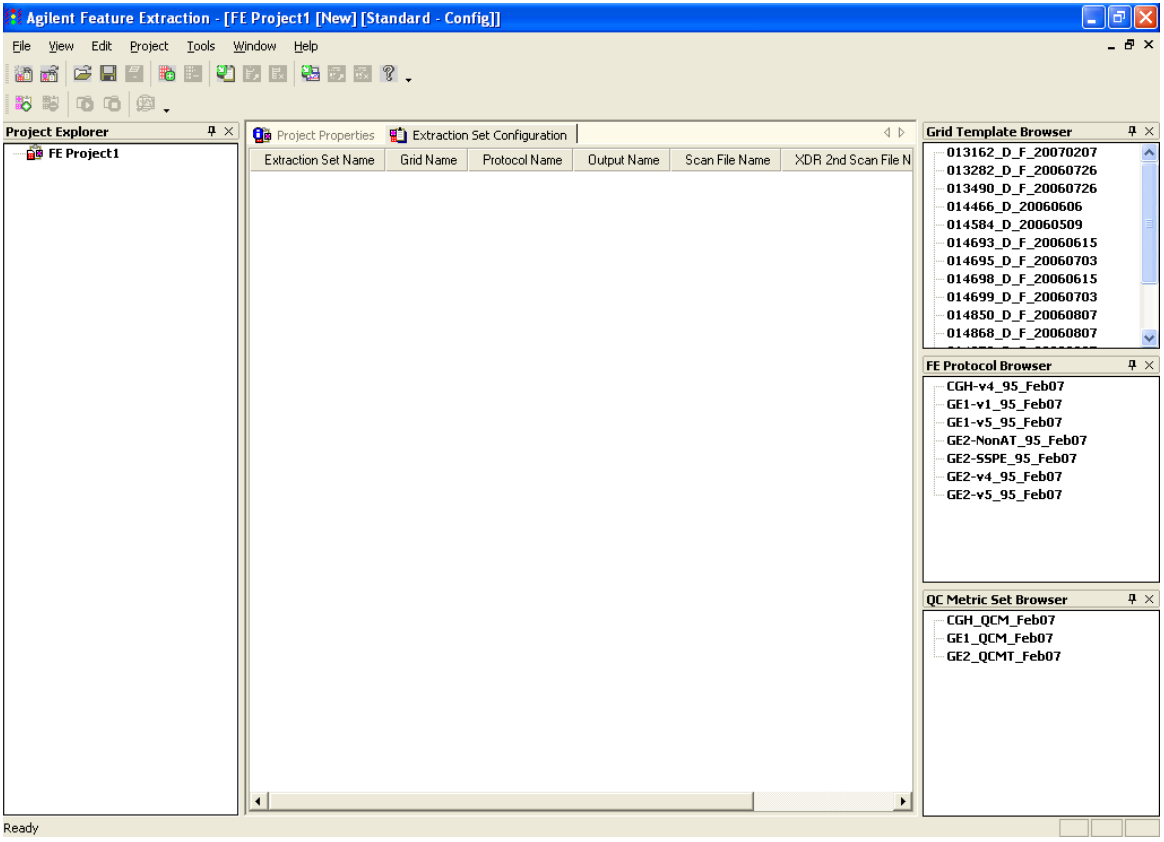


Figure 2 FE Standard Project Configuration window

If you do not see the QC Metric Set Browser when you first open the software, select ***View > QC Metric Set Browser****.*

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

Set up/edit a project to run existing image files (standard project)

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.

1 Open a standard project, if one is not already open.

See "Create a standard project" on page 31.

2 Change Project Properties.

See "Access project properties for individual projects" on page 36.

3 Add the image files to make up extraction sets.

See "Working with extraction sets" on page 62.

4 View the Extraction Set Configuration tab sheet to make certain all information is correct for each extraction set.

See "View/edit the Extraction Set Configuration table" on page 67.

5 If any of the extraction sets does not contain a grid template or grid file, add one.

See "Add a grid template to an extraction set" on page 69.

6 If the extraction set does not contain a protocol, add one.

See "Add a protocol to an extraction set" on page 81.

7 Select a QC metric set from those installed with FEv9.5, or import a new one if desired.

This allows you to display a QC Chart showing the selected QC metrics after the batch is processed.

See "Working with QC Metric Sets" on page 83.

8 Save the project.

See "Save a project" on page 33.

Set up/edit a project to run image files in real time (on-time project)

You can run Feature Extraction for Agilent microarrays in real time as image files are saved after scanning. 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 images, not non-Agilent ones.
- You do not add image files or extraction sets (image files, grid templates, protocols) to the on-time project.
- With an on-time project, you can run Feature Extraction only on Agilent images that have grid templates based on their design files loaded in the database and a default protocol specified within the grid template properties.
- You can enter 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 32.

2 Change Project Properties.

See "Select a time to terminate real-time Feature Extraction" on page 46.

3 Save the project.

See "Save a project" on page 33.

Create a project

Create a standard project

To create a new standard project:

- Select **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 Project N or not or close FE Project N 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

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

  Notice that the Project Explorer and the Extraction Configuration tab sheet do not appear on screen.

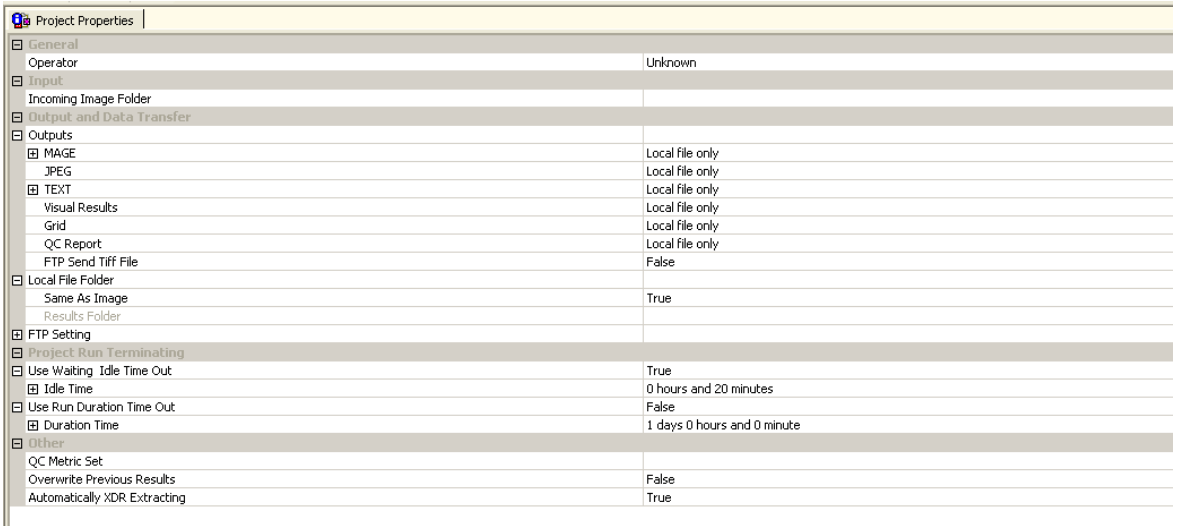


Figure 3 On-Time Project Configuration window

Open a project

To open an existing standard or on-time project:

1 Select **File > Open > Project**,
Or
Click the **Open project file** icon,

2 Select the project name, and click **Open**.

Save a project

To save a standard or on-time project to a new name:

1 Select **File > Save As**.

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

2 Enter the new name, then click **Save**.

If you are saving the project for the first time, you can click the Save icon and enter 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

Program Files\Agilent\Microarray\FeatureExtraction\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 appears.

Back up and restore the database

You cannot restore a v.9.1 database to v.9.5 software, nor can you restore a v.9.5 database to v.9.1 software.

You can use this utility to restore a v.9.1 database to v.9.1 software, or a v 9.5 database to v9.5 software. For instructions, see "Restore a v.9.1 database to v.9.1 software" on page 35.

You can easily back up and restore the database containing protocols, grid templates, grid files and design files. We recommend that you back up your database every month.

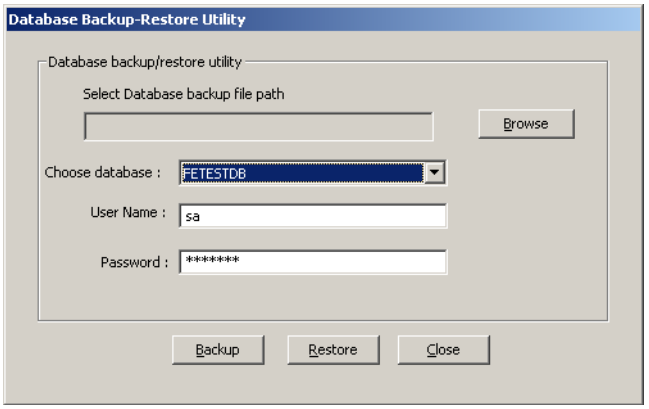


Figure 4 Backup Restore Utility dialog box

Backup and restore the current database

1. Select **Start > (All) Programs > Agilent > Feature Extraction > Tools > Backup Restore Utility**.
2. To set the path for your backup or restore database, click **Browse**.
3. Select the directory, and click **Open**.
4. Select the database that you want to back up or restore.
5. For the **User Name**, enter "`sa`", if not already there.
6. For the **Password**, enter "`welcome`", if not already there.
7. Click **Backup**, or click **Restore**.

Restore a v.9.1 database to v.9.1 software

If, after you have installed v.9.5 software, you want to use your v.9.1 database, you must use it with v.9.1 software. Follow these instructions to be able to use this utility to restore the database.

1 Uninstall v9.5 software.

   This process uninstalls the Backup Restore utility.

2 Install v9.1 software.

3 Copy the Backup Restore utility from the CD to the directory,

   C:\Program Files\Agilent\Microarray\Feature Extraction

4 Follow the instructions to "Backup and restore the current database" on page 34.

Changing properties for individual projects

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" on page 48.

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

Access 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 the figure below.

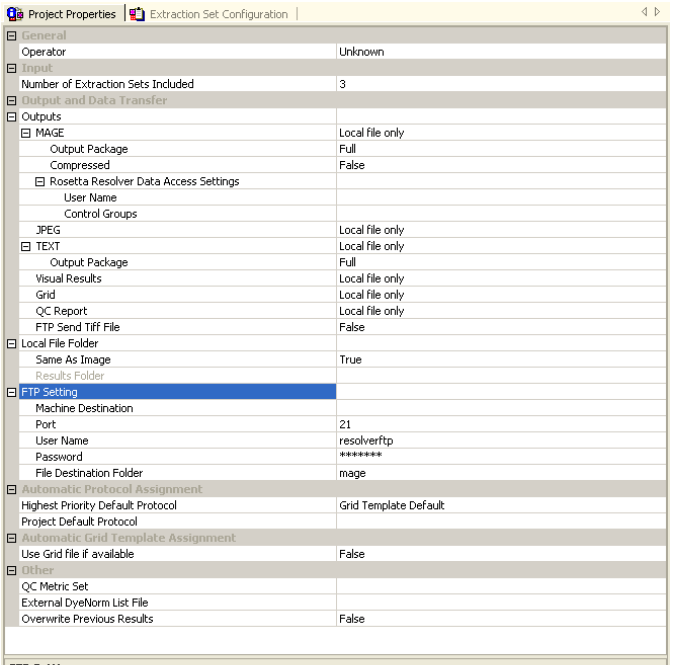


Figure 5 Standard Project Properties tab, expanded

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 sheet.

The Project Properties tab for On-time Projects looks like the figure on the next page.

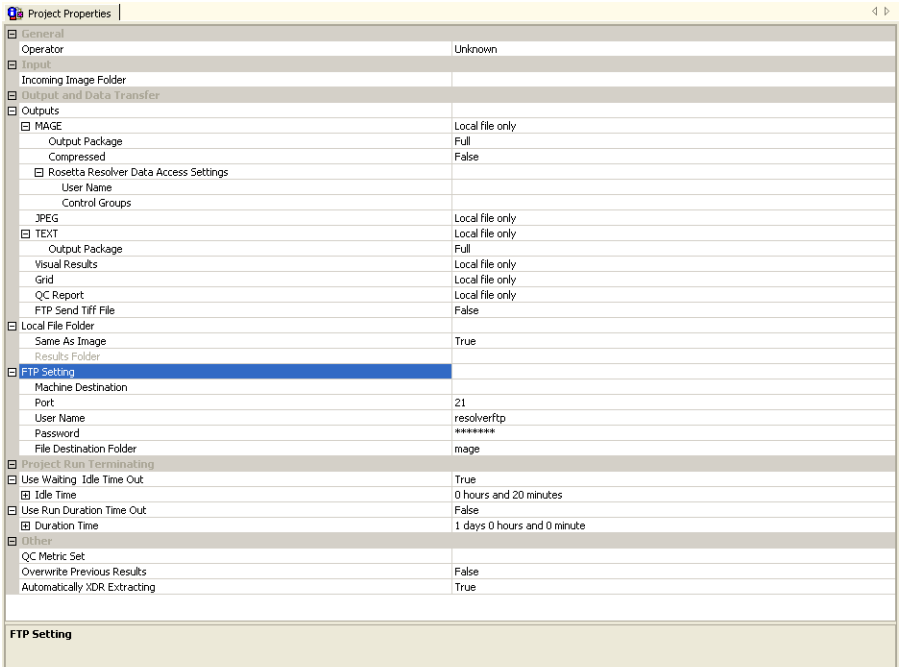


Figure 6 Project Properties tab for on-time projects, expanded

The parameters and settings in this sheet 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 "Select the incoming image folder" on page 45.

- Capacity to terminate the run automatically

  See "Select a time to terminate real-time Feature Extraction" on page 46.
- Capacity to automatically extract XDR images

  See "Turn off the option to automatically extract XDR images" on page 47.

Change general project property settings

Change to new values

- Double-click on 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 on the arrow to the right of the cell, and select from the list.

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 select **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 enter 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.

Select the location for saving and/or sending output files

See the 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 40 and "Select the local file folder for saved results" on page 42.

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 42.

FTP send only Select to send the files to an external location. See "Change the FTP settings to send result files automatically to Rosetta Resolver/Luminator" on page 42.

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/Luminator and Array Express databases.

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

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** 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 enter names for these settings, no access control information is entered into the MAGE-ML file. Luminator cannot accept access control information.

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

Rosetta Resolver Data Access Settings: Control Groups Enter 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, enter "`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 "Change the compression factor for the JPEG output" on page 204 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, or a **Compact** (reduced) set.

Visual Results Visual results from the feature extraction process are saved to a .shp or "shape" file, which only Agilent FE software can read.

Grid This is the grid information saved to Agilent grid.csv and 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. The type of QC Report saved is determined by the protocol type for the extracted data: 2-color gene expression (GE), 1-color GE or CGH, each one containing, or not containing, spike-in data, or the miRNA protocol type.

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 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** to select **False**.
2. Click the cell next to **Results Folder**.
3. Click the **Browse** button.
4. Select the directory to contain the results.
5. Click **OK**.

Change the FTP settings to send result files automatically to Rosetta Resolver/Luminator

If you use the same settings repeatedly, the values specific to your installation should be set up in the Preferences dialog box. See "Change the default FTP settings for automatic data transfer" on page 55.

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

Machine Destination This field is specific to your Rosetta Resolver or Luminator server name. Please contact your system administrator for the name to enter.

Port Enter the number of the FTP port on the server. The default is port 21.

User name Enter 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 and **luminatorftp** for Luminator servers.

Password This field already contains the default password for Rosetta Resolver and Luminator 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.

Set up to automatically assign protocols

Select the 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. For more information on extraction sets, grid templates and protocols, see "Working with extraction sets" on page 62.

1 Click the 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 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.

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.

Select to use a grid file

- 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.

Select to use a QC metric set

Agilent QC Metric Sets are named as follows:

microarray name_QCM (or QCMT)_date (monthyear)

QCM stands for QC Metrics. QCMT stands for QC Metrics and Thresholds

Example: GE2_QCMT_Feb07

1 On the Project Properties tab, click the cell to the right of **QC Metric Set**.

2 Click the down arrow and select the metric set from the list.

When the software was installed, Agilent QC Metric Sets were also installed and appear in this list.

To import additional QC Metric Sets

1 Select **Tools > QC Metric Set > Import**...

2 Select a metric set or sets provided by Agilent or that you created in the QC Chart Tool, and click **OK**.

OR

1 Select **View > QC Metric Set Browser**.

The QC Metric Set Browser appears below the FE Protocol Browser in the lower right-hand corner of the main window.

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

3 Select a metric set or sets provided by Agilent or that you created in the QC Chart Tool, and click **OK**.

See "QC Chart Tool Instructions" accessible through http://www.agilent.com/chem/feqcmetrics.

Select an external DyeNorm list file

See "Create a new internal DyeNorm gene list" on page 76 to learn how to create an external DyeNorm list from the dye norm list editor.

If the protocol calls for a DyeNorm gene list and the image must be extracted with a grid file, use an external DyeNorm list. This list must have the order of the column names as Probe Name, Gene Name, Systematic Name. This is the same list order for a dye norm list exported by the dye norm list editor.

1 Click the cell next to **External DyeNorm List File**.

2 Click the **Browse** button.

3 Select the .txt file, and click **Open**.

Select to overwrite previous results

- 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.

Select the incoming image folder

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**.

Select a time to terminate real-time Feature Extraction

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 Enter the times.

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 Agilent scanner will be extracted in a combined fashion and a single output file set covering a broader range of data will be generated.

If the setting False, the two images will be extracted seperately and two output file sets will be generated.

Changing default project properties

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

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, you open the Preferences dialog box.

1 Select **Tools > Preferences**.

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

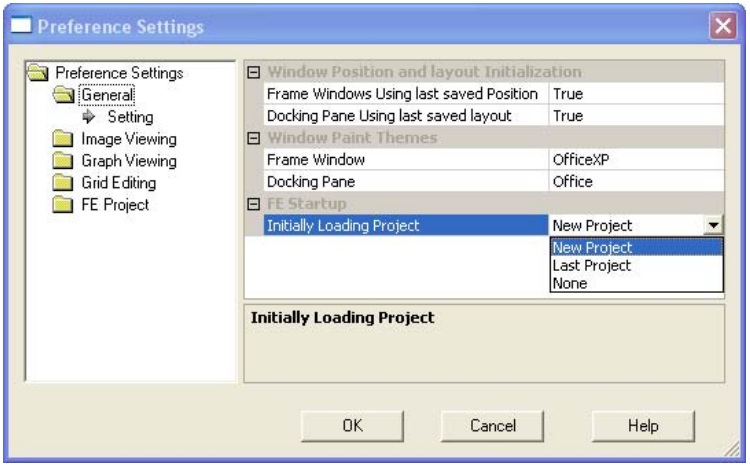


Figure 7 Preferences dialog box

Access the General default project properties

1 Open the Preferences dialog box.

2 Open the **FE Project** folder, if not already open.

3 Open the **General** folder, if not already open.

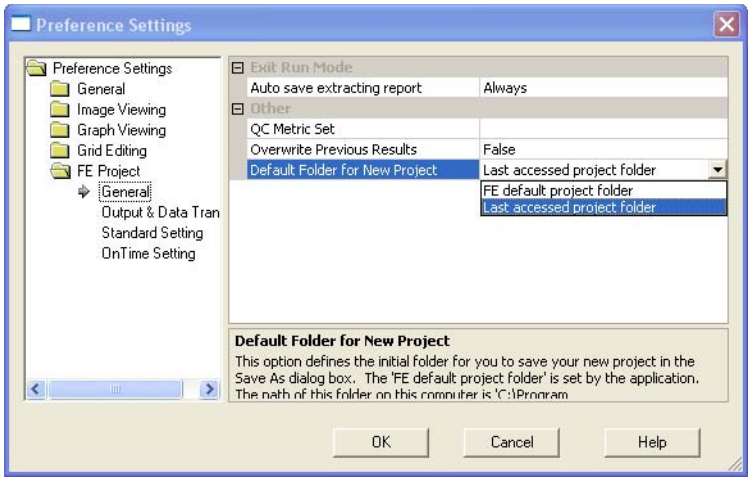


Figure 8 Preferences dialog box—FE Project General settings

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 Access the General default project properties.

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 will never be saved as a separate .rtf file after generation. You will be 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 will not be 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.

Select a default QC metric set

Agilent QC Metric Sets are named as follows:

microarray name_QCM (or QCMT)_date (monthyear)

QCM stands for QC Metrics. QCMT stands for QC Metrics and Thresholds

Example: GE2_QCMT_Feb07

1 Access the General default project properties.
2 Click the cell to the right of **QC Metric Set**.
3 Click the down arrow and select the metric set from the list.

When the software was installed, Agilent QC Metric Sets were also installed and appear in this list.

To import additional QC Metric Sets

1 Select **Tools > QC Metric Set > Import**...
2 Select a metric set or sets provided by Agilent or that you created in the QC Chart Tool, and click **OK**.

OR

1 Select **View > QC Metric Set Browser**.

The QC Metric Set Browser appears below the FE Protocol Browser in the lower right-hand corner of the main window.

2 Right-click the QC Metric Set Browser, and select **Import...**
3 Select a metric set or sets provided by Agilent or that you created in the QC Chart Tool, and click **OK**.

See “QC Chart Tool Instructions” accessible through http://www.agilent.com/chem/feqcmetrics.

Select to overwrite previous results

1 Access the General default project properties.

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.

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 Access the General default project properties.

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:\Program Files\Agilent\Microarray\FeatureExtraction\FEProjects”.

Access the default properties for Output and Data Transfer

1 Open the Preferences dialog box.

2 Open the **FE Project** folder.

3 Open the **Output & Data Transfer** folder.

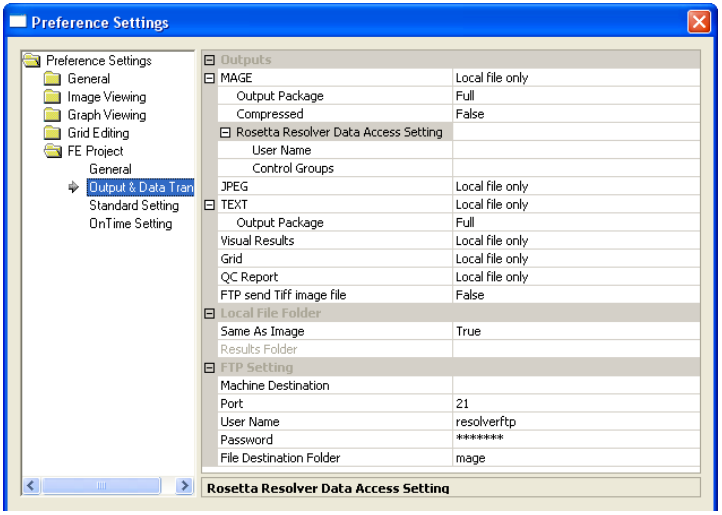


Figure 9 Preferences dialog box—Output & Data Transfer settings

Select the default location for saving and/or sending output files

- Access the default properties for Output and Data Transfer

See the 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 53 and "Select the default local file folder for saved results" on page 55.

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).

Local file only Select to save the files to a local file folder. See "Select the default location for output files" on page 53.

FTP send only Select to send the files to an external location. See "Change the default FTP settings for automatic data transfer" on page 55.

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/Luminator and Array Express databases.

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

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 enter names for these settings, no access control information is entered into the MAGE-ML file. Luminator cannot accept access control information.

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

Rosetta Resolver Data Access Settings: Control Groups Enter 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, enter "`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 "Change the compression factor for the JPEG output" on page 204 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, or a **Compact** (reduced) set.

Visual Results Visual results from the feature extraction process are saved to a .shp or "shape" file, which only Agilent FE software 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 This is the grid information saved to Agilent grid.csv and 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. The type of QC Report saved is determined by the protocol type for the extracted data, either 2-color gene expression (GE), 1-color GE or CGH, each one either containing spike-in data, or not.

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** to select **False**.
2. Click the cell next to **Results Folder**.
3. Click the **Browse** button.
4. Select the directory to contain the results.
5. Click **OK**.

Change the default FTP settings for automatic data transfer

1. Access the default properties for Output and Data Transfer
2. In the Project Properties tab, expand the **FTP Setting** folder, and enter the values for these FTP settings:

Machine Destination This field is specific to your Rosetta Resolver or Luminator server name. Please contact your system administrator for the name to enter.

Port Enter the number of the FTP port on the server. The default is port 21.

User name Enter 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 and **luminatorftp** for Luminator servers.

Password This field already contains the default password for Rosetta Resolver and Luminator 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.

Access the default properties for standard projects

1. Open the Preferences dialog box.
2. Open the **FE Project** folder.
3. Open the **Standard Setting** folder.

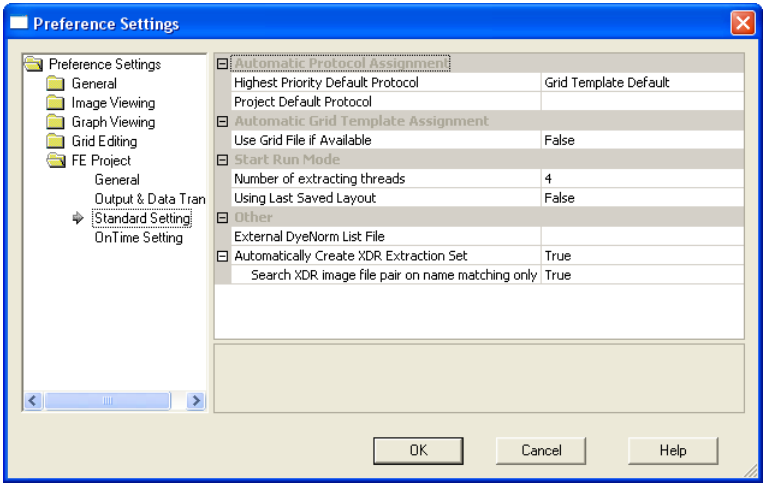


Figure 10 Preferences Dialog Box—Default standard project properties

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 62.

1. Access the default properties for standard projects.
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.

Change the default protocol for all projects

1. Access the default properties for standard projects.
2. Click the cell next to **Project Default Protocol**.
3. Select a protocol from the list.

Select to use a grid file

1. Access the default properties for standard projects.
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.

Change the Start Run Mode setting

These settings specify the maximum number of extractions that can be run simultaneously and the number of running monitors.

1 Access the default properties for standard projects.

2 Change the settings for the two options.

Number of extracting threads The choices are 1, 2 and 4. If 4 is selected and 3 extraction sets are to run, the software uses only two "threads" at a time to run them. This means that the software runs two extraction sets simultaneously. When one is complete, the extraction of the third set begins even if the other set has not completed.

If your PC has enough memory, using multiple threads can speed up batch extraction. A PC with 1GB of RAM can handle 4 threads running 44k images at a 10 micron scan. A PC with 2GB of RAM can handle 2 threads running 244k images at a 5 micron scan.

If you see an error message that says "MemoryMapImage: There is not enough memory for the application to open this image...", you should reduce the number of threads being used.

Using Last Saved Layout If True, software uses the layout that was saved in the previous project before exiting the software. If false, it uses the default layout from FE installation.

Select an external DyeNorm list file

See "Create a new internal DyeNorm gene list" on page 76 to learn how to create an external DyeNorm list.

If the protocol calls for a DyeNorm gene list and the image must be extracted with a grid file, use an external DyeNorm list. This list must have the order of the column names as Probe Name, Gene Name, Systematic Name. This is the same list order for a dye norm list exported by the dye norm list editor.

1 Access the default properties for standard projects.
2 Click the cell next to **External DyeNorm List File**.
3 Click the **Browse** button.
4 Select the .txt file, and click **Open**.

Select to create XDR extraction sets automatically

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

1 Access the default properties for standard projects.
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.

Access the on-time default properties

1 Open the Preferences dialog box.
2 Open the **FE Project** folder.
3 Open the **OnTime Setting** folder.

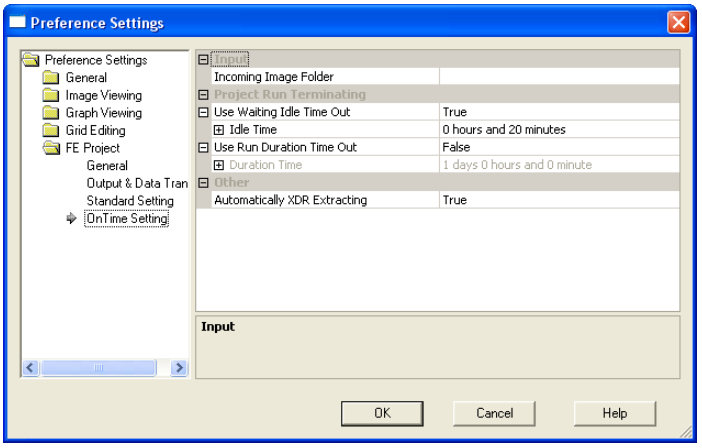


Figure 11 Preferences Dialog Box—Default on-time project properties

Select the default image folder

1. Access the on-time default properties.
2. Click on the cell next to **Incoming Image Folder**.
3. Click the **Browse** button.
4. Double-click the image folder, and click **OK**.

Select a default time to terminate real-time Feature Extraction

1 Access the on-time default properties.

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 Enter the times.

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 Access the on-time default properties.

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.

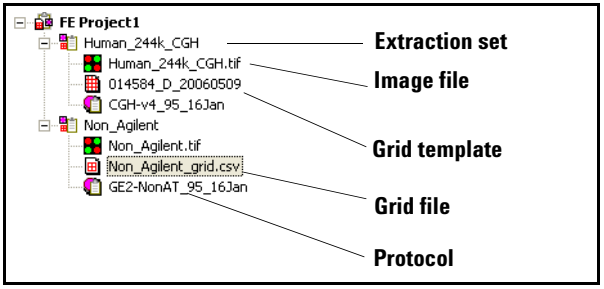


Figure 12 Extraction sets in Project Explorer

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 and is represented with an icon that marks it as an XDR extraction. See Figure 13.

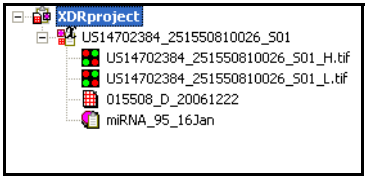


Figure 13 XDR extraction image files marked with XDR icon

grid template Grid information from Agilent 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 selects a grid template from the database that matches the design file associated with the image file.

For most Agilent microarray images, the grid template should be automatically assigned. 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 should 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.

Add extraction sets (image files) to the standard project

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

- Dragging and dropping the image file onto Project Explorer
- Selecting the extraction set through the FE Project toolbar.
- Selecting the image file through the Edit menu.
- Selecting 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.

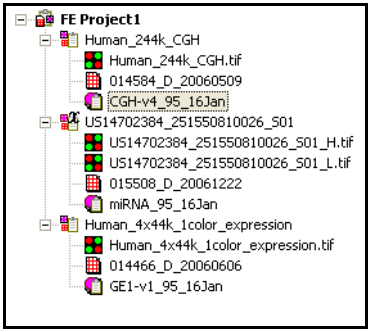


Figure 14 Project window after adding image files (extraction sets)

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 Select **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 Select **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 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).

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 Select **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.

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 select **Delete**.

The extraction set is immediately deleted.

OR

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

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 with the CTRL key.

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

The extraction sets(s) are immediately deleted

OR

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

If an extraction set or sets in the Extraction Set Configuration table are selected, a message appears 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 appears 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.

View/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.

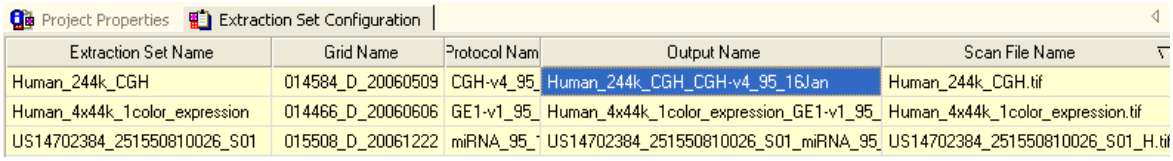


Figure 15 Extraction Set Configuration table

The Extraction Set Configuration tab sheet 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.

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 belonging to the extraction set.

- Double-click this name to see the image in the scan file.

Scan File Path The directory containing the scan file.

Working with grid templates/files

Show or hide the Grid Template Browser

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

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

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

Add a grid template to an extraction set

If you add a 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, but gridding or annotation may be incorrect.

If there is no 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.

Add a grid file to an extraction set

Although you can add a grid file for Agilent microarrays, you usually add a grid file for non-Agilent microarrays.

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

Add or remove a grid template to or from the database

To add a grid template to the database:

1. Select **Tools > Grid Template > Add**...

   OR

   Right-click the Grid Template Browser, and select **Add**...

2. Select a design file or grid file from the list, and click **Open**.

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

To remove a grid template from the database:

1. Select **Tools > Grid Template > Remove**.

   OR

   Right-click the Grid Template Browser, and select **Remove**.

2. Click **Yes** when asked if you're sure you want to remove the grid template from the database.

Change the grid template used in an extraction set

If you want to use a different grid template than that associated with the image file, 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, replacing the previous grid template.

Change the grid file used in an extraction set

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

1 Right-click the second line of the extraction set, and select **Select grid file** from the menu.

2 Select the .csv file that you want to use with the image file, and click **Open**.

Access 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.

You can also access grid mode from the image display. See "Access grid mode from the image file" on page 100.

Access from a grid template

1 In Project Explorer, right-click the grid template in the extraction set.

2 Select **Create and use grid file....**

The image file for the extraction set appears in grid mode, ready for gridding.

Access from a grid file

1 In Project Explorer, right-click the grid file (.csv) in the extraction set.

2 Select **Edit Grid...**.

The image file for the extraction set appears in grid mode, ready for gridding.

See the next chapter in this guide to learn how to set up or edit a grid file.

Update grid template information online (via Agilent eArray)

As new genomic information becomes available, Agilent updates its design files. These updates are available for downloading through the Agilent eArray web site.

Online Update requires an eArray user login. If you don't have one, you should sign up for one at www.agilent.com/chem/earray.

Feature Extraction provides a link to load new grid templates into the Feature Extraction database that contain the latest annotation available on eArray.

1 Right-click anywhere in the Grid Template Browser, and select **Online Update...**.

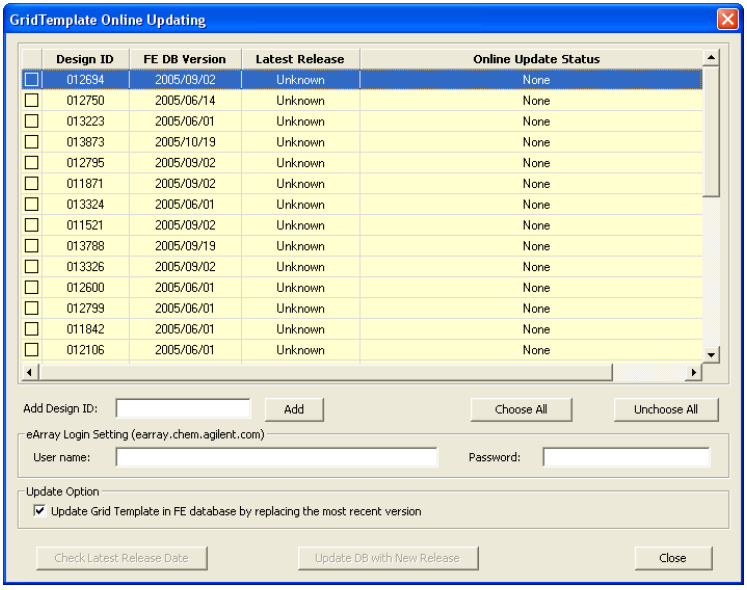


Figure 16 Grid Template Online Updating dialog box

2 Mark the Design IDs whose release dates you want to check or whose updates you want to download from eArray.

More than one version of the grid template can reside in the database. The update function lets you choose to add the latest version from eArray to already existing versions or to overwrite the most recent version already in the database.

Because the database has limited space, leave the default setting to overwrite the old version unless you have a compelling reason not to.

If you want to update or check the release dates of all of them, click **Choose All**.

If you want to add a Design ID not in the list, in the **Add Design ID** field, enter the ID and click **Add**.

3 To activate the bottom buttons, enter your **User name** and **Password** for eArray.

4 To check the release dates, click **Check Latest Release Date**, or
To overwrite the most recent grid template in the database, click **Update DB with New Release**.

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

View or change grid template properties

1 Double-click the grid template whose properties you want to view or change

2 View the grid template information.

You can view all of the grid information; you can change the default protocol (Agilent or GenePix 2-color, CGH or miRNA), the default 1-color protocol and the default DyeNorm gene list.

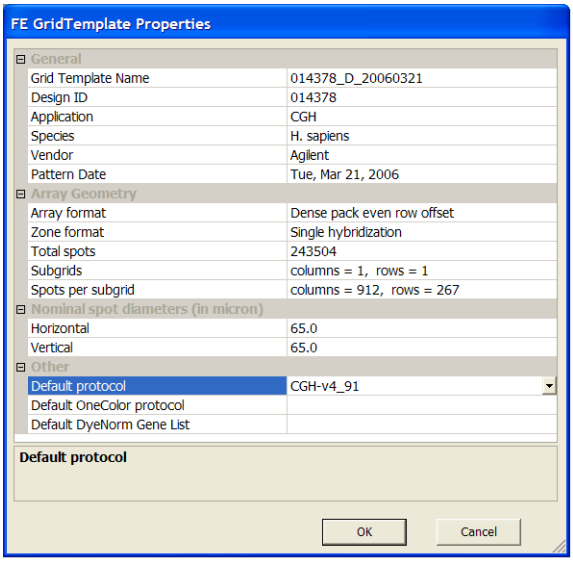


Figure 17 Grid Template Properties dialog box

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.

Add or change the default 1-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 recognizes to use the 1-color protocol when the 1-color scan file is loaded.

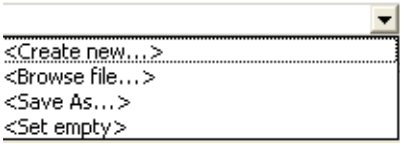
1 Double-click the grid template whose protocol you want to add or change.

2 Click the cell to the right of **Default Onecolor protocol**, and select a 1-color protocol from the list.

Change the default DyeNorm gene list

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

1 Click the arrow to the right of the cell next to **Default DyeNorm Gene List**.



2 Select one of the following five options:

DyeNorm Gene List Name This is blank if no DyeNorm Gene Lists are available or if you select Set Empty. If you previously created an internal DyeNorm gene list or lists or selected an external gene list for the default list, select a name from this list for the default list. Double-click the list to bring up the DyeNorm Editor for editing the list.

Create New Select this option if you want to create a new internal list from the existing probes. See the task on the next page.

Browse file Select this option to browse and select an external gene list for the default list and for editing.

Save As Saves the default list as an external list with a new name.

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

Create a new internal DyeNorm gene list

1 Click the cell to the right of **Default DyeNorm Gene List**.

2 Select **Create New**.

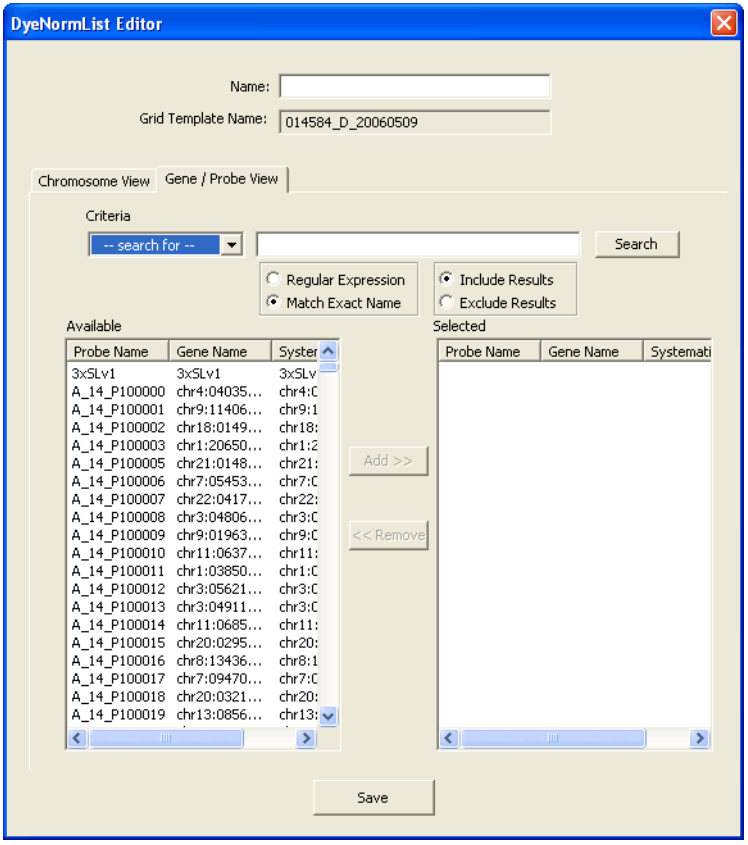


Figure 18 DyeNorm List Editor—Gene/Probe View

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.

1. If necessary, click the **Gene/Probe View** tab.
2. Select **Probe Name**, **Gene Name** or **Systematic Name** as the type of name to search for from the pull-down menu.
3. Select to search according to the exact name, **Match Exact Name**, or to any part of the name, **Regular Expression**.
4. Enter the name or part of the name.
5. Select to include or exclude results.
6. Click **Search**.

Match Exact Name If selected, you must enter the exact probe name, gene name or systematic name, depending on the type of name you selected in step 1.

Regular Expression If selected, you can enter any part of the probe name, gene name or systematic name.

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.

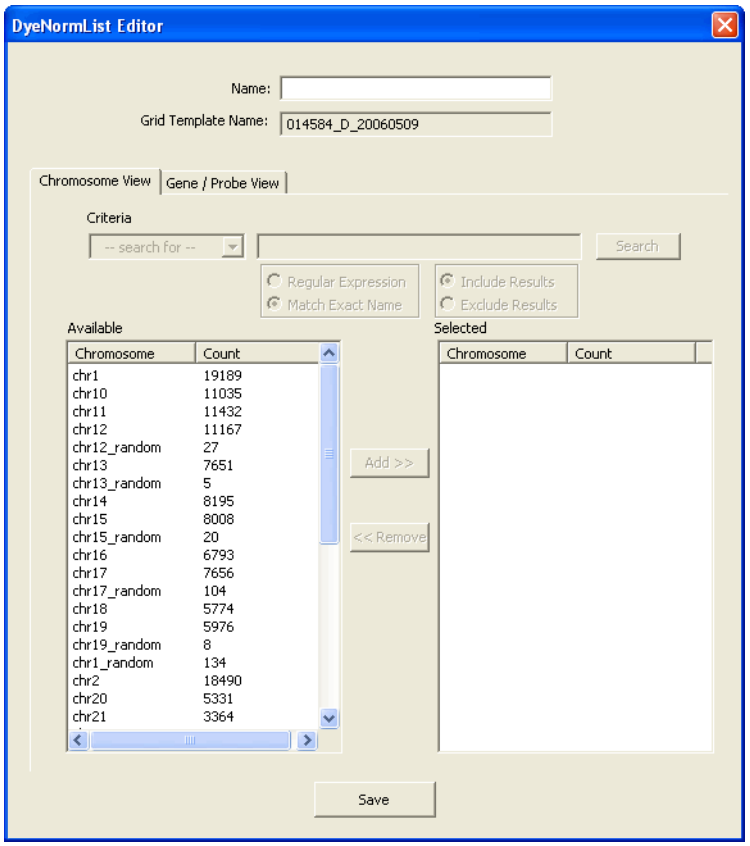


Figure 19 DyeNorm List Editor—Chromosome View

2 Add/remove genes or chromosomes to list interactively.

Add/remove genes or chromosomes to list interactively

1 Highlight the gene or genes of interest in the Available Genes on Array list (**Shift** for contiguous names, **CTRL** 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 DyeNorm gene list to internal file

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

1. Enter the **Name** of the DyeNorm gene list
2. Click **Save**.

When you click **Save**, a list is created with the Dye Norm List Name and placed in the internal list available to the grid template.

Edit a DyeNorm gene list

1. Click the cell to the right of **Default DyeNorm Gene List**.
2. Select an existing gene list.
3. Double-click the list to bring up the **DyeNorm List Editor**.
4. Add and remove any genes or chromosomes that you want.
5. Click **Save**.

Import external Dye Norm List to create a new internal list

1. If you want to create a new list from an external .txt file, click the cell next to **Default DyeNorm Gene List** and click the down arrow.
2. Click **Browse file**, select the 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.
3. Click the down arrow to the right of the cell next to **Default DyeNorm Gene List,** and select the external gene list as the default list.

4 Double-click the list to bring up the DyeNorm Editor.

5 Add or remove genes and/or chromosomes from the list.

6 Enter the **Name** for the DyeNorm gene list, and click **Save** to create an internal list.

Create an external DyeNorm gene list

After you create an external DyeNorm gene list file, you can then select it for a project.

1 Create a new internal DyeNorm gene list.

2 Click the down arrow to the right of the cell next to **Default DyeNorm Gene List**, and select the new gene list.

3 Click **Save As**.

4 Name the file, if necessary, and click **Save**.

The file is saved as an external .txt file.

Working with protocols

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:

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

Add a protocol to an extraction set

If there is no protocol in the extraction set, follow these instructions:

1. Select a protocol from the FE Protocol Browser.
2. Drag and drop the protocol onto the third line in the extraction set.

   The new protocol appears in the extraction set.

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. Drag and drop the protocol onto the extraction set.

   The new protocol appears in the extraction set, replacing the previous protocol.

Change protocol settings

- Double-click a protocol in the FE Protocol Browser to bring up the Protocol Editor.
- Change the settings, and click **OK**.

See Chapter 4, “Changing Protocol Settings””.

Import or export a protocol to or from the database

To import a protocol to the database:

1 Select **Tools > FE Protocol > Import**...

OR

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

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

To export a protocol from the database:

1 Select **Tools > FE Protocol > Export**...

OR

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

Remove a protocol from the database

*The **Derived from** message can help you keep track of user-created protocols. If the protocol is Agilent-created, the message indicates this. If the protocol is user-created, the message lists the original Agilent protocol from which the new protocol was derived.*

1 Select **Tools > FE Protocol > Remove**.

OR

Right-click the FE Protocol Browser, and select **Remove**.

2 Click **Yes** when asked to make sure that you want to remove this protocol.

Agilent-created protocols cannot be removed from the database.

Working with QC Metric Sets

You have the option of using QC metric sets that give you additional statistics to those within the default QC report 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.

You can use the Agilent QC Metric Sets installed with the FE software, additional or updated Agilent QC Metric Sets downloaded from the web site, or QC metrics sets created with the QC Chart Tool.

Go to http://www.agilent.com/chem/feqcmetrics to download updated or new metric sets.

Show or hide the QC Metric Set Browser

This Browser contains the most recent Agilent QC metric sets after the software is installed. To show or hide the QC Metric Set Browser:

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

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

To import a QC Metric Set to the database:

1 Select **Tools > QC Metric Set > Import**...

OR

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

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

To export a QC Metric Set from the database:

1 Select **Tools > QC Metric Set > Export**...

OR

Right-click the QC Metric Set Browser, and select **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 remove a QC Metric Set from the database

1 Select **Tools > QC Metric Set > Remove**.

OR

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

2 Click **Yes** when asked to make sure that you want to remove this protocol.

To associate a QC Metric Set with a project

You associate a QC Metric Set with a project on the Project Properties sheet. You cannot associate a QC Metric Set with an individual extraction set.

1 Click the **Project Properties** tab.

2 Right-click the cell next to **QC Metric Set**.

3 Click the down arrow, and a the **QC Metric Set** from the list.

Running Feature Extraction

Several criteria must be met before you can run Feature Extraction on a project.

- No extraction set in the project can contain a red X.
- The Extraction Configuration sheet must contain all the intended names, path files, grid templates and protocols.

Start Feature Extraction

1 Click the **Project Run mode On/Off** icon, or select **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 enter a new name, and click **Save**.

If you run Feature Extraction on the full image of Agilent multiplex (formerly known as multipack) microarrays (44k or 244k formats) with subarrays, FE produces one of each type of file for the multiplex microarray, not for each subarray.

Up to four monitors are automatically turned on and appear at the bottom of the screen. The number depends on these criteria:

- The number of extraction sets being extracted
- Whether you turned any monitors off
- The number of extracting threads you set in the Preferences dialog box

See "Access the default properties for standard projects" on page 56.

NOTE

If you run Feature Extraction on a 2-color microarray with a 1-color protocol, Feature Extraction only operates on green channel data. If you run Feature Extraction on a 1-color microarray with a 2-color protocol, an error will occur in the Dye Normalization step.

Turn on/off monitoring for Feature Extraction

See "Change the Start Run Mode setting" on page 58 to understand how to set up the monitors during a run.

1 Select **View** > **Running Monitor** > **Monitor N**, where N is a number from 1 to 4.

2 Mark or clear the check box next to the monitor you want on or off.

Stop Feature Extraction

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

Create a new project with remaining extraction sets if run fails

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 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. FE 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.

Viewing results

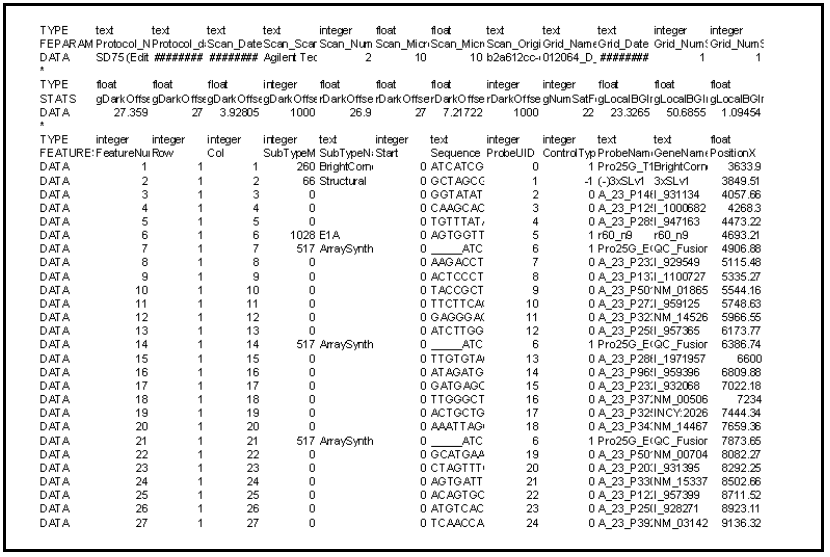
View the text result file in Microsoft Excel

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 directory specified for containing the results.
2. Right-click the text results file (protocol name on end), and select **Open with...**.
3. Select **Microsoft Excel**.



TYPE	text	text	text	text	integer	float	float	text	text	text	integer	integer
FEPARAM	Protocol_N	Protocol_di	Scan_Date	Scan_Scar	Scan_Num	Scan_Micr	Scan_Micr	Scan_Origi	Grid_Name	Grid_Date	Grid_Num	Grid_Num
DATA	SD75 (Edit	########	########	Agilent Tec	2	10	10	b2a612cc-(	012064_D_	########	1	1
*												

TYPE	float	float	float	integer	float	float	float	integer	integer	float	float	float
STATS	gDarkOffs	gDarkOffs	gDarkOffs	gDarkOffs	rDarkOffse	rDarkOffse	rDarkOffse	gNumSat	FigLocalBG	rgLocalBG	rgLocalBG	rgLocalBGr
DATA	27.359	27	3.92805	1000	26.9	27	7.21722	1000	22	23.3265	50.6855	1.09454
*												

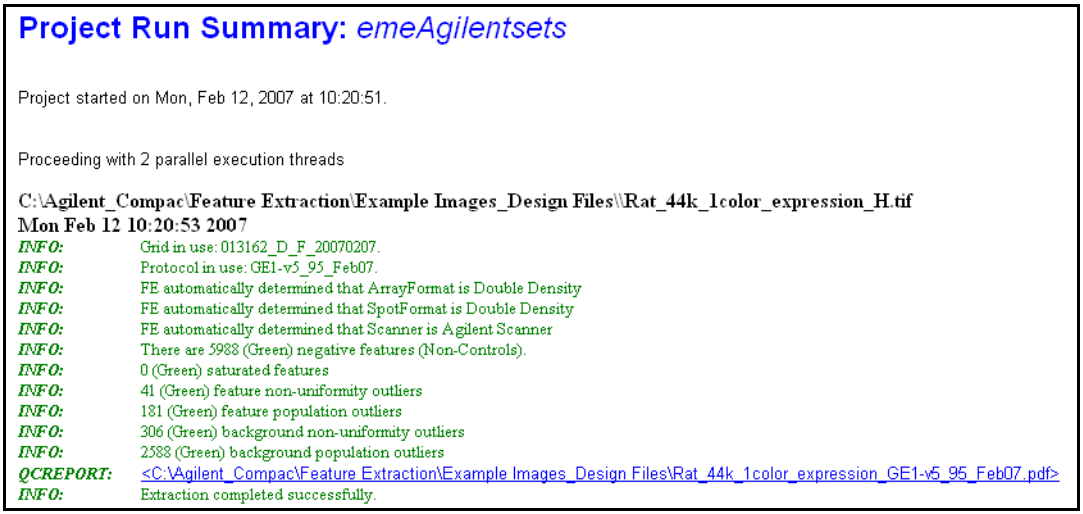
TYPE	integer	integer	integer	integer	text	integer	text	integer	integer	text	text	float
FEATURE	FeatureNu	Row	Col	SubTypeM	SubTypeNa	Start	Sequence	ProbeUID	ControlTyp	ProbeNam	GeneName	PositionX
DATA	1	1	1	260	BrightCorn	0	ATCATCG	0	1	Pro25G_T1	BrightCorn	3633.9
DATA	2	1	2	66	Structural	0	GCTAGCC	1	-1	(-)3xSLv1	3xSLv1	3849.51
DATA	3	1	3	0		0	GGTATAT	2	0	A_23_P14	I_931134	4057.66
DATA	4	1	4	0		0	CAAGCAC	3	0	A_23_P12	I_1000682	4268.3
DATA	5	1	5	0		0	TGTTTAT/	4	0	A_23_P28	I_947163	4473.22
DATA	6	1	6	1028	E1A	0	AGTGGTT	5	1	r60_n9	r60_n9	4693.21
DATA	7	1	7	517	ArraySynth	0	_____ATC	6	1	Pro25G_E	QC_Fusior	4906.88
DATA	8	1	8	0		0	AAGACCT	7	0	A_23_P23	I_929549	5115.48
DATA	9	1	9	0		0	ACTCCCT	8	0	A_23_P13	I_1100727	5335.27
DATA	10	1	10	0		0	TACCGCT	9	0	A_23_P50	NM_01865	5544.16
DATA	11	1	11	0		0	TTCTTCA	10	0	A_23_P27	I_959125	5748.63
DATA	12	1	12	0		0	GAGGGA	11	0	A_23_P32	NM_14526	5966.55
DATA	13	1	13	0		0	ATCTTGG	12	0	A_23_P25	I_957365	6173.77
DATA	14	1	14	517	ArraySynth	0	_____ATC	6	1	Pro25G_E	QC_Fusior	6386.74
DATA	15	1	15	0		0	TTGTGTA	13	0	A_23_P28	I_1971957	6600
DATA	16	1	16	0		0	ATAGATG	14	0	A_23_P96	I_953396	6809.88
DATA	17	1	17	0		0	GATGAGC	15	0	A_23_P23	I_932068	7022.18
DATA	18	1	18	0		0	TTGGGCT	16	0	A_23_P37	NM_00506	7234
DATA	19	1	19	0		0	ACTGCTG	17	0	A_23_P32	INCY:2026	7444.34
DATA	20	1	20	0		0	AAATTAG	18	0	A_23_P34	NM_14467	7659.36
DATA	21	1	21	517	ArraySynth	0	_____ATC	6	1	Pro25G_E	QC_Fusior	7873.65
DATA	22	1	22	0		0	GCATGAA	19	0	A_23_P50	NM_00704	8082.27
DATA	23	1	23	0		0	CTAGTTT	20	0	A_23_P20	I_931395	8292.25
DATA	24	1	24	0		0	AGTGATT	21	0	A_23_P33	NM_15337	8502.66
DATA	25	1	25	0		0	ACAGTGC	22	0	A_23_P12	I_957399	8711.52
DATA	26	1	26	0		0	ATGTCAC	23	0	A_23_P25	I_928271	8923.11
DATA	27	1	27	0		0	TCAACCA	24	0	A_23_P39	NM_03142	9136.32

Figure 20 Portion of text results file in Microsoft Excel

Note that for 244k arrays, the results will not all fit in an Excel spreadsheet (only the first 65536 rows). Results can be viewed in GeneSpring, Chip Analytics or CGH Analytics.

View the Project Run Summary Report

1. Go to the directory containing the project whose report you want to view.
2. Double-click the file named **YourComputerName_ LastBatchReport.rtf** to view the last saved report for the project.



Project Run Summary: *emeAgilentsets*

Project started on Mon, Feb 12, 2007 at 10:20:51.

Proceeding with 2 parallel execution threads

C:\Agilent_Compac\Feature Extraction\Example Images_Design Files\\Rat_44k_1color_expression_H.tif
Mon Feb 12 10:20:53 2007

INFO:	Grid in use: 013162_D_F_20070207.
INFO:	Protocol in use: GE1-v5_95_Feb07.
INFO:	FE automatically determined that ArrayFormat is Double Density
INFO:	FE automatically determined that SpotFormat is Double Density
INFO:	FE automatically determined that Scanner is Agilent Scanner
INFO:	There are 5988 (Green) negative features (Non-Controls).
INFO:	0 (Green) saturated features
INFO:	41 (Green) feature non-uniformity outliers
INFO:	181 (Green) feature population outliers
INFO:	306 (Green) background non-uniformity outliers
INFO:	2588 (Green) background population outliers
QCREPORT:	<C:\Agilent_Compac\Feature Extraction\Example Images_Design Files\Rat_44k_1color_expression_GE1-v5_95_Feb07.pdf>
INFO:	Extraction completed successfully.

Figure 21 Project Run Summary

OR

Double-click the file named **FE Project N_timestamp.rtf** to view a summary report for a specific run of the project.

You can access 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.

- More than 50% of the features along any edge of the microarray are non-uniform or not Found.
- The average difference between the nominal grid location and the spot centroid is greater than 10 microns.
- The grid location does not fit within the image.

  In this case, the software stops the extraction and does not produce a grid file output.

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 97 for instructions for a successful grid placement and extraction.

If there is no barcode in the MAGE-ML output...

Absence of a barcode may be due to several reasons:

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

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 "Add or change the New Array Identifier" on page 211.

View the QC Report

Seven possible QC reports can be generated:

- 2-color GE QC report with or without spike-ins
- 1-color GE QC report with or without spike-ins
- CGH QC report with or without spike-ins
- MicroRNA (miRNA) report

You can also access the QC Report from the Project Run Summary. Double-click the blue link next to QC Report.

If the FE license is a full license, the QC Report is in .pdf format.

If the FE license is a demo license, the QC Report is in .html format.

The report generated depends on the protocol used.

1 Go to the directory specified for containing the results.

2 Double-click the QC Report .pdf or .html file.

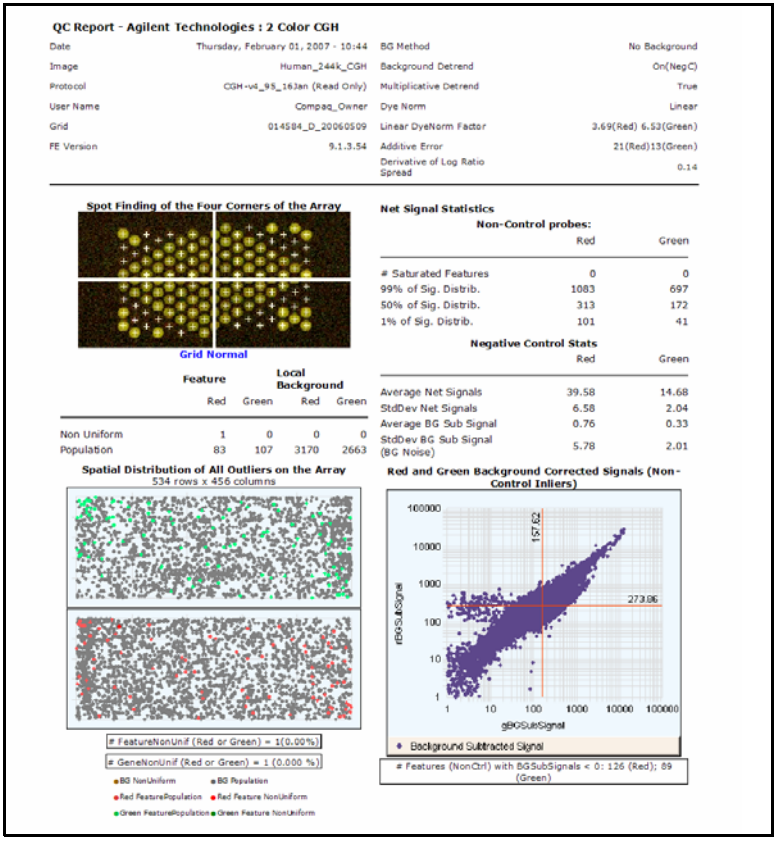


Figure 22 CGH QC Report—Page 1

Print the QC Report

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

1. Select **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. Select **File > Print Preview**.
7. Adjust the margins to display all the images.
8. Click **Print**.

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

View the QC Chart

If you generated a QC chart with the QC Chart tool, you can view this chart with this instruction.

- Select **Project > Show QC Chart**.

If you generated a QC chart by associating QC metrics with a project, you can view the chart with this instruction.

- Double-click the blue link next to **QC Chart** in the Project Run Summary report.

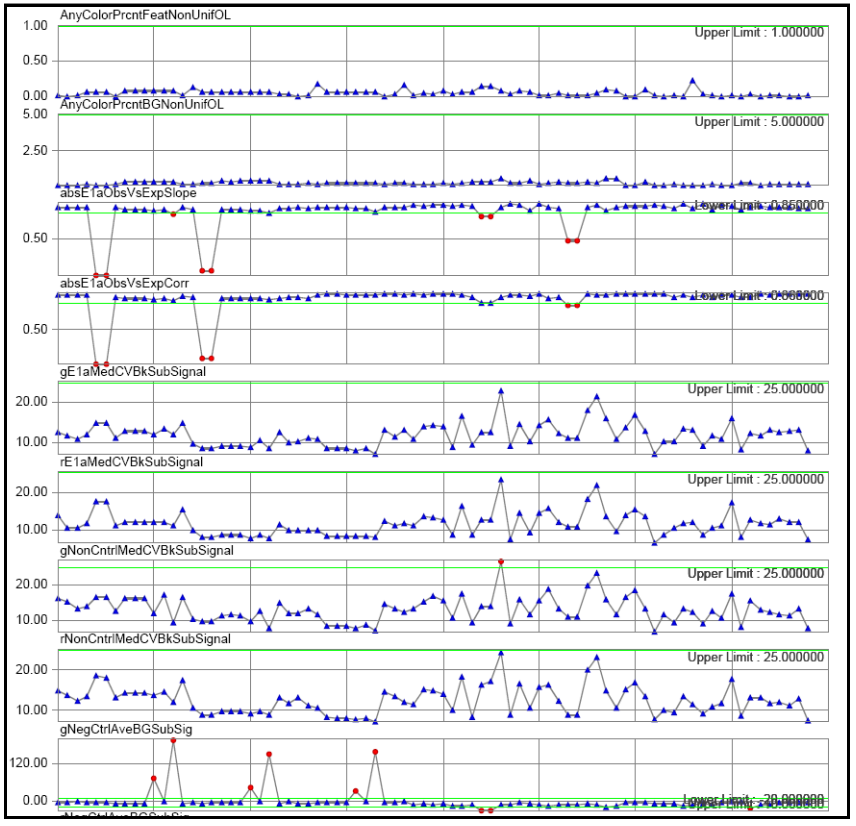


Figure 23 QC Chart

View visual results

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

1 Right-click a processed extraction set.

2 Select **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.

For multiplex Agilent arrays (more than 1 microarray on a single glass slide), a single .shp file is generated for the entire image and covers all of the arrays on the slide.

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

For more details on how to load the shape file from the Image workspace, see "Load Feature Extraction visual results" on page 212.

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

You can also open the image from Project Explorer or the Extraction Set Configuration tab sheet (double-click).

2 Select **Feature Extraction > Load Visual Result**.

3 Select the .shp file associated with the image, and click **Open**.

Work with the visual results

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

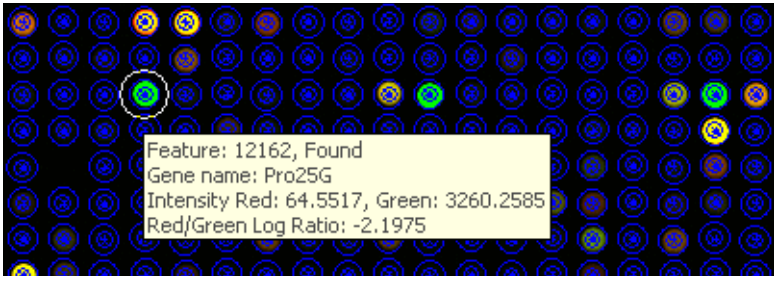


Figure 24 Visual results for extracted microarray image

3 Creating Grid Files and Templates

Feature Extraction incorporates automation with the use of grid templates, which contain general gridding information for a class of Agilent microarrays, not for specific microarrays.

If gridding fails for a particular Agilent microarray, you may have to create an Agilent grid file to use for the extraction. For non-Agilent microarrays in a project, you must create an Agilent grid file for the specific non-Agilent microarray you intend to extract.

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

When you create an Agilent grid file or redo a grid template, you locate the spot *centroids* 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.

This algorithm should locate most, if not all, of the spots where you want them. Any that are incorrect can be moved using the “Spot navigator” (square box). This ensures that the Feature Extraction software identifies the correct feature and local background.

Each section of this chapter describes the actions you can take to properly locate 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.

If automatic gridding for an Agilent image fails...

Follow these instructions to use a grid file if automatic gridding fails:

1. Create a new project with remaining extraction sets if run fails.
2. Access grid mode from a project extraction set.
3. Position grids, locate spots and save grid file.
4. Set up and rerun Feature Extraction.

You can also try rerunning Feature Extraction with a protocol that contains a different grid placement method. See "Use a different protocol for the extraction" on page 99.

Create a new project with remaining extraction sets if run fails

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. 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. FE 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.

Access grid mode from a project extraction set

This is the quickest way to access grid mode after automatic grid placement has failed.

1 In Project Explorer, right-click the grid template in the extraction set.

2 Select **Create and use grid file...**.

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

Also, the grid file replaces the original grid template in the extraction set.

Position grids, locate spots and save grid file

1 Position main grid, if necessary.

See "Positioning the main grid" on page 116.

2 Position subgrids.

See "Positioning the subgrids" on page 120.

3 Click the **Calculate spot size and centroids** icon.

See "Calculating spot size and spot centroids" on page 126.

Note that you cannot run this for multiplex (formerly known as multipack) microarrays.

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

See "Locating spot centroids manually" on page 127.

5 Save the grid file.

"Save the grid file" on page 135.

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.

Use a different protocol for the extraction

1. Double-click the protocol used for the extraction.
2. In Protocol Editor, click **Place Grid**.
3. In the cell next to **Placement Method**, double-click **Allow Some Distortion**.

   Place and Rotate Only appears in the cell.

   The placement method is inside the automatically determine section of the protocol. Pick the format of the array to uncover the placement method.
4. Click **Save As** to create a new protocol.
5. Enter the new **Protocol Name**, and click **OK**.
6. Create a new project with the same image file and new protocol.
7. Rerun Feature Extraction.

Creating and using a grid file for non-Agilent images

Follow these instructions to create and use a grid file for non-Agilent images or for any image where you prefer to use a grid file rather than a grid template. You must create a grid file for non-Agilent images to run Feature Extraction.

1 Access grid mode from the image file.

2 Position grids, locate spots and save grid file.

3 Extract the image file with its new grid file.

Access grid mode from the image file

1 Open the Agilent image file.

See "Open an image file" on page 102.

2 Set up the initial grid with the grid template used in the first extraction and the original protocol.

See "Set up the initial grid for an Agilent image" on page 104.

3 In the Geometry Information dialog box, click **Manual Fit** (since AutoFit has failed previously on this image).

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

Extract the image file with its new grid file

To test the new grid file or grid template, you can create the project and add the extraction set within grid mode.

1. Select **Feature Extraction > Create a New FE Project**.
2. Modify the project properties, and select to **Use Grid file if available** (**True**).
3. Add the original image file to the extraction set.
4. Add protocol used in the grid file to the extraction set.

   The grid file should have been automatically added.
5. Run Feature Extraction to see if the the grid works.

Setting up an initial grid

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 Work Area.

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

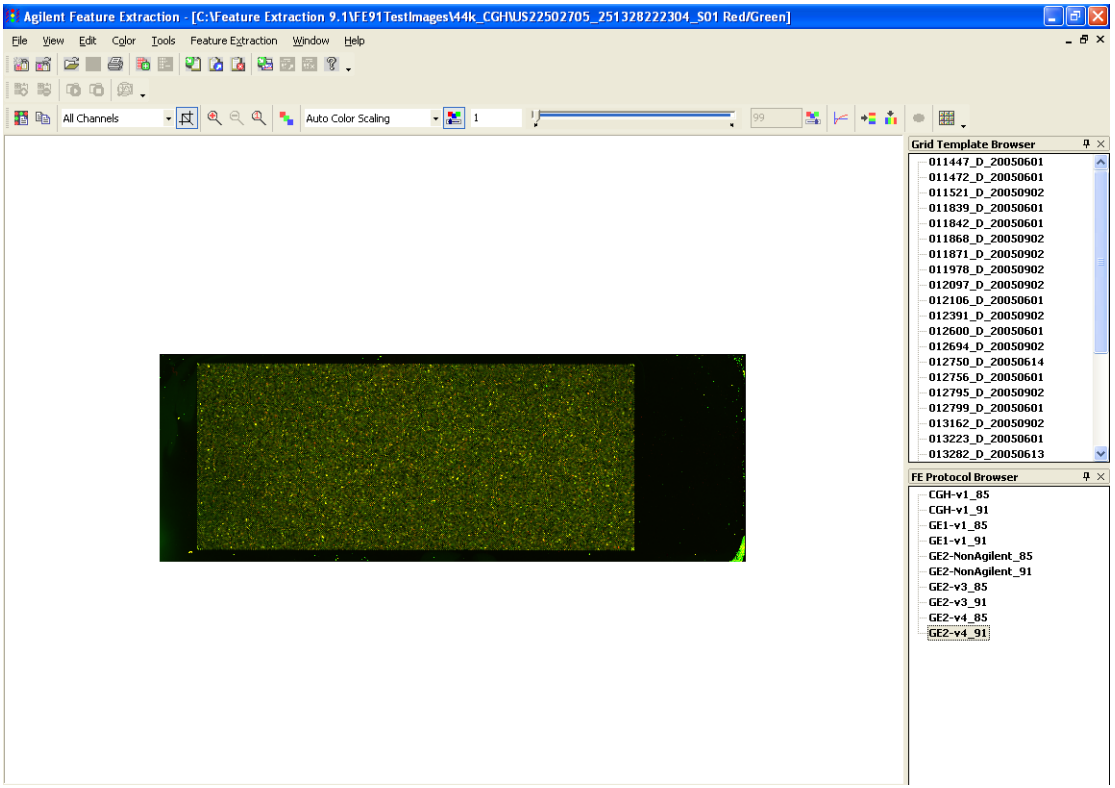


Figure 25 TIFF image for Agilent microarray

3 To see the spots more clearly, click the **Toggle Log Scale** icon and increase the image size.

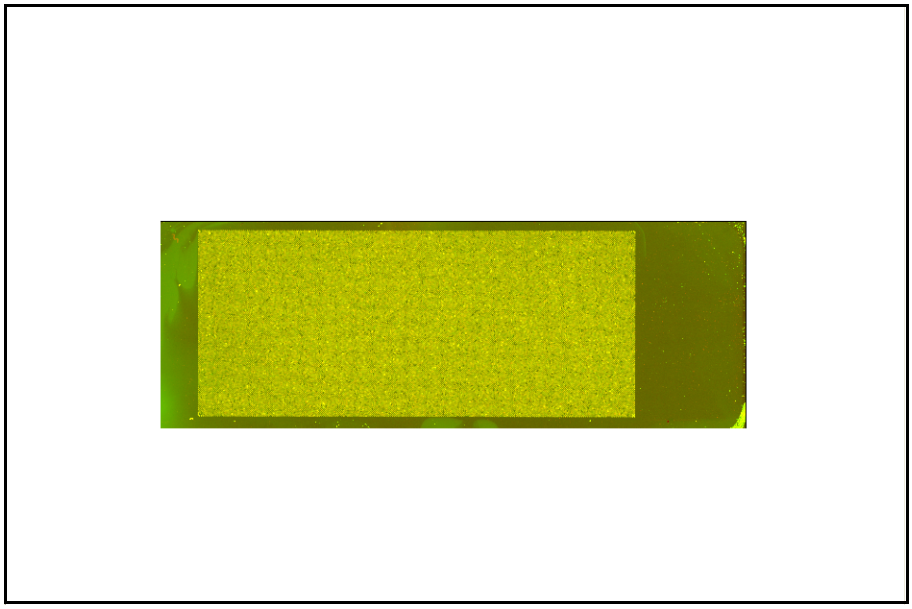


Figure 26 Image view with log scale on and image maximized

Change the Portrait/Landscape view of the image

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.

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.

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.

Set up the initial grid for an Agilent image

Select the gene list type for an Agilent image

1 Click the Grid button on the image toolbar. 

2 Select the gene list type to let you position the grid.

Typically, for an Agilent image you would select a Grid Template whose data orientation is landscape, as is that of the Agilent image.

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.

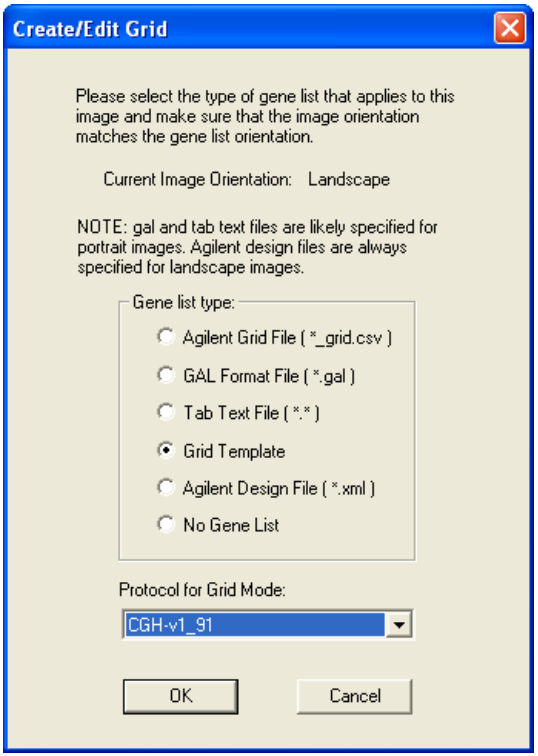


Figure 27 Create/Edit Grid dialog box

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.

Grid Template

This is the preferred selection. Select a grid template if you are attempting to create a grid replacement for an Agilent image whose grid was not placed properly on the image.

Agilent Design File (*.xml)

Select this option if you have only a design file available. Your image must be in a landscape orientation because the design information specifies a landscape orientation.

3 Select a protocol.

This dialog box gives you the opportunity to select a protocol to help with grid placement and spot finding. Grid placement and spot finding algorithms differ between protocols. 11k and 22k are the same, 44k and 44k_CGH are different, and non-Agilent is different. You must select a protocol that matches the image for grid mode. If not, the grid and spot placement may not be correct.

4 Click **OK**.

What happens next depends on the gene list type that you select.

- For Grid Template, see page 106.
- For Agilent Grid File, see page 108.
- For Agilent Design File, see

If you select Grid Template as the gene list type

The Select Grid Template list appears.

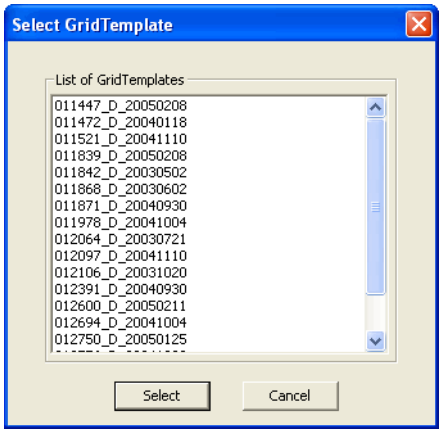


Figure 28 Select Grid Template list

1 Select the grid template that you want to change to better fit the Agilent image you have opened.

2 Click **Select**.

The Geometry Information dialog box appears.

3 Change the default values if you need to.

4 Click **Manual Fit**.

The main grid view appears with the grid out of position. See next page.

If you click Auto Fit, the software gives you the same grid placement that was used in the extraction that produced an error, that is, if you are also using the same protocol. If you change the Placement Method in the protocol, AutoFit may work.

See "Geometry information that you can change" on page 109".

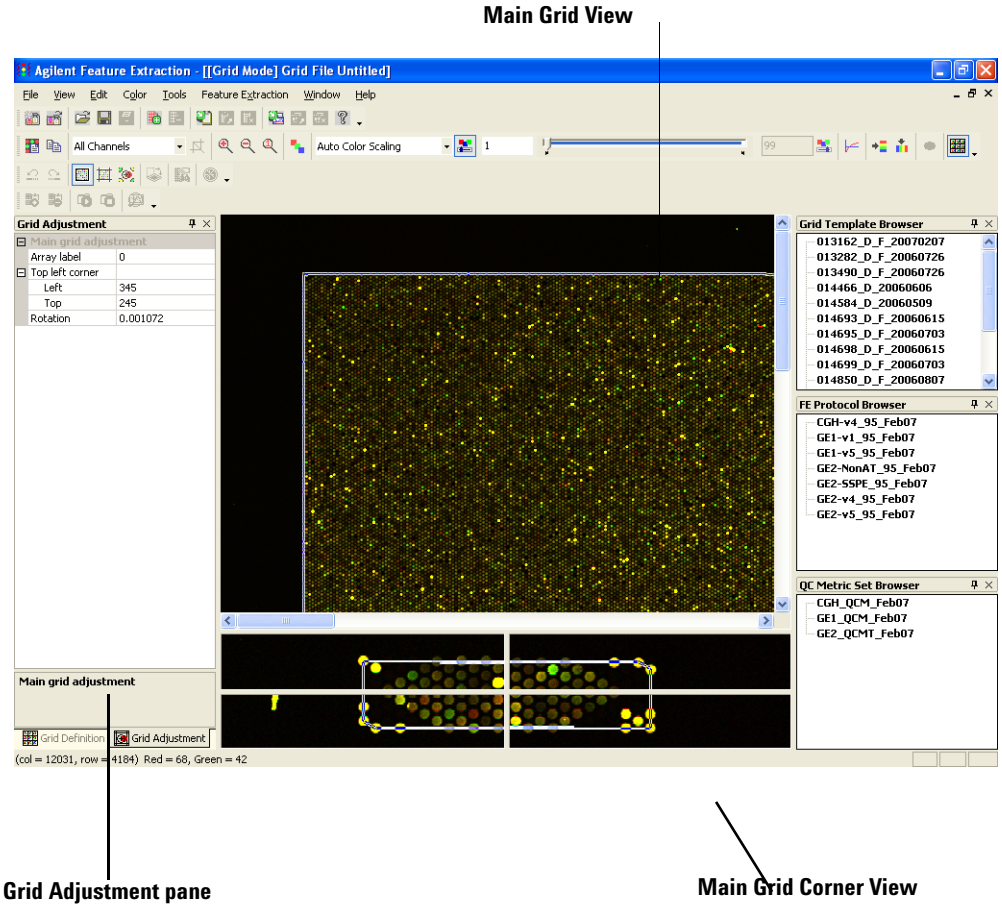


Figure 29 Main Grid View for Agilent image with Agilent grid template

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 an Agilent Design File as the gene list type

A Browser appears from which to select the design file.

1. Select the design file whose values you want to change to better fit the Agilent image you have opened.
2. Click **Open**.

   The Geometry Information dialog box appears.
3. Change the default values if you need to.
4. Click **Manual Fit**.

   The main grid view appears with the grid out of position. See page 107.

   If you click Auto Fit, the software gives you the same grid placement that was used in the extraction that produced an error, that is, if you are also using the same protocol. If you change the Placement Method in the protocol, AutoFit may work.

See "Geometry information that you can change" on page 109".

Geometry information that you can change

The geometry information that you can change depends on your selection of gene list type.

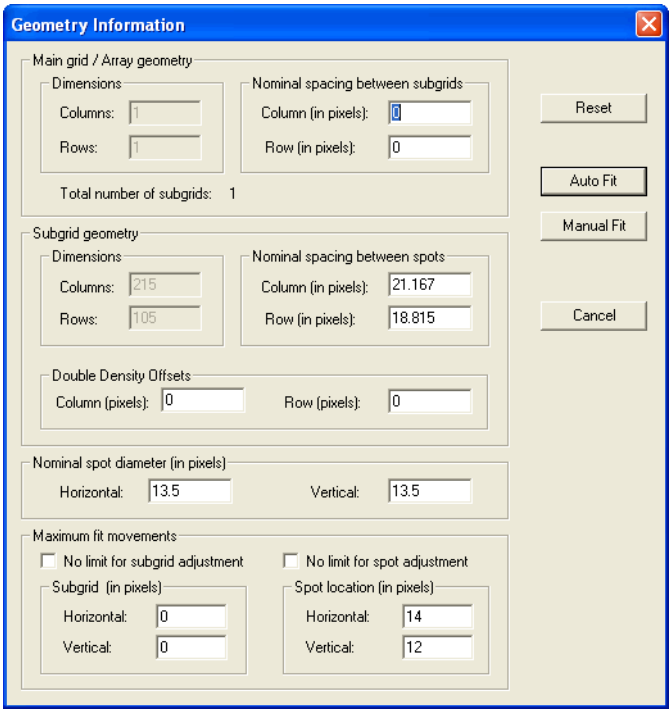


Figure 30 Geometry Information dialog box

If you are going to manually fit an Agilent image that has failed automatic grid placement, you do not need to change any of these settings.

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.

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

Maximum fit movements The software restricts your ability to move subgrids and spots on the grid because in most cases, you just need a fine adjustment. For some images you may need to move a grid or spot centroid over a large distance. You can change the default maximum movements or elect to have no limits on the movement.

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.

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 "one iteration" grid, the software identifies the spot centroids.

Table 1 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
Grid Template	Yes	No	No	Yes	Yes
Agilent Design	Yes	No	No	Yes	Yes
No gene list	Yes	Yes	Yes	Yes	Yes

* You can always change the nominal spacing between subgrids and the maximum fit movements.

† 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.

Set up the initial grid for a non-Agilent image

1 Click the **Grid** button on the image toolbar. 

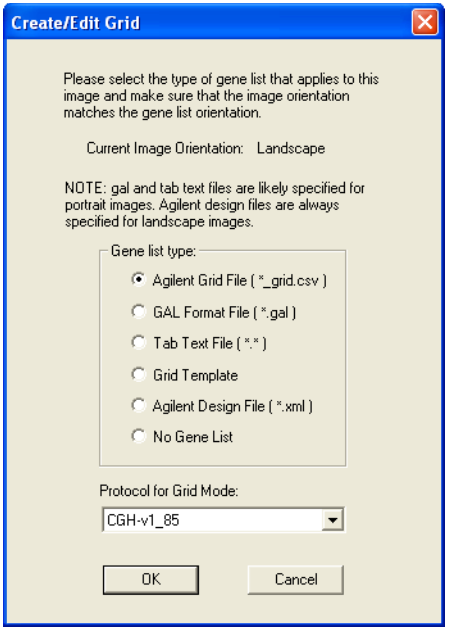


Figure 31 Create/Edit Grid dialog box

2 Select the gene list type to let you position the grid.

For a non-Agilent image you typically select a gal or tab text file if such a file is associated with the image, or you may have to create a grid with no gene information.

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.

Gal Format File (*.gal) Select this file if the image that you opened has an associated GAL file.

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.

For these types of gene list files, you must use a flexible protocol that allows some distortion because there is no nominal grid spacing to rigidly place.

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 "XML Control Type output" on page 181 of the *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.

3 Select a protocol.

4 Click **OK**.

If you 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 that you can change" on page 109.

The Main Grid View for this image appears.

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

The Tab Text Grid File Preview window appears.

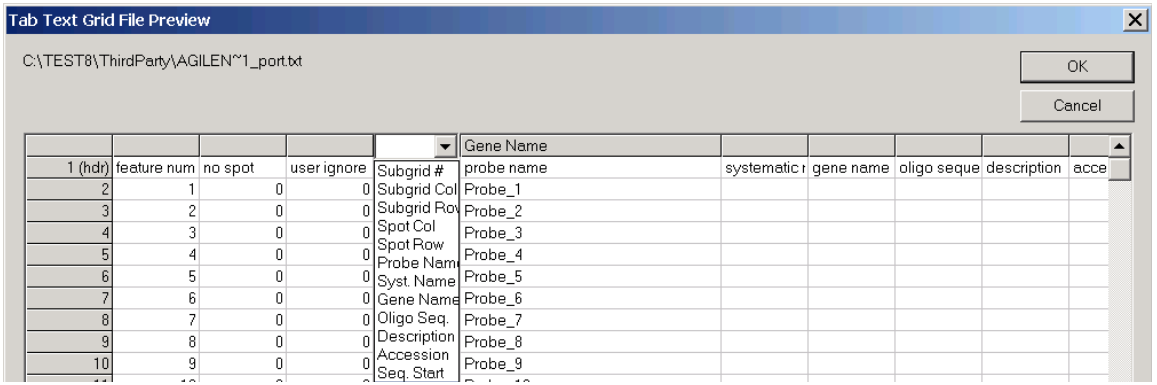
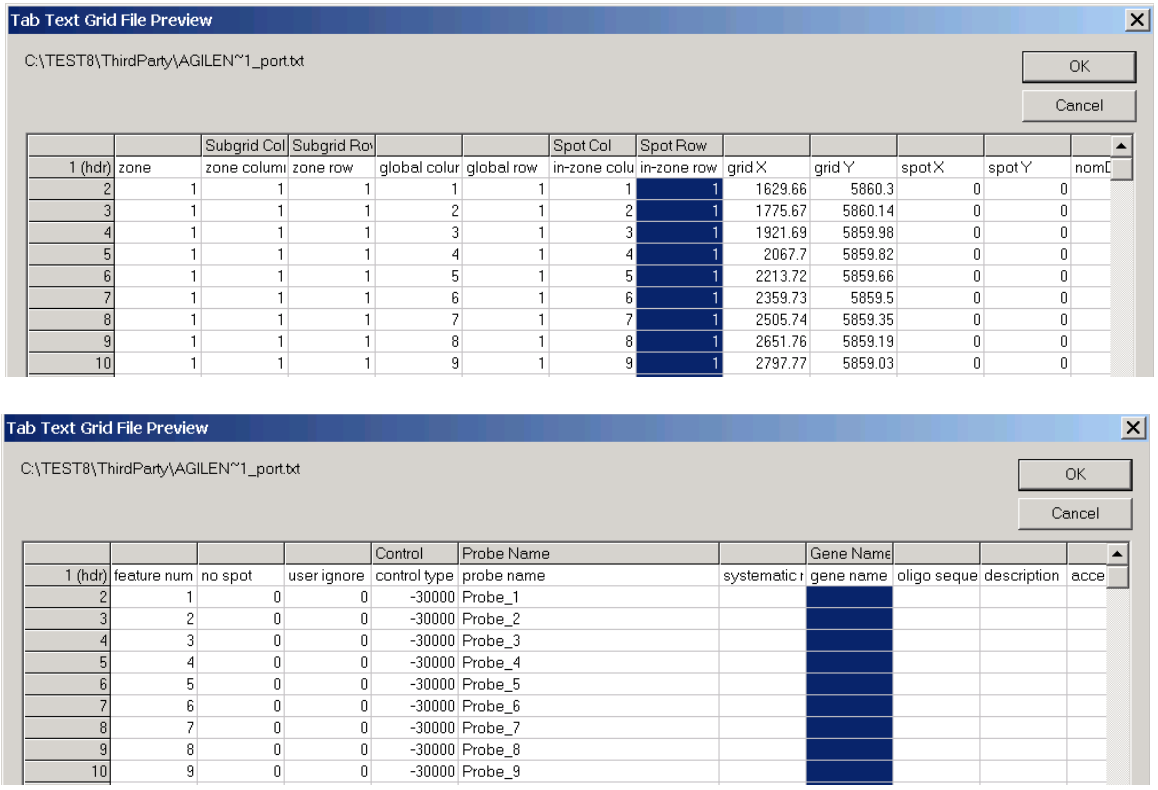


Figure 32 Tab Text Grid File Preview window with header selected

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.



Tab Text Grid File Preview

C:\TEST8\ThirdParty\AGILEN~1_port.txt

OK

Cancel

		Subgrid Col	Subgrid Rov			Spot Col	Spot Row					
1 (hdr)	zone	zone columi	zone row	global colur	global row	in-zone colu	in-zone row	grid X	grid Y	spot X	spot Y	nomL
2	1	1	1	1	1	1	1	1629.66	5860.3	0	0	
3	1	1	1	2	1	2	1	1775.67	5860.14	0	0	
4	1	1	1	3	1	3	1	1921.69	5859.98	0	0	
5	1	1	1	4	1	4	1	2067.7	5859.82	0	0	
6	1	1	1	5	1	5	1	2213.72	5859.66	0	0	
7	1	1	1	6	1	6	1	2359.73	5859.5	0	0	
8	1	1	1	7	1	7	1	2505.74	5859.35	0	0	
9	1	1	1	8	1	8	1	2651.76	5859.19	0	0	
10	1	1	1	9	1	9	1	2797.77	5859.03	0	0	

Tab Text Grid File Preview

C:\TEST8\ThirdParty\AGILEN~1_port.txt

OK

Cancel

				Control	Probe Name		Gene Name			
1 (hdr)	feature num	no spot	user ignore	control type	probe name	systematic r	gene name	oligo seque	description	acce
2	1	0	0	-30000	Probe_1					
3	2	0	0	-30000	Probe_2					
4	3	0	0	-30000	Probe_3					
5	4	0	0	-30000	Probe_4					
6	5	0	0	-30000	Probe_5					
7	6	0	0	-30000	Probe_6					
8	7	0	0	-30000	Probe_7					
9	8	0	0	-30000	Probe_8					
10	9	0	0	-30000	Probe_9					

Figure 33 Tab Text Grid File Preview window with headers selected

3 Click **OK**.

The Geometry Information dialog box appears.

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

For a description of all the fields and buttons in the Geometry Information dialog box, see "Geometry information that you can change" on page 109.

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

If you select No Gene List

The Geometry Information dialog box appears.

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 that you can change" on page 109.

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

Positioning the main grid

Zoom in or out of the Views

See "Change the magnification (zooming)" on page 224 of this guide for more details on zooming.

- Select **View > Zoom** to choose a % zoom from 10% to 1000%. A 100% zoom is defined and you can select a zoom above or below this in 10% increments. You can also change the default zoom settings.

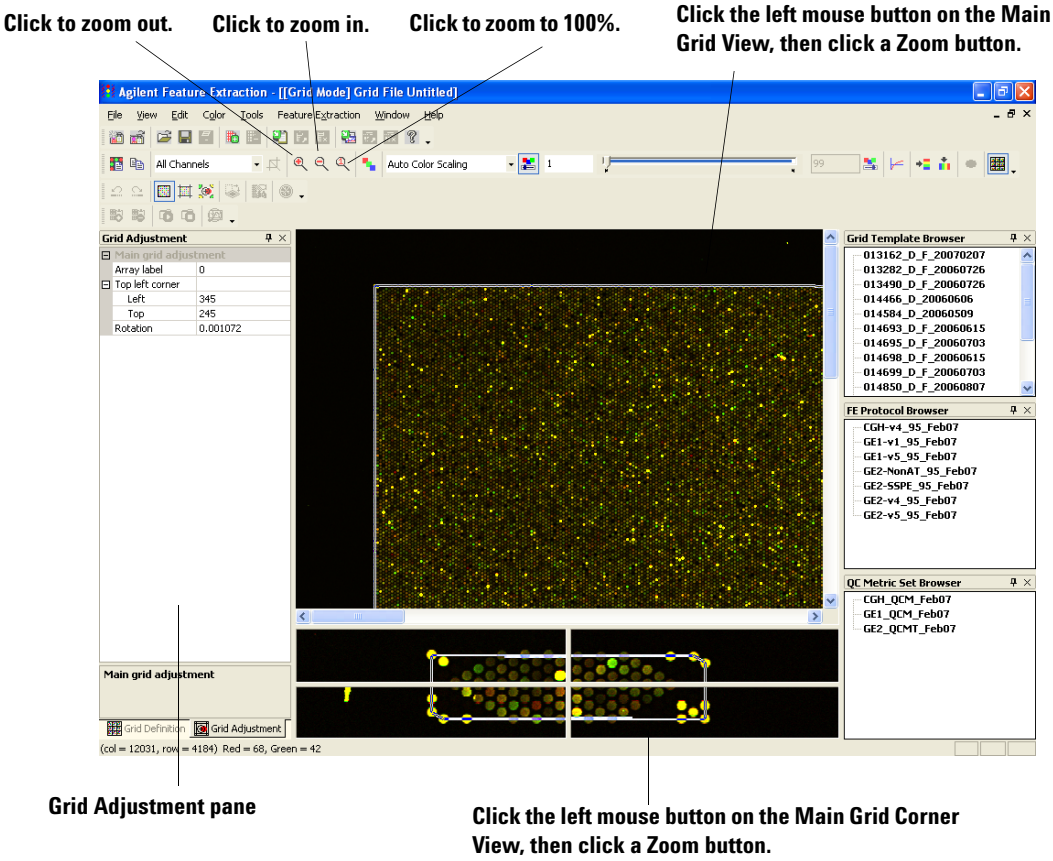
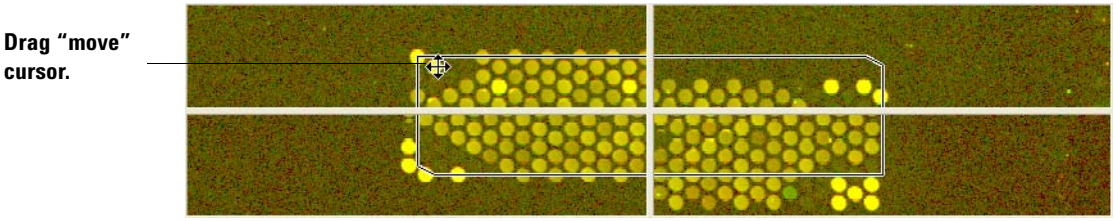


Figure 34 Main Grid View with Grid Adjustment panel for 244k array

Move the position of the main grid

Note that the thick line represents the main grid and the thin line represents the subgrids.

1. Make sure that the color log scale has been turned on to view the corner spots on the main grid more easily.
2. Move the mouse along the Main Grid Corner View until the "move" cursor appears.
3. Drag the "move" cursor to position the grid on the four corners of the microarray image.

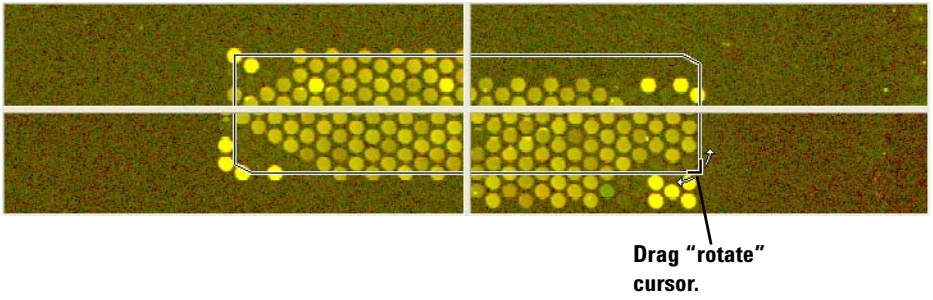


Rotate the position of the main grid

To move or rotate the main grid another way, see "Move or rotate the main grid with Grid Adjustment pane" on page 118.

1. Make sure that the color log scale has been turned on to view the corner spots on the main grid more easily.
2. Move the mouse to a corner of the Main Grid Corner View.

   The "rotate" cursor appears.
3. Drag the "rotate" cursor to rotate and optimize the position of the grid on the four corners of the microarray image.



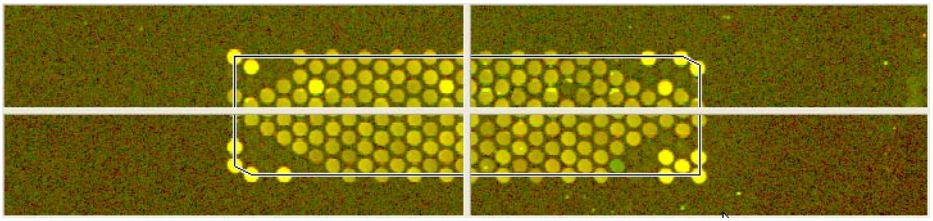


Figure 35 Final position of main grid on 244k microarray

Undo or redo main grid position movements

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

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 139.

1 Click the **Grid Adjustment** tab, if necessary.

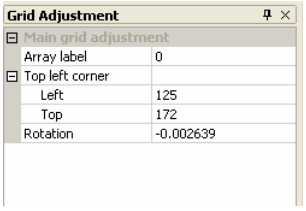


Figure 36 Grid Adjustment pane for Main Grid View

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

Array Label This number identifies which of the two 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.

Positioning the subgrids

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.

You can also manually adjust the nominal spot centroids for individual spots. See the next section.

Prepare to adjust the subgrids

1. Make sure that you turn the color log scale on if you have not already done this.
2. Click the **Adjust Subgrid** button on the toolbar to bring up the Subgrid View and Subgrid Corner View.
3. Click anywhere on the Subgrid View to select a subgrid to appear in the Subgrid Corner View.

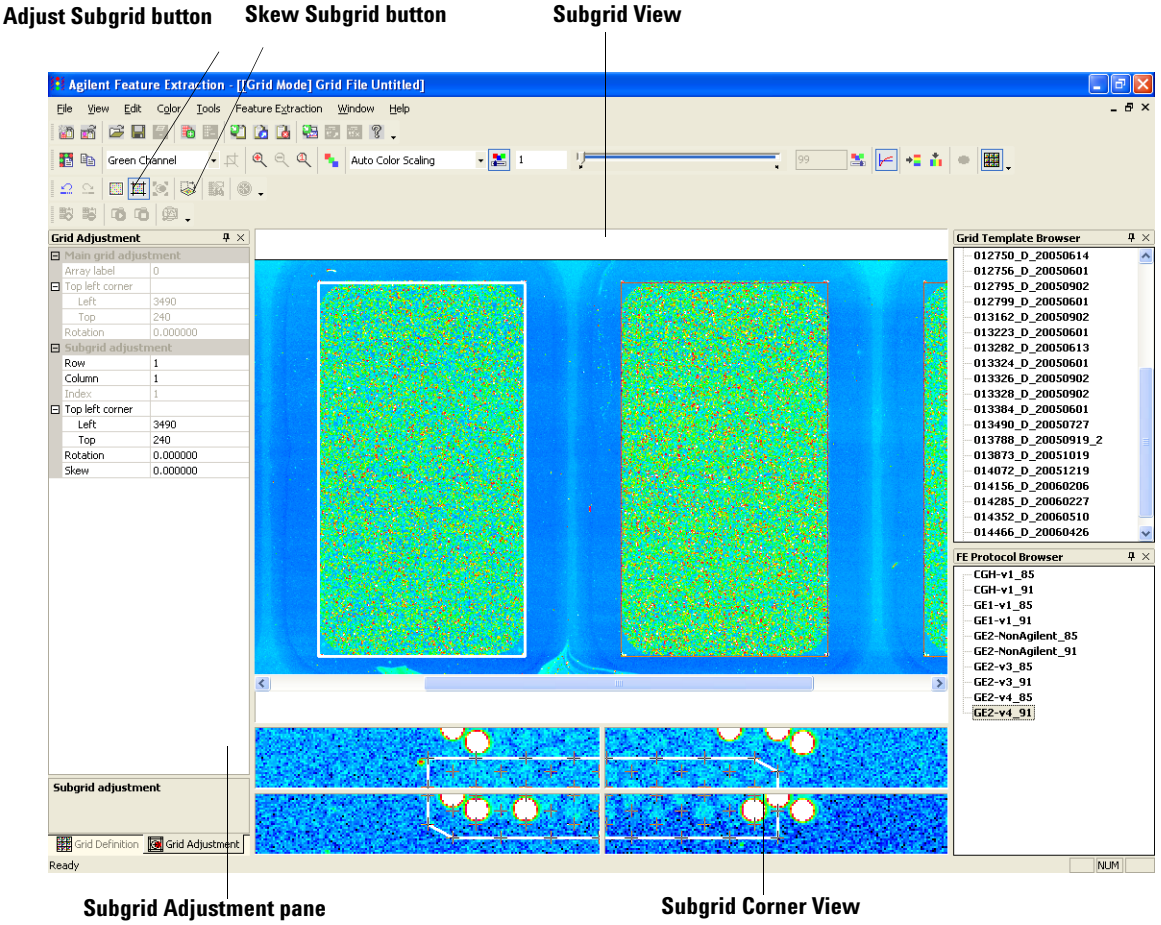
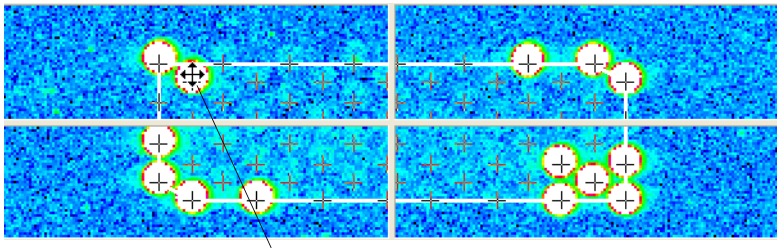


Figure 37 Subgrid View with Grid Adjustment pane for a 1-color 4X44k multiplex (formerly known as multipack) microarray

Move the position of the subgrid

1 Move the mouse along the Subgrid Corner View until the "move" cursor appears.

2 Drag the "move" cursor on the Subgrid Corner View to position the subgrid on the four corners of that portion of the microarray image. (Shift + Arrow Key also moves a subgrid.)



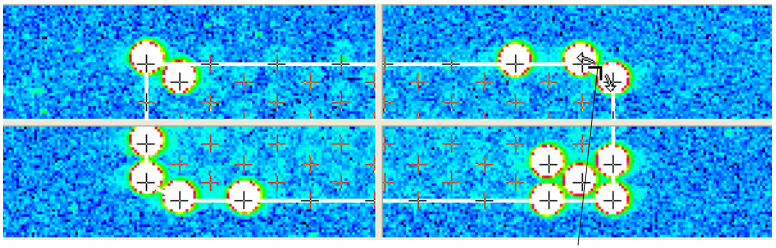
Rotate the subgrid

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

1 Move the mouse to a corner of the Subgrid Corner View.

The "rotate" cursor appears.

2 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.



Skew the subgrid corner

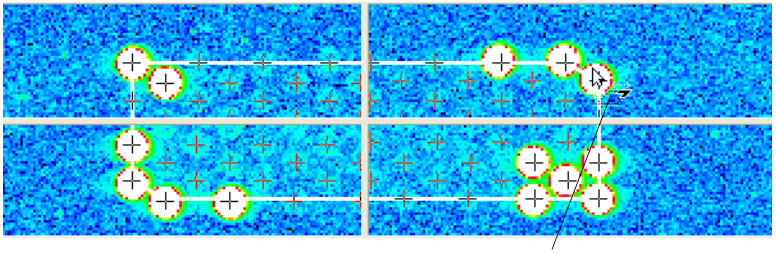
If neither of the above methods work, try skewing the subgrid corner.

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.



Undo or redo subgrid position movements

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

Position subgrids with Grid Adjustment pane

When the Subgrid View appears, 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 139.

1 Click the **Grid Adjustment** tab, if necessary.

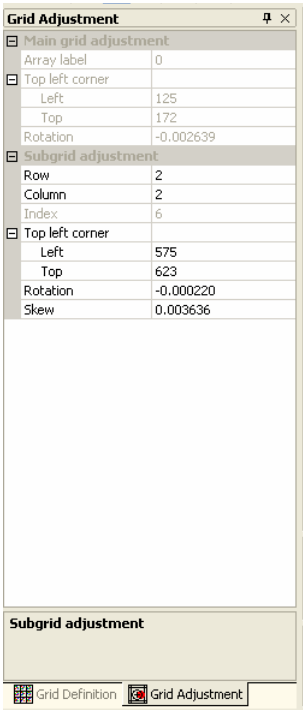


Figure 38 Grid Adjustment pane for subgrids

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.

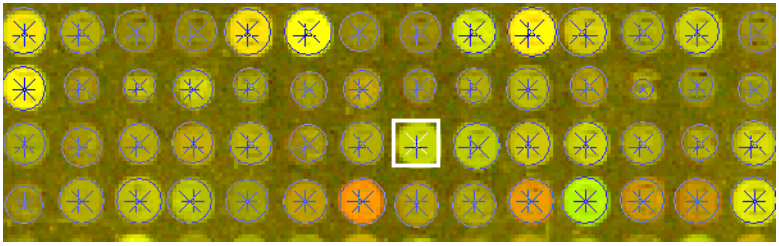
Rotation Enter a number to rotate the top left corner of the subgrid.

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

Calculating spot size and spot centroids

This is the same algorithm used for calculating spot size and finding spot centroids by the software during an automatic extraction. If a grid file is used in an extraction set, this calculation must saved as part of the file.

- 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).



CAUTION You do not have access to this function for multiplex (formerly known as multipack) microarrays unless you crop out a single pack.

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.

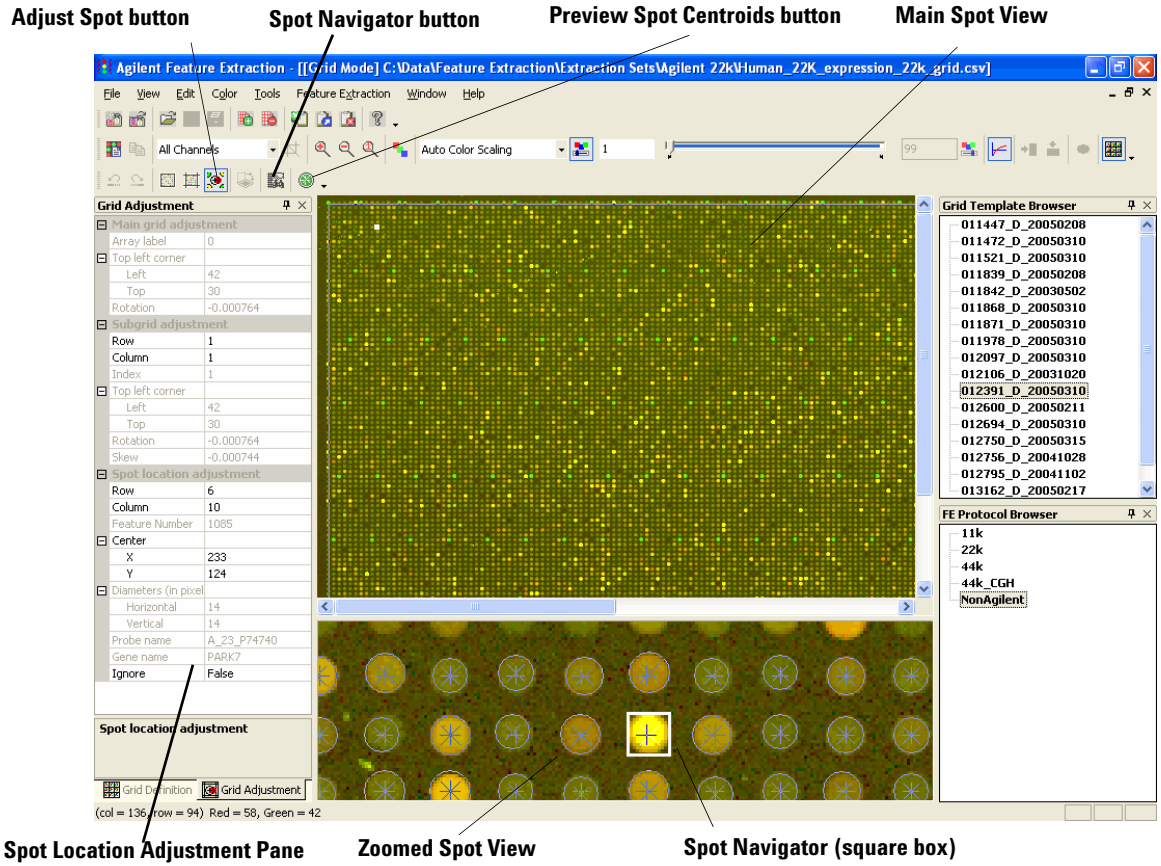
Locating spot centroids manually

If a spot centroid has been manually moved, the result called IsManualFlag is set to true. See Table 23, "Feature results contained in the FULL output text file (FULL FEATURES table)," on page 147 of the Reference Guide for the result.

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.

Find a spot to adjust

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



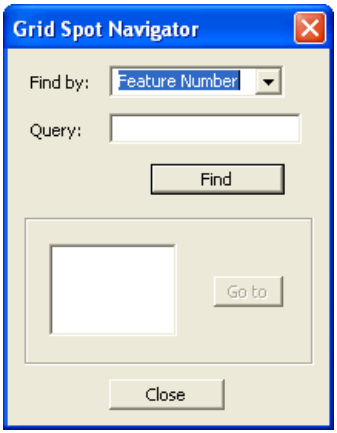
The calculated spot centroids appear when you enter Adjust Spot mode.

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.



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 displays in the bottom box.

7 Select a feature from the list, and click **Go to**.

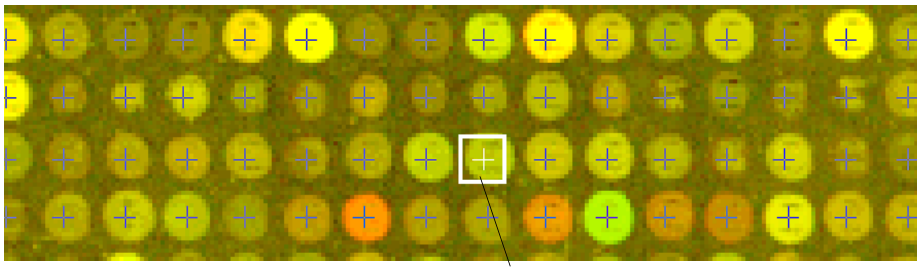
The software moves the Spot Navigator to that feature.

8 Click **Close**.

The software positions the Spot Navigator (square box) in the center of the Zoomed Spot View.

Position the spot centroid

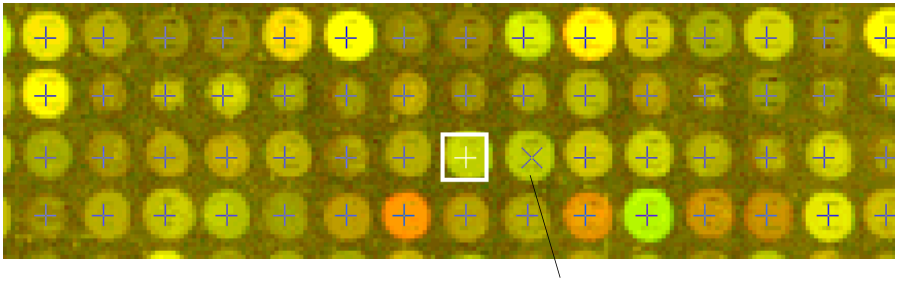
Before you position the Spot Navigator to find the centroid, the position of the nominal spot centroid could look like this:



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:



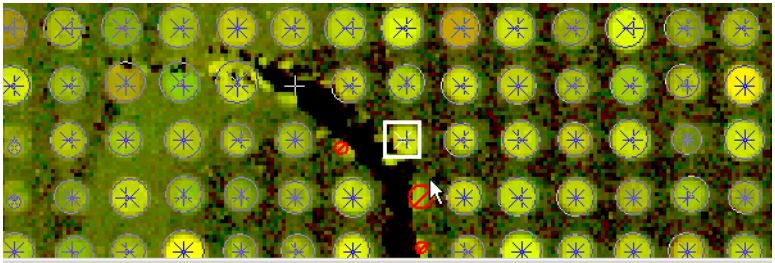
New estimated spot centroid (X) position;

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

The software will not use this moved estimated spot centroid in the spotfinder calculation, but it will use it to define a feature and local background.

Ignore spots

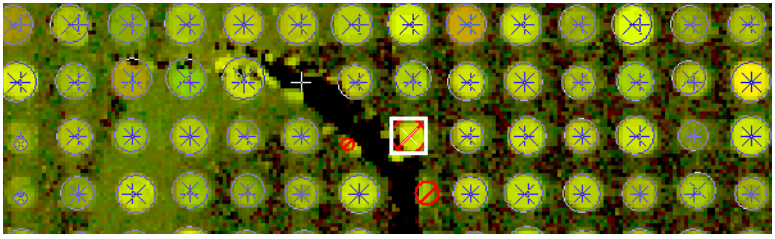
1 Position the **Spot Navigator** where you want to ignore a spot.



2 Mark the **Ignore spot** checkbox.

3 Press **Enter**.

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



Undo or redo spot position movements

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

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.

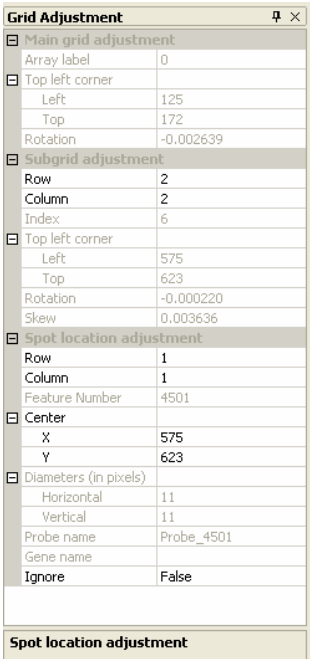


Figure 39 Spot Location Adjustment pane

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 139.

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

3 Press **Enter**.

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 spot	Mark this check box for the selected spot, and click Apply to ignore the selected spot.

Finishing grid setup

Refit a grid

You may want to refit a grid under the following circumstances mentioned in the table below. The solution to the problem is also described:

Table 2 Conditions for refitting grids

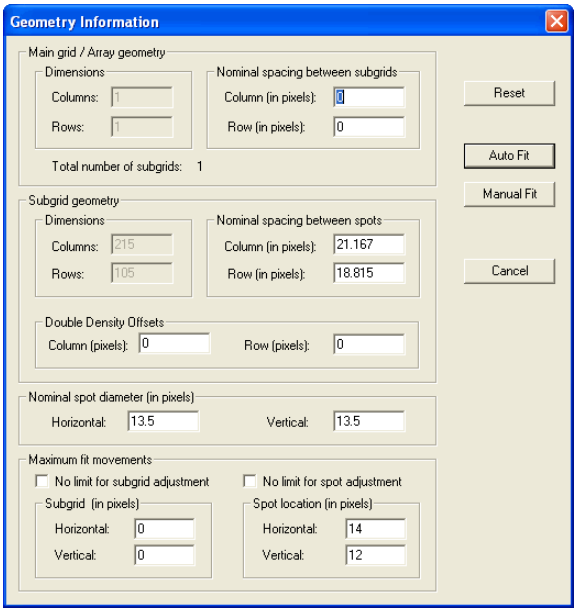
Grid Issue	Solution
Main grid is not properly placed.	Position the main grid manually, then refit the grid with autofit to fit the subgrids.
Subgrids are not correctly spaced.	Refit grid to adjust the spacing between subgrids followed by a manual fit. Assumption is that the subgrids are evenly spaced.
Spot spacing appears off.	Refit grid to adjust the spacing between the spots.
Average spot size is not correct.	Refit grid to adjust the spot diameter. This diameter is used by the Calculate Spot Size and Centroids algorithm and the WholeSpot algorithm.

To refit a grid, follow these steps.

To learn about these fields, see "Geometry information that you can change" on page 109.

1 Select **Edit > Refit Grid**.

The Geometry Information dialog box appears.



Click Reset to re-enter the original number before you changed it.

Figure 40 Geometry Information dialog box to refit Agilent grid

2 Change entries, and click **Auto Fit** or **Manual Fit**.

Save the grid file

1 Click the **Save Grid** button on the image toolbar.

2 Name the grid, then click **Save**.

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.

Save a grid file as a grid template

You can do this only by adding the grid file as a grid template to the database. If you have a set of non-Agilent microarrays that have the same gene list and basic geometry, you can create a grid file for one, add this grid file as a grid template to the database and then run extractions on the other non-Agilent microarray images with this grid template.

1 Select **Tools > Grid Templates > Add**.

2 Select the .csv file to add to the database, and click **Add**.

The grid file is now a grid template and is listed in the Grid Template Browser on the left of the screen.

CAUTION A grid file contains absolute locations for the spots. A grid template contains a grid that is aligned to the data during extraction. When a grid file is saved to the database as a grid template, the new template contains no absolute locations. Therefore, when you run Feature Extraction with a grid template derived from a grid file, the extracted results may differ from those run with the original grid file.

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.

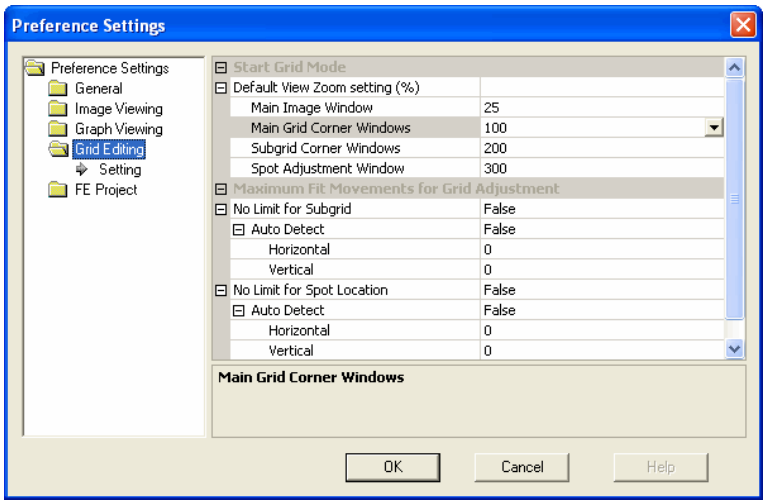


Figure 41 Preference Settings—Grid Editing selected

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 above.

Maximum Fit Movements

These are the maximum distances that you can move a grid or spot centroid. The software restricts your ability to move subgrids and spots on the grid because in most cases, you just need a fine adjustment. For some images you may need to move a grid or spot centroid over a large distance.

- You can elect to have no limit on the movements.
- If Auto detect is off, you can change the default maximum movements.

- If Auto detect is selected, the software determines the default values, which you can change in the Geometry Information dialog box when you create or refit a grid.

  You can also change these values in the Grid Definition panel for any grid or spot mode.

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 3. In addition, there are a number of Ctrl and Alt key operations that help you work with grids. These are presented in Table 4.

Table 3 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 or corner windows.
	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 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 or corner windows.
	Rotate subgrid	Edit value, then press Enter.	NA	Drag on corner in corner window.
	Skew subgrid	Edit value, then press Enter.	NA	Click Skew Subgrid button, then drag on corner in 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** checkbox, 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 4 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

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

See Chapter 1, “Default Protocol Settings” in the *Reference Guide* for a listing of the default settings and parameter values for the protocols shipped with the software.

Changing protocol settings — an introduction

Protocol templates

Agilent ships protocol templates for running Feature Extraction with various types of scan data. The name of each protocol template follows the format:

Agilent microarray type-lab protocol version_software version_date

The table below lists the names of the templates that Agilent ships, along with their descriptions:

Table 5 Unremovable Protocol Templates shipped with software

Protocol Template Name	Description
CGH-v4_95_Feb07	CGH microarray with values for version 4 of lab protocol and version 9.5 of software
GE1-v1_95_Feb07	1-color gene expression microarray with values for first version of lab protocol and version 9.5 of software
GE1-v5_95_Feb07	1-color gene expression microarray with values for version 5 of lab protocol and version 9.5 of software
GE2-SSPE__95_Feb07	2-color gene expression microarray with values for SSPE version of lab protocol and version 9.5 of software
GE2-v4_95_Feb07	2-color gene expression microarray with values for version 4 of lab protocol and version 9.5 of software

Table 5 Unremovable Protocol Templates shipped with software

Protocol Template Name	Description
GE2-v5_95_Feb07	2-color gene expression microarray with values for version 5 of lab protocol and version 9.5 of software
GE2-nonAT_95_Feb07	2-color gene expression microarray with values for a non-Agilent image and protocol and version 9.5 of software

NOTE

Agilent miRNA microarrays are currently in development. Please check the Agilent website for the latest information.

To download the miRNA protocol template when available or updates for other templates, go to http://www.agilent.com/chem/feprotocols and follow the instructions for importing the templates into the Protocol Browser. See "Import or export a protocol to or from the database" on page 82.

New protocols from protocol templates

See Chapter 5, "How Algorithms Calculate Results" in the 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. Open a protocol.
2. View or change the protocol properties.
3. Change the protocol steps to run.
4. Change the settings for each step.
5. Save a protocol to a new name.

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

Create a new protocol from a template or existing protocol

To create a protocol for a specific class of microarray and format, you ***must*** use a protocol template or existing protocol for the same class or format of microarray.

For example, if you want to create a 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 protocol is by looking at the Derived From field of the general section of the protocol. ("Change the protocol steps to run" on page 147)

CAUTION

Protocol templates provide both visible and **hidden** settings whose values are specific to the class or format of microarrays. Although you can change the visible settings so that any two protocols of different class or format *appear* identical, you **cannot change the hidden settings** that distinguish these protocols from one another.

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.

The first screen that appears, no matter the protocol opened, is the default screen for the step that was used just before a protocol was closed or saved.

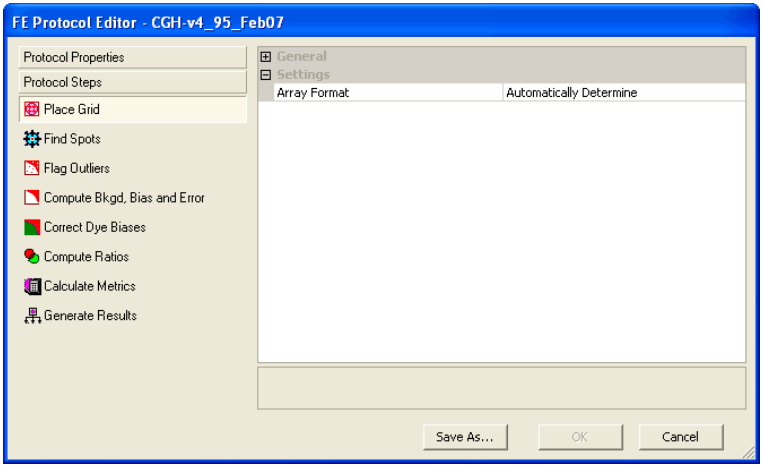


Figure 42 FE Protocol Editor

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.

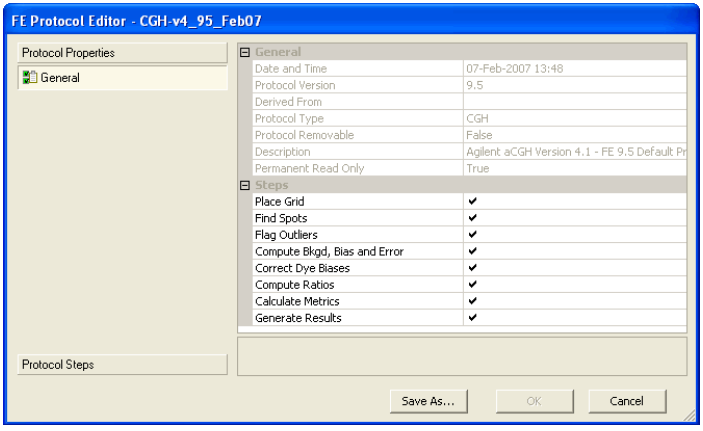


Figure 43 Protocol Properties page of the Protocol Editor

Date and Time Timestamp for the creation of the protocol

Protocol Version Version of the protocol template, corresponding to software version with which it was shipped

Derived From Name of original Agilent protocol template from which the current protocol is derived; blank until you exit and reopen modified protocol

Protocol Type Corresponds to type of scan data for which the protocol is to be used for Feature Extraction: 2-color gene expression, 1-color gene expression, CGH or miRNA data

Protocol Removable Describes if the protocol is removable from the database with user-created protocols removable and Agilent protocol templates not removable

Description Description of the type of microarray and format for which the protocol was designed

Permanent Read Only Select True if you do not want the protocol to be changeable and False if you do.

NOTE Agilent miRNA microarrays are currently in development. Please check the Agilent website for the latest information.

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 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, or Feature Extraction algorithms, on the right.
2. Mark each checkbox to enable a step that you want to run, or Clear each checkbox to disable a step that you do not want to run.

Refer to Figure 43 on the previous page to see the FE Protocol Editor.

Protocol steps for GE 2-color and CGH images

Below is a description of each of the protocol steps for GE 2-color and CGH images, as seen in Figure 43, in the order that the software runs them.

Place Grid

- A grid is placed on the microarray image.

  This step or algorithm finds the grid to define the nominal positions of the spots on the microarray.

Find Spots

- The spot centers are located and/or optimized starting with 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.
- Pixels of too low or high intensity are then removed from the defined features and background.

- Statistics, such as the mean, median and standard deviation (SD), are calculated on the remaining inlier pixels within the feature and background.
- Features with saturated pixels are also flagged.
- The protocol decides whether the mean or median signals are used for further analysis.

Flag Outliers

- Features and background regions are flagged if their pixel distributions are non-uniform.
- Also, population statistics are applied to local background regions and to features that are replicates of the same probe sequence to produce distributions of signal intensity.
- This analysis flags features or background regions with signals that are less than or greater than thresholds set from those distributions.

Compute Bkgd, Bias, and Error

- Background signals are subtracted from raw signals using the area either on or around the features.
- The background-subtracted signals can be corrected for systematic gradients, especially low signals, using spatial detrending.
- The error is calculated at this point, using the choice between the Universal error model or the larger of the error estimates from either the Universal error model or the propagated error.
- 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”.
- A multiplicative detrending algorithm is also used to account for gradients proportional to the signal.

Correct Dye Biases

- Differences in the red and green channel signals caused by different efficiencies in labeling, emission and detection are estimated and corrected.

Compute ratios

- Log ratios of the red and green channels are calculated, as well as a log ratio error and a p-value for each feature. These metrics 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 3 types of text results or 3 files. You can set the compression for the jpeg file.

Protocol steps for GE 1-color and miRNA images

GE 1-color Feature Extraction analyzes GE 1-color images using the same steps listed for GE 2-color images except for the steps for correcting for dye biases and computing ratios.

Agilent miRNA microarays are currently in development. Please check the Agilent website for latest information.

miRNA The miRNA analysis also does not correct for dye biases nor compute ratios. Instead, the analysis estimates the expression level by sampling multiple probes with multiple features per probe and reporting a total gene signal independent of microarray design, rather than a processed signal (counts/pixel).

The protocol, then, does not contain the Correct for Dye Biases step nor the Compute Ratios step, but it does contain a MicroRNA Analysis step after the Calculate Metrics step.

Change the settings for each 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** button 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.

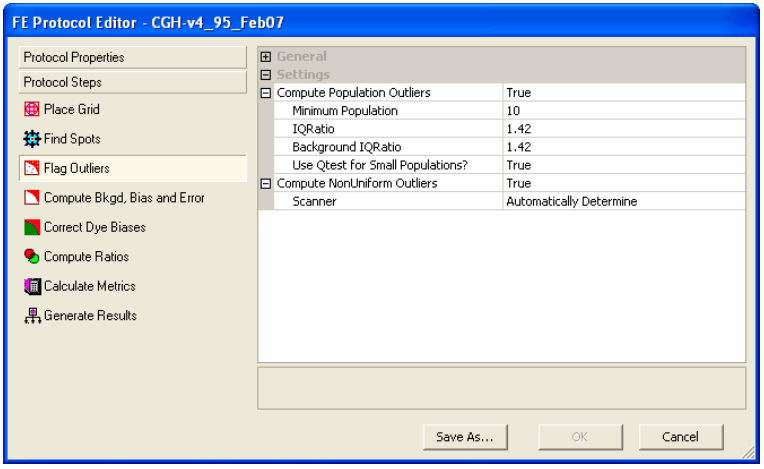


Figure 44 FE Protocol Editor with Protocol Steps on left

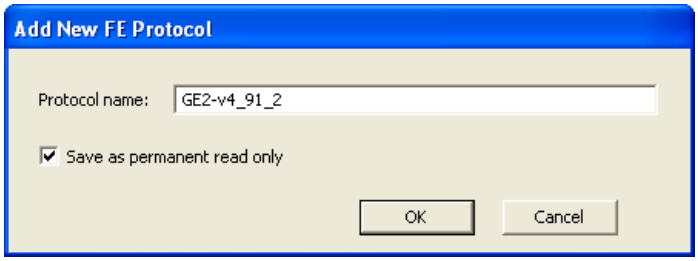
See the remaining sections in this chapter to learn how and why you change parameters for each protocol step.

Save a protocol to 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**.



3 Enter a **Protocol name**.

4 If you want to make the protocol editable, clear the **Save as permanent read-only** check box.

For standard protocols that you use frequently, you may want to save these protocols as permanent, read-only protocols so that others cannot alter them.

5 Click **OK**.

Change 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.

Access the settings to place a grid

- Click the **Place Grid** button to bring up the settings, if not already there.

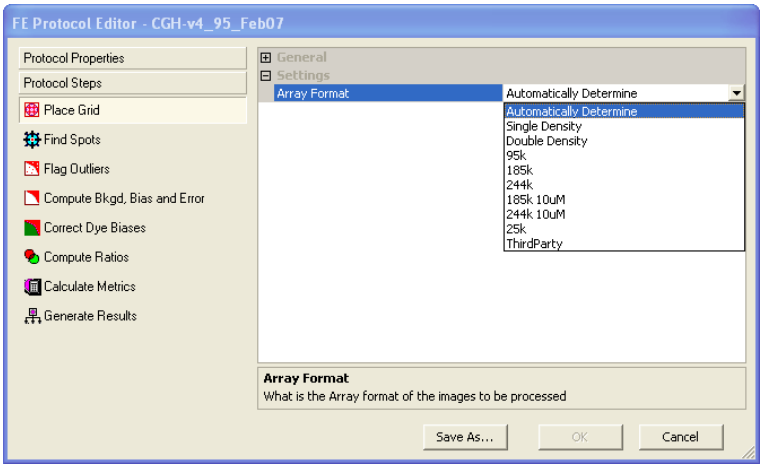


Figure 45 FE Protocol Editor with Place Grid selected – Array Format

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.

The Placement Method option now appears.

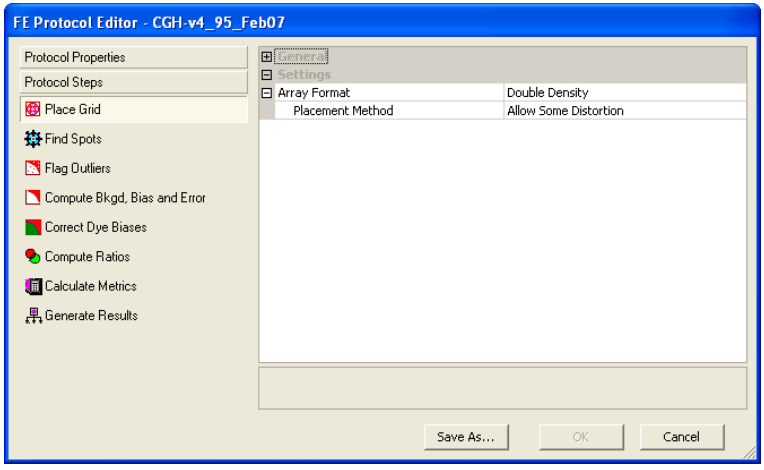


Figure 46 FE Protocol Editor with Place Grid selected – Placement Method

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

Protocols can automatically detect the array format (22k, 44k, 244k) and the scan type (Agilent or GenePix and 5 or 10 microns). 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 Measure Spots).

With this feature, you do not need different protocols for 22k, 44k and 244k array format (as well as Agilent and GenePix scans). 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.

Select a method for placing the grid

1. Click the cell next to **Placement Method**, and click the **Up** 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.

Change Protocol Step 2: Find Spots

See Chapter 1, "Default Protocol Settings", of the Reference Guide *for a setting comparison between the various microarray formats.*

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

Access the settings to find spots when format is automatically determined

- Click the **Find Spots** button in FE Protocol Editor.

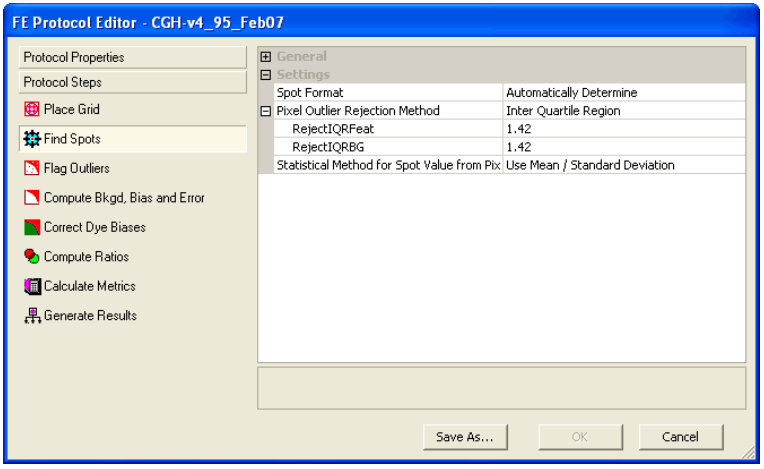


Figure 47 FE Protocol Editor with Find Spots selected—Autodetermine

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.

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

Select the statistical method to reject pixel outliers

1 Click the cell next to **Pixel Outlier Rejection Method**, and click the **Up** 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 "Step 5: Reject outliers" on page 204 of the *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.

Select whether to use mean or median for data analysis

1 Click the cell next to **Select Method for Spot Value from Pixes**, 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 to use for downstream analysis.

Access the settings to find spots when the format is selected

1 Click the cell to the right of the **Spot Format** cell.

2 Click the down arrow, and select from the list of spot formats.

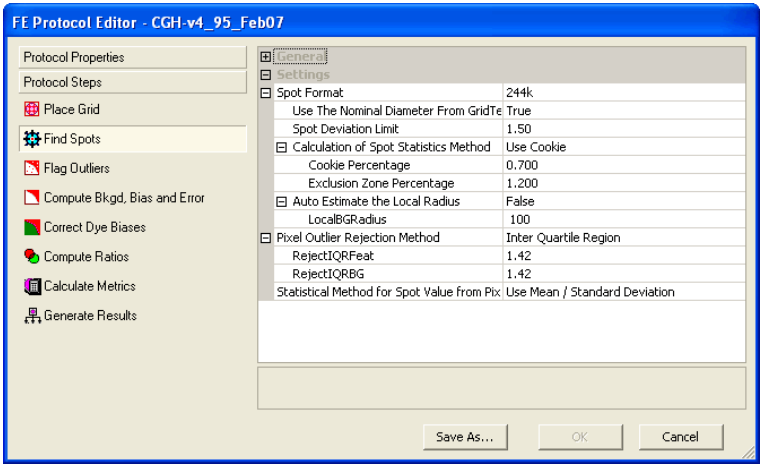


Figure 48 FE Protocol Editor with Find Spots and Spot Format selected

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.

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.

Select a spot statistics method to define features

1 Select either **Use Cookie** or **Use WholeSpot**.

CookieCutter Method

The CookieCutter 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 CookieCutter method is usually used for Agilent microarray spots because of their uniformity.

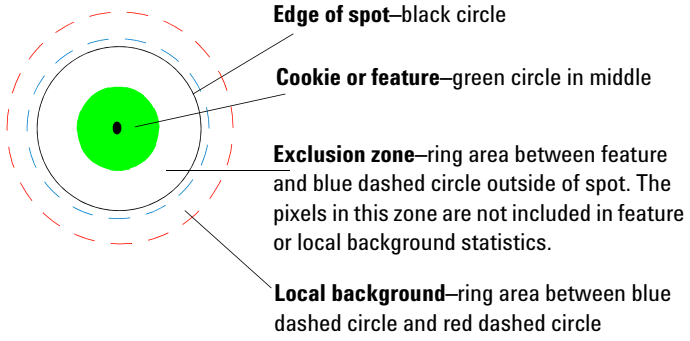


Figure 49 Feature and background definitions – CookieCutter method

WholeSpot 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.

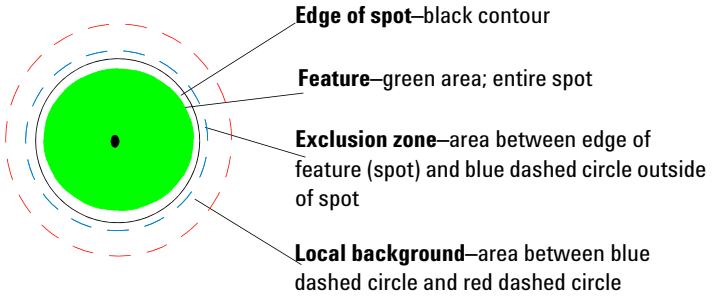


Figure 50 Feature and background definition – WholeSpot method

2 If you selected **Use Cookie**, change the **Cookie Percentage** and/or the **Exclusion Zone Percentage**, if necessary.

If you selected **Use WholeSpot**, 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

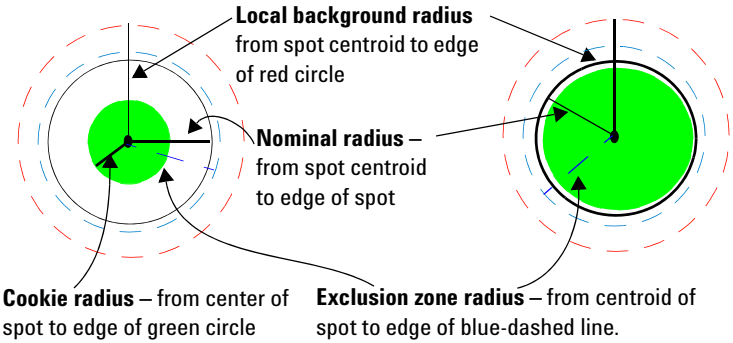


Figure 51 Definitions of different radii

See "Step 4: Estimate the radius for the local background" on page 201 of the Reference Guide.

The radii depicted in Figure 51, "Definitions of different radii," on page 160 are calculated differently for the two methods.

Select to autoestimate the local radius to define local background

This radius method estimates background that is local to a feature of interest.

1 Click the cell next to **Autoestimate the local radius** and select **True**.

2 Select **True** or **False**.

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.

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.

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.

See "Step 4: Estimate the radius for the local background" on page 201 in the Reference Guide for details on this calculation.

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, 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 effective saturation limit could be as much as 20x higher than 65535.

Propagation of Find Spot errors to other algorithms

Contact your local Agilent support representative for a copy of the Agilent paper (confidential) on error modeling to learn more about error propagation.

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 are propagated to the background subtraction algorithm.

Change Protocol Step 3: Flag Outliers

Not every feature is created, hybridized and scanned perfectly. And 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.

Access the settings to flag outliers

- Click the **Flag Outliers** button to change settings for flagging abnormal features and backgrounds.

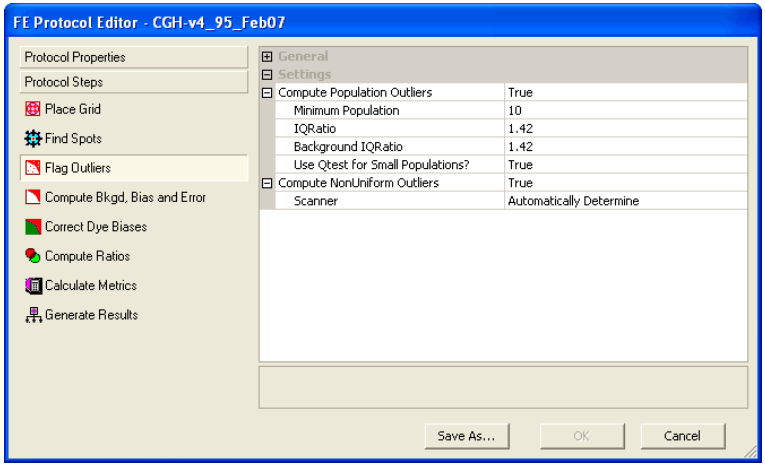


Figure 52 FE Protocol Editor with Flag Outliers selected

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 backgrounds that have intensities beyond the confidence limit for the distribution of the intensities for the replicate features and local backgrounds on the microarray.

- Change settings to flag non-uniform features and background (outliers)

  This setting is dependent on the scanner used to scan the microarrays: Agilent scanner or GenePix scanner.

Change settings to flag population outliers

A second 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 **Up** arrow.

2 Select **True** or **False**.

3 If you select **True**, change any of the default parameters that you need to.

The Flag Outliers sheet 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 signal lies outside of this range, Feature Extraction flags that feature or local background 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 x IQ Range (Interquartile Range) encompasses 99% of normal distribution. The Interquartile Range is the 75^{th} percentile - 25^{th} 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 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 *Reference Guide* to learn more about Q-test calculations.

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.

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 (Agilent or GenePix) 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 **Up** arrow.

2 Select **True** or **False**.

"True" is the default setting for all protocol templates. If True is set, move on to the next step.

3 Click the cell next to **Scanner,** and click the **Up** arrow.

4 Select either the **Agilent Scanner** or **GenePix Scanner.**

The settings change depending on your selection.

To learn more about how the algorithm calculates the terms in the polynomial equation below, see "Flag Outliers" on page 207 of the Reference Guide.

5 Click the cell next to **Automatically Compute OL Polynomial Terms**.

6 Select **True** or **False**.

A different set of parameters appears for a True or a False setting.

7 Enter the values for the parameters.

Parameters for a "False" setting

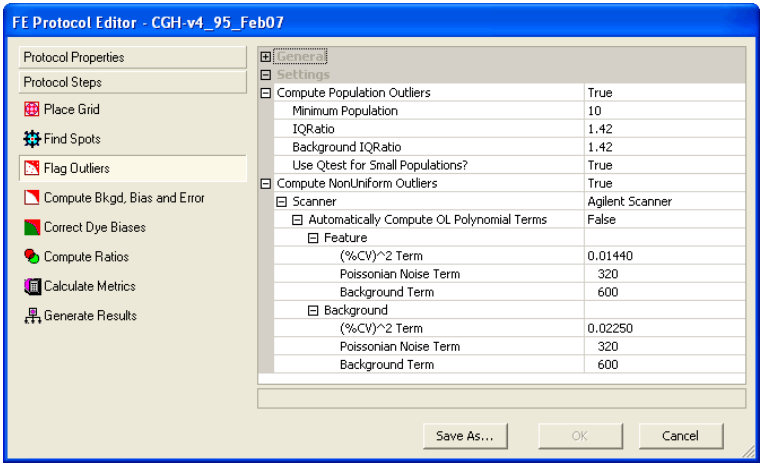


Figure 53 FE Protocol Editor for Flag Outliers: Agilent scanner selected

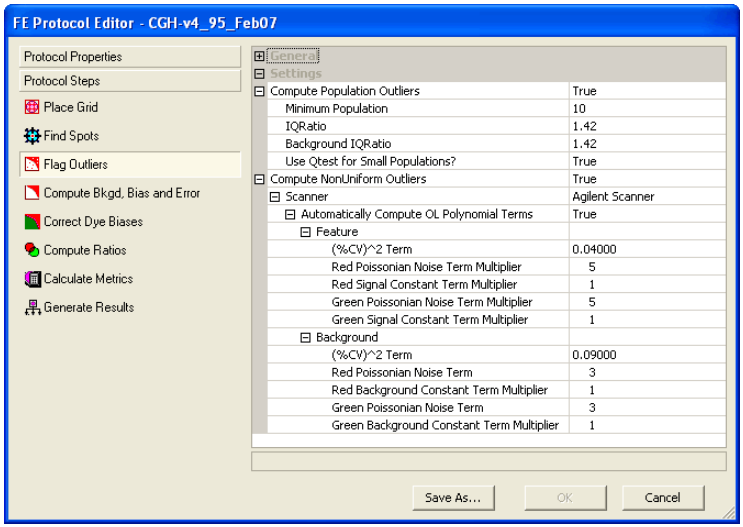
A noise model, based upon a polynomial equation, is made up of three variables whose values are shown in the Automatically Compute OL Polynomial Terms folder:

(%CV)^2 Term (A) This term is the square of the CV (the constant coefficient of variation), or the intensity-dependent variance.

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).

Parameters for a "True" setting



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 multipliers for the B and C terms in both color channels (2-color) or one channel (1-color).

$(\%CV)^2$ Term (A) Same as for the "False" case above

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 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 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 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 features or background vs variance.

CAUTION

The default values for these terms have been optimized for Agilent microarrays, either with an Agilent scanner or GenePix scanner, and recommended lab protocols. You must perform optimization experiments to determine the optimum values for your microarrays and lab protocols.

Change Protocol Step 4: Compute Bkgd, Bias and Error

At this stage of the Feature Extraction process, several systematic sources contribute to background signal that may still remain in 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 "Statistical results (STATS)" on page 133 in the *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 levels can also vary depending on the dye label and color channel used.

Access the settings to compute background, bias and error

- Click the **Compute Bkgd, Bias and Error** button to select a method and change settings for background subtraction.

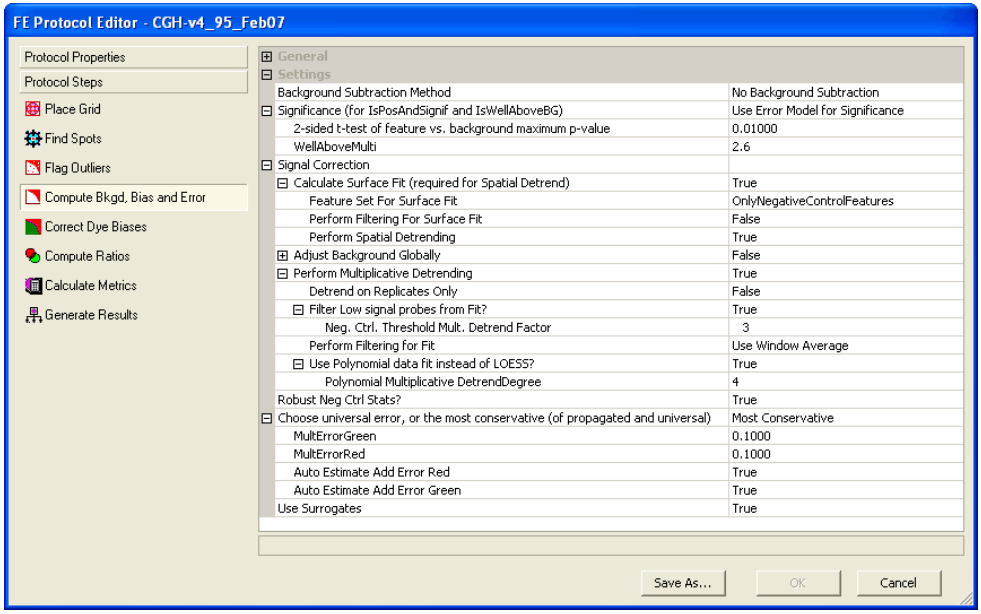


Figure 54 FE Protocol Editor for Correct Bkgd and Signal Biases

The selections for this protocol step let you decrease the contribution of the background signal to the feature signal.

To view the values used for background subtraction depending on the subtraction and correction methods selected, see Table 22, "Stats results contained in the text output file (STATS table)," on page 133 of the Reference Guide.

- 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. This is especially useful for 1-color and CGH microarrays.
- Choose to globally adjust the initial background-subtracted signals.

- Choose an error method to estimate the error in the calculations for the processed signal.
- 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.

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 **Up** arrow.

2 Select the background subtraction method most relevant to your microarray and processing conditions.

The default method for Agilent microarrays loaded with a design file is *No background subtraction*. The default method for 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.

Local background method

This method subtracts the local background estimated for each spot using the radius method in the Find and Measure 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 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 If a microarray system does not show additivity of the local background signal, then a global background subtraction method is needed. Global methods require that both non-uniformity outlier and population outlier methods be turned on. See the topic "Change Protocol Step 3: Flag Outliers" on page 163 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 and Measure 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.

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.

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.

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.

Agilent provides validated negative controls for in situ microarrays. These probes have been shown to reflect the non-specific signal level that is present on probes.

If there is more than one probe sequence on the microarray annotated in the design or grid file as a negative control, this method analyzes the feature population 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.

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.

Choose to remove the surface trend found in the data

What is detrending?

To learn more about the detrending calculations, see "Step 11: Calculate the feature background-subtracted signal (BGSubSignal)" on page 213 of the 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). For 1-color data, Feature Extraction also removes any "domed" surface trend that cannot be cancelled through the use of 2-color ratios, a process called *multiplicative detrending*.

This option is also available for 2-color data and is the default choice with newer protocols. 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.

Background Spatial Detrending

For spatial detrending, Agilent separates the surface fit and subtraction processes that must exist together to remove the surface trend found in the data.

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.

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 or subtracting the surface fit through this foreground, the data becomes more consistent and reproducible.

For some very small microarrays, the algorithms are unable to calculate a surface fit. If an error appears that says, “Execution error: Spatial Detrending: too few data points available to calculate the fit. ()”, turn off the Calculate Surface Fit option.

Instructions

Follow the steps below to set up or change the settings for background detrending.

1. Click the cell next to **Calculate Surface Fit**, and click the **Up** arrow.
2. Select **True** or **False**.
3. If you select True, change the settings that you need to.

Feature Set for Surface Fit

Select one of the following sets of features for which you intend to calculate the surface fit for a set of features with low signal.

AllFeatureTypes Select if microarray has a set of very low intensity features evenly distributed on the slide – default for gene expression protocols

OnlyNegativeControlFeatures Select this if you have enough negative controls (around 2% on a 44k microarray) 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 the time needed to calculate the surface without reducing the accuracy of the fit, especially for 244k arrays.

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 or linear trend found from the surface fit from the data.

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 **Up** 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 combination of local background subtraction, spatial detrend and global background adjustment provides a hybrid background correction method that is unique to Agilent Feature Extraction algorithms.

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.

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.

- 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 will be normalized to the respective replicate averages before going into the fitting set.

Filter Low signal probes from Fit? You can chose 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.

Use Window Averaging You must decide whether to use all of the data points or use local averages or medians.

Surface Type You can chose a Polynomial surface fit or LOESS. If you chose polynomial you can choose the order of the polynomial.

If you are setting up to extract 1-color data, double-click the cell next to Perform Multiplicative Detrending to select True or False.

All 1-color protocols have a default setting of True. Newer 2-color protocols have a default of True, whereas older 2-color protocols have a default of False.

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 Chapter 5, "How Algorithms Calculate Results", in the *Reference Guide* for more information on how this algorithm works.

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 **Up** 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.

See **"Parameters for Universal and Most Conservative error models"** 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 (Tables 12-16 in Chapter 6). If you select **True**, then FE 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 or Agilent microarrays with non-Agilent protocols, 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 multiplicative error does not vary much from microarray to microarray. But the additive noise term does 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.

For older Agilent microarrays where 8 x formats have 1.9k features, the auto-estimation of the additive error uses only the standard deviation of the negative control because spatial

detrending is not recommended. When an 8 x format image is extracted, a message appears to remind you that the auto-estimation used only negative control statistics.

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 (e.g., 25 for Agilent *in situ* oligo microarrays). For high quality microarrays processed correctly, the values are quite similar for the two methods of calculation. If, however, 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.

Effect of additive error selection on correction method for dye normalization You can use any upstream DyeNorm method with the auto-estimation method. However, if the auto-estimation option is cleared and the additive error is set to a constant that was previously optimized for the Agilent microarray-based Gene Expression system (e.g. constant = 25 for *in situ* microarrays with linear-amp target), then you should choose the Normalization Method = "Linear&LOWESS". The default constant of 25 was derived with the Linear&LOWESS method enabled.

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 will be 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 **Up** 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 its background standard deviation (i.e., BGSubSignal < BGSDUsed)

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 standard deviation (SD) of the background used for that feature in the green or red channel. The SD is either a population SD, in the case of global background methods, or it is a pixel level SD, in the case of the local radius method.

For the case when the "No Background" option is chosen, the SD is the pixel level SD of the local background. These SDs are reported in the Feature Extraction output columns, "g(r)BGSDUsed" (“Feature results (FEATURES)” on page 147 of the *Reference Guide*). These SDs are the same SDs that are utilized for the IsPosAndSignif algorithm of the Background Subtraction tab (“Change settings to test the significance of a feature” on page 183 of this guide).

An example of these cases can be found in the clinical diagnostic industry, for example, a glucose assay. It is common practice to determine the noise of replicated measurements of samples that have [glucose] = 0. The standard deviation (SD) of these signal measurements, or background SD (Bk_SD), is then entered into the calibration curve for that assay to calculate the lowest limit of detection (LLD) for that assay. Any test sample that yields a signal below the Bk_SD will produce a result such as "[glucose] < x ug/ml" (where x ug/ml is the LLD).

Propagation of errors from the Compute Bkgd, Bias and Error algorithm to other algorithms

Contact your local Agilent support representative for a copy of the Agilent paper (confidential) on error modeling to learn more about error propagation.

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 "Select the method to estimate the error (Error Model)" on page 178 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.

Change settings to test the significance of a feature

To view the values used for the t-test depending on the subtraction and correction methods selected, see Table 22, "Stats results contained in the text output file (STATS table)," on page 133 of the Reference Guide.

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.

- Change the default parameter values for the two tests, if necessary.

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 FE 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 {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.

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)”. The first method, set as the default in older protocols, uses pixel statistics for 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 193 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 (FE) text output columns "g(r)BGSDUsed".

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.

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.

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 FE 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 FE text output column.

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.

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.

Change Protocol Step 5: Correct Dye Biases

A *dye bias* is the difference in the red and green signals caused by different efficiencies in labeling, emission or detection.

Systematic differences in dye bias must be removed from the 2-color GE and CGH data. To remove dye bias, the Feature Extraction software uses a dye normalization algorithm. This step is not used with GE 1-color and miRNA data.

Access the settings to correct dye biases

- Click the **Correct Dye Biases** button to change settings to remove difference in dye bias.

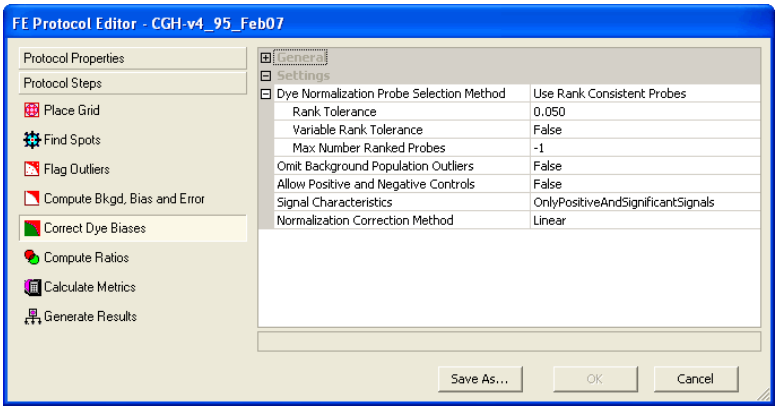


Figure 55 FE Protocol Editor with Correct for Dye Biases selected

You can change five types of settings for this algorithm.

- 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 at a different rate than in the other.

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.

Choose a method to select a set of dye normalization probes

Feature Extraction supplies four different 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 **Up** arrow.

2 Select one of the four methods described below and on the next page.

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 defining features (> 50%).
- They are not population outliers in either channel.

You can enter a *Rank Tolerance* value if you select this 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 %. 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.

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.

Use List of Normalization Genes

This method allows you to specify a file containing a list of normalization genes, often referred to as "house-keeping genes". This is useful if you have identified a set of genes that should have no differential expression, in that particular experiment, and therefore would be good normalization genes.

You can use an external file for the list of normalization genes, or you can create your own list. See "Extracting Microarrays Automatically" on page 27 to learn how to use an external DyeNorm list file or create your own list.

The dye norm list requires that at least 1% of the features on the microarray pass the following filters:

- Feature is not a control probe (e.g. Control Type = 0)
- Feature signal is positive and significant compared to background (e.g. IsPosAndSignif = 1)
- Feature signal is not flagged as non-uniform (e.g. NonUnifOL = 0)
- Feature has fewer than 5% saturated pixels

Use Rank Consistent List of Normalization Genes

For this option you use a gene list subjected to the rank consistency filter. The description for the options Use Rank Consistent Probes and Use List of Normalization Genes holds for this option. You can also choose to allow variable rank tolerance, which means a rank tolerance is done after grouping the data into signal groups. In each group the same X percentage most rank-consistent data are included.

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 X percentage most rank-consistent data are included.

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 244k arrays.

Select to omit background population outliers

1. Click the cell next to **Omit Background Population Outliers**, and click the **Up** 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.

Select to allow positive and negative controls

If you turn this function on, FE uses positive and negative controls in the dye normalization calculation.

1 Click the cell next to **Allow Positive and Negative Controls?**, and click the **Up** arrow.

2 Select **True** or **False**.

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 **Up** arrow.

2 Select one of the following options.

OnlyPositiveandSignificant Signals Use only positive and significant signals for ranking in the rank consistency filter.

AllPositiveSignals Use all positive signals for ranking in the rank consistency filter.

AllNegativeAndPositiveSignals Use all negative and positive signals for ranking in the rank consistency filter.

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 **Up** arrow.

2 Select from one of the following methods:

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 description below) to the linearized data set.

LOWESS 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.

Change Protocol Step 6: Compute Ratios

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.

To determine if a feature is differentially expressed, three measures are important: *log ratio*, *log ratio error* and the *p-value*.

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.

This step is not used with GE 1-color and miRNA data.

Access the settings to compute ratios

- Click the **Compute Ratios** button to change settings to calculate these measures.

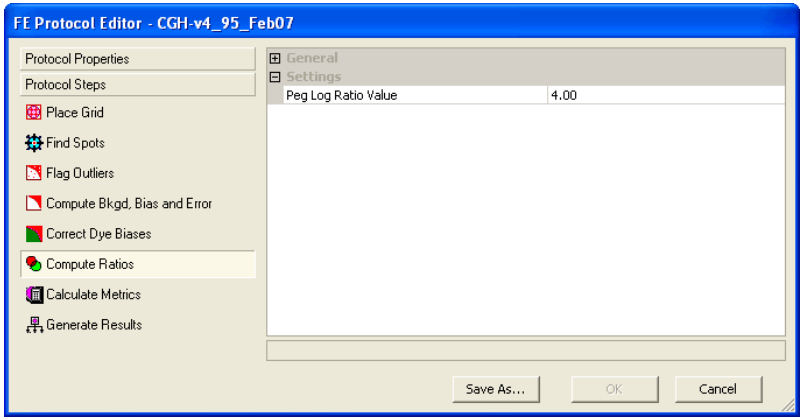


Figure 56 FE Protocol Editor with Compute Ratios selected

You can change only one parameter for this protocol step:

- Enter the value for the log ratio limit

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 Up arrow in cell next to **Peg Log Ratio Value**.

2 Enter a positive or negative value up to 4 that you want.

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 6 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.

Table 6 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*.
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.

*Log ratios are = or between -4 (ratio = .0001) and +4 (ratio = 10000). If (r/G) > 1 or (R/g) < 1, log ratio = 0 and p-value = 1.

Case 1: R/G Both channels use DyeNormSignals. P-value and log ratio are calculated as usual. For signals not using surrogates, the g(r)DyeNormsignal is equal to the g(r)ProcessedSignal, used to calculate log ratios.	Case 2: r/G r = rSurrogateUsed G = gDyeNormSignal P-value and log ratio are calculated as usual. If r/G > 1, then FE software automatically sets LogRatio = 0 and pValueLogRatio = 1.
Case 3: R/g R = DyeNormSignal g = gSurrogateUsed P-value and log ratio are calculated as usual. If R/g < 1, then FE software automatically sets LogRatio = 0 and pValueLogRatio = 1.	Case 4: r/g Both channels use surrogates. FE software automatically sets LogRatio = 0 and pValuLogRatio = 1. For signals using surrogates, the g(r)Processed Signal is equal to the g(r)SurrogateUsed value, used to calculate log ratios. (Cases 2 & 3)

Figure 57 Summary–Use of surrogates for calculations

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.

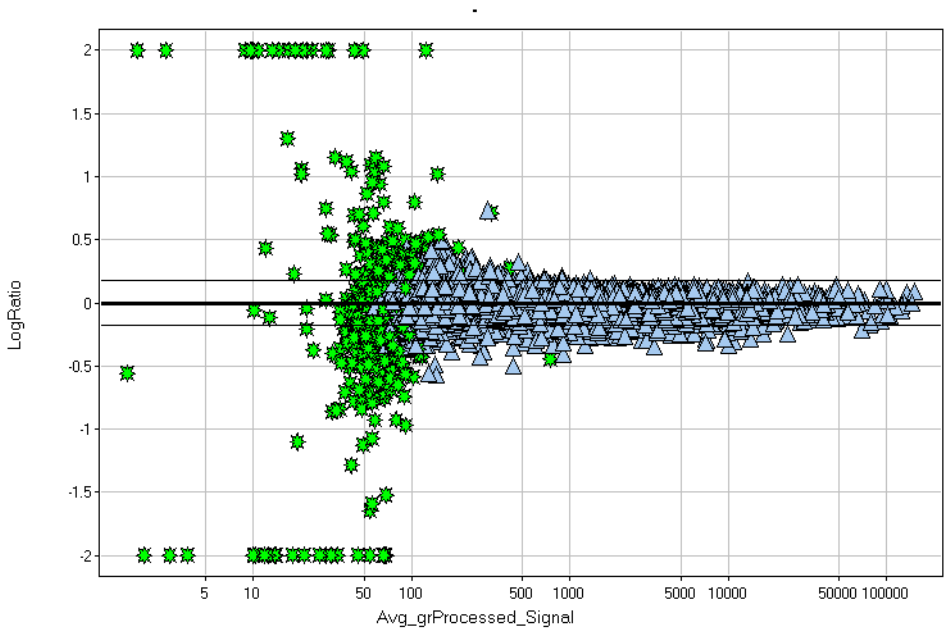


Figure 58 Plot of LogRatio vs Avg_grProcessedSignal–No surrogates

Green stars are data points affected by use of surrogates in Figure 59.

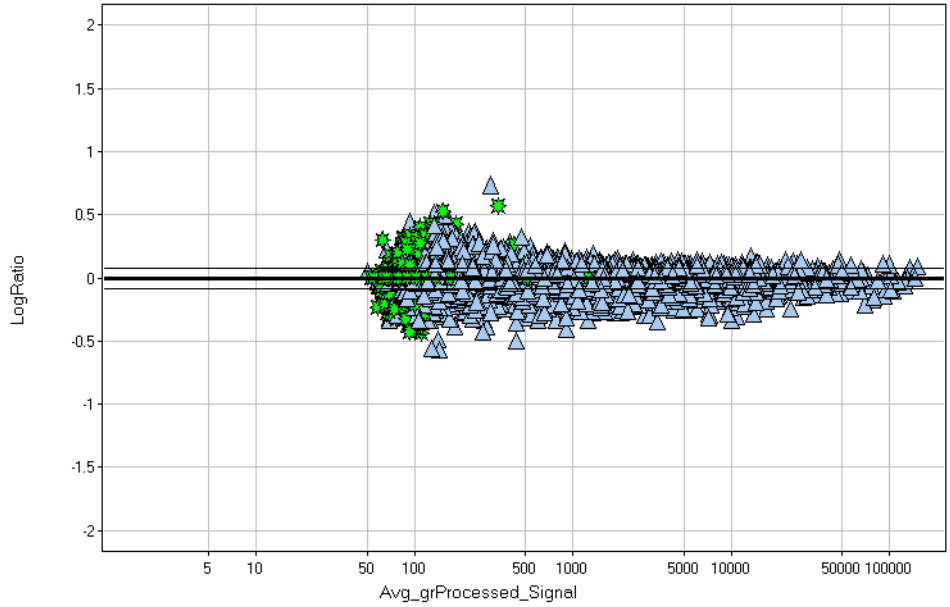


Figure 59 Plot of LogRatio vs Avg_grProcessedSignal–With surrogates

Green stars are data points affected by use of surrogates. (See Figure 58 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 FE sets the log ratio = -4. If green channel < 0, then FE sets the log ratio = +4. If both red and green channels < 0, then FE sets the log ratio = 0.

Change Protocol Step 7: Calculate Metrics

In this step you can change the settings for the metrics that appear in the QC report.

Access the settings to calculate metrics

- Click the **Calculate Metrics** button to change the values for these parameters.

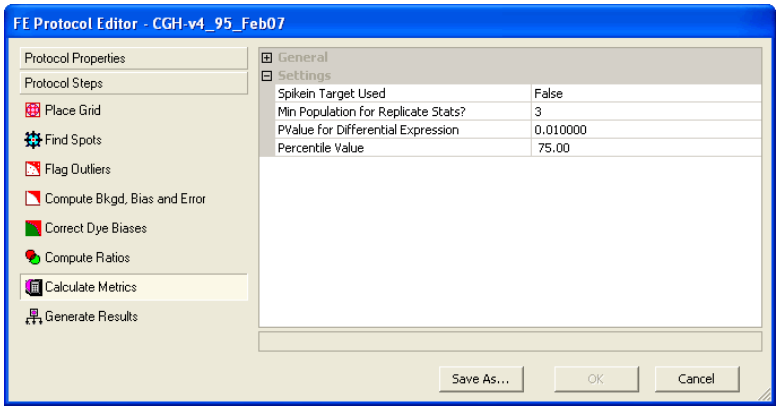


Figure 60 FE Protocol Editor with Calculate Metrics selected

You can change four types of 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.
- Enter the p-value for differential expression.
- Enter the percentile value.

Select to use spikein targets

1. Click the cell next to **Spikein Target Used?**, and click the **Up** 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 spikein controls.

Change the minimum population for replicate statistics

The QC report includes statistics based on replicate probes (%CV, log ratio uniformity). FE 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.

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.

Change the percentile value

This value is an indicator of the distribution of signals of the non-control features and is useful in downstrem 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.

Change Protocol Step 7b: MicroRNA Analysis

This step will be used when the miRNA protocol becomes available.

This step appears in the protocol only for microRNA protocols. With this step you set up to calculate a Total Gene Signal. See the *Reference Guide* for more details on how Feature Extraction processes a miRNA image and on the QC report and text file results, algorithms and protocol default values.

Both the protocol parameters and default values may change over time. Check the web site for updated protocols: http://www. agilent.com/chem/feprotocols.

Access the settings to analyze microRNA (miRNA) images

Agilent miRNA microarays are currently in development. Please check the Agilent website for latest information.

- Click the **MicroRNA Analysis** button to change the settings for the analysis.

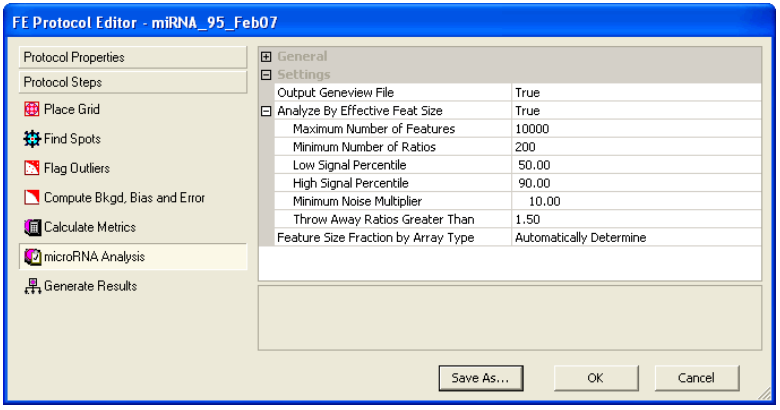


Figure 61 FE Protocol Editor with MicroRNA Analysis selected

You can change two settings for this protocol step.

- Choose to generate the results to a GeneView file, or not.
- Choose to scale the analysis with the effective feature size parameters, or not.

Choose to generate the results to a GeneView file

After running Feature Extraction with the microRNA protocol, a new file is generated in addition to the tab-text file, MAGE-ML file, QC Report and any other files selected. This file is generated in the same directory as the other files and is named ArrayName_GeneView.txt.

Each row includes a “gene” on the microarray, or a unique Systematic/Gene name in the design

This file is a tab-text file with the following columns:

- SystematicName
- ControlType
- gTotalGeneSignal
- gTotalGeneError
- IsGeneDetected

1 Click the cell to the right of **Output GeneView File**.

2 Click the down arrow, and select **True** or **False**.

Choose to scale the analysis with the effective feature size parameters

The calculation of the feature size is different for miRNA images. Even though you can change the default value from True to False, you will almost never have to do this.

1 Click the cell to the right of **Analyze By Effective Feat Size**.

2 Click the down arrow, and select **True** or **False**.

Even though the default value is True and these parameters appear, you will almost never have to change them:

Maximum Number of Features Maximum number of features used to compute the effective size

Minimum Number of Ratios Minimum number of ratios used to compute the effective size

Low signal Percentile	Signal filter for computing the effective size of the spot. Uses signal in at least n (default–50%) percentile of all signals
High Signal Percentile	Signal filter for computing the effective size of the spot. Uses signal in at most n (default–90%) percentile of all signals.
Minimum Noise Multiplier	Makes sure the signals are at least x times above the background noise for use in computing the effective spot size
Throw Away Ratios Greater Than	Ratios > 1 mean the outside of the cookie is brighter than the inside of the cookie. This may be due to crescents or non-uniformity or other anomalies.

Change Protocol Step 8: Generate Results

Access the settings to generate results

- Click the **Generate Results** button to change the values for output.

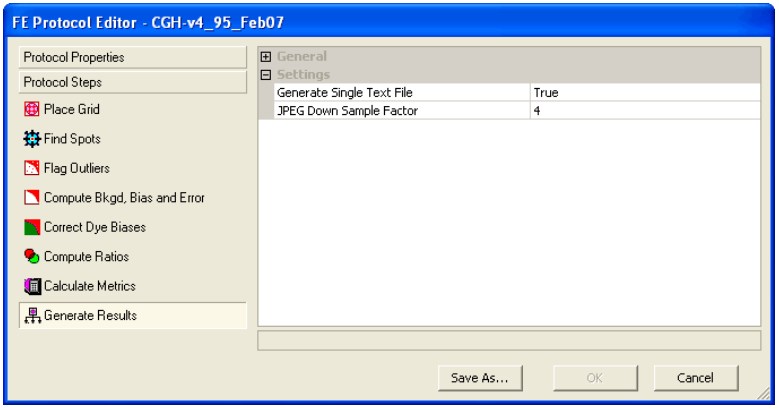


Figure 62 FE Protocol Editor with Generate Results selected

You can change only two settings for this protocol step.

- Select to generate a single file for the text output
- Change the compression factor for the jpeg output

Select to generate a single file for the text output

1 Click the cell next to **Generate Single Text File?**, and click the **Up** 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).

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 **Up** arrow.

2 Enter a number.

4 is the factor used in version 7.5. If you select 1, the JPEG output is compressed by a factor of 4.

5
Changing Image Displays

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

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.

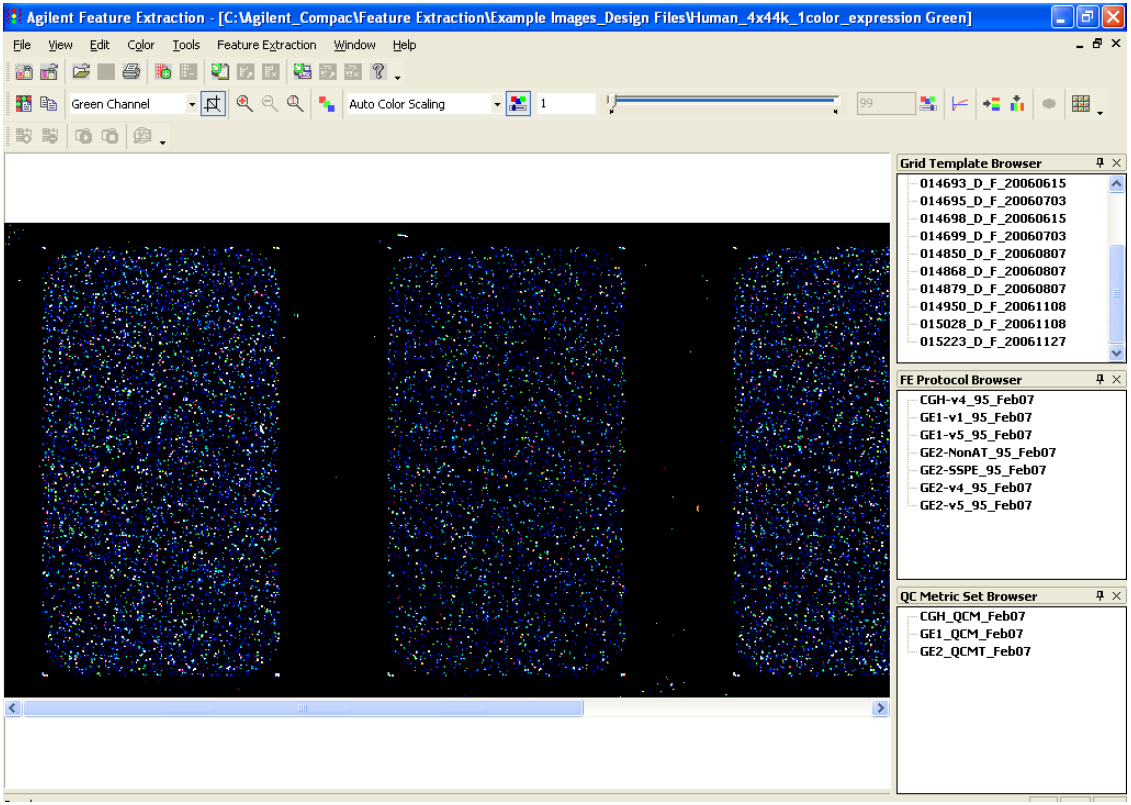


Figure 63 Uncropped TIFF image for Agilent microarray

You can also open an image file double-clicking the image file in in a project extraction set. For information on how to do this, see Chapter 2, “Extracting Microarrays Automatically”.

Crop an image

Crop an image using the cursor

You no longer 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 probably turned ON because this is the default setting when the software ships. (See “Change default image viewing settings” on page 233.) Therefore, when you move the cursor to form a rectangle around a section of the image, you crop, not zoom, the image.

Images should be cropped so that there is a border whose distance from the perimeter of the microarray is two to five times the spacing between two spots.

1 If not already on, click the **Cropping Mode on/off** button.
2 Position the cursor at the top left corner of the image that you want to crop, and hold the left mouse button down.
3 Drag the cursor to the opposite corner.
4 Release the left mouse button to display the cropped image.

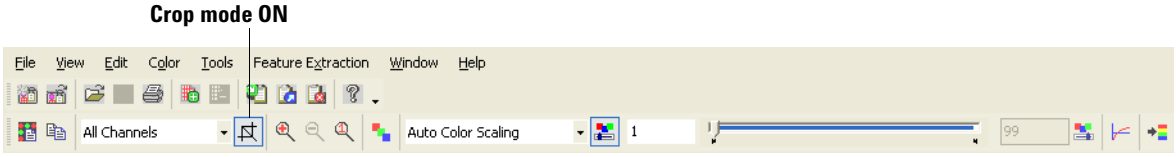


Figure 64 Image top menu and toolbar in Feature Extraction Tool

You can cancel the cropping operation by using the (Esc) key, or click the Cropping mode on/off button to turn cropping off.

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.

1 Select **Edit > Manual Crop**.

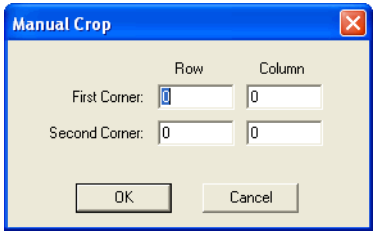


Figure 65 Manual Crop dialog box

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**.

Save a cropped image to a new file

If the image file is from an Agilent microarray with a barcode:

1 Select **File > Save Modified Image** to save the cropped image to a new file.

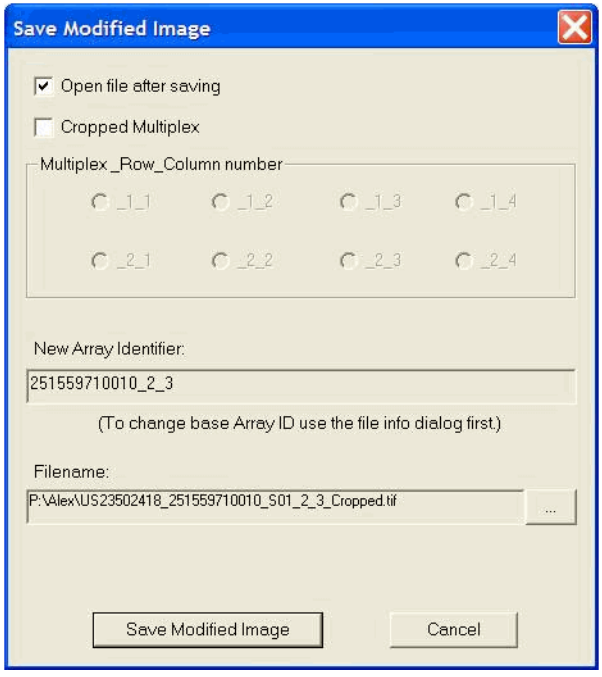


Figure 66 Save Modified Image dialog box for Agilent microarray

2 To close the file after saving, clear the **Open file after saving** check box.

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 "Add or change the New Array Identifier" on page 211.

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:

- Select **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. In version 8.5 you can add an identifier. See "Add or change the New Array Identifier" on page 211.

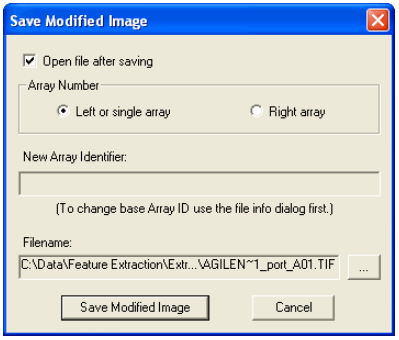


Figure 67 Save Modified Image dialog box for non-Agilent microarray

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. Select **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 select **File > Save As** before the change is accepted.
4. Click **OK**, and select **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.

Load Feature Extraction visual results

You can select to store visual results either in the same directory as the image file or in the results directory. In either case the visual results file is stored with the same name as the results file name with an extension of .shp (shape).

Version 9.5 can read the visual results that have been saved using version 6 and later of Agilent Feature Extraction Software.

1 Open the image for which you want to see the visual result.

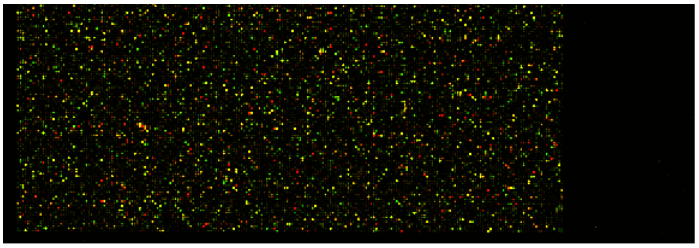


Figure 68 Image of Agilent single density microarray

2 Select **Feature Extraction > Load Visual Result** to load previously saved visual results for the currently displayed image file.

   The software opens the directory 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 directory separate from the image file directory, you must move to that directory.

3 Select the shape file, and click **Open**.

4 Click the **Zoom** icon repeatedly until you can see the marked features.

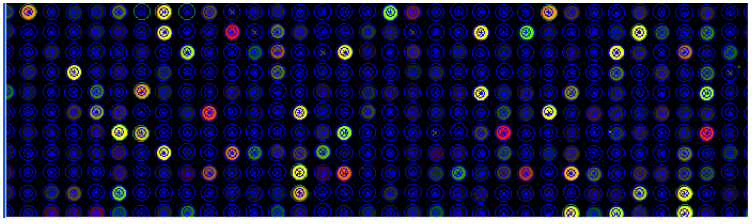


Figure 69 Visual results for image in Figure 68

If more than one image file reside in the image file directory, then you must remember the name of the extraction set for the shape file you want to view. This is also true if you or someone else changed the name of the original extraction set that produced the shape file.

If you saved the shape file to another results directory, you must remember the name of the directory and the name of the extraction set for that shape file.

Save an image (original or cropped) for import to another program

1. Select **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** or **Jpeg Files**.
4. Enter the name of the image.
5. Click **Save**.

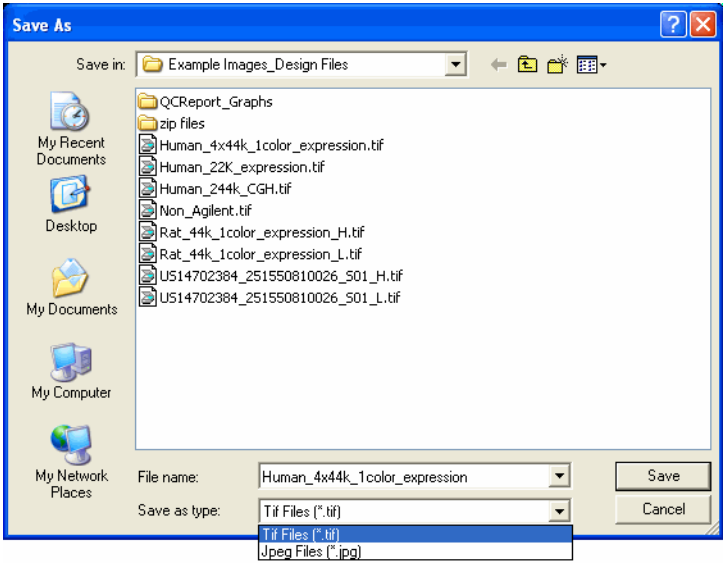


Figure 70 Save As dialog box

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

- Select **View > Image File Info** to display a tabbed dialog box that displays all available information about an image file.

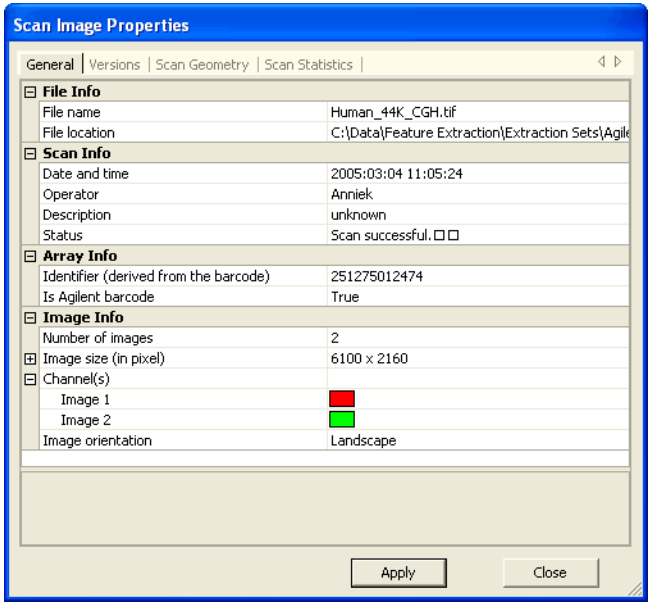


Figure 71 Scan Image Properties dialog box

The only field that you can change is the Identifier field. See "Add or change the New Array Identifier" on page 211.

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.

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.

1 Select **Tools > Flip Upper Left to Lower Right (Landscape/Portrait)** from the top menu.

   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.

2 Select **File > Save modified image**.

   If you do not save the image in its new orientation, you could run into problems with Feature Extraction later.

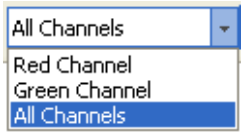
Change the channels that display

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.

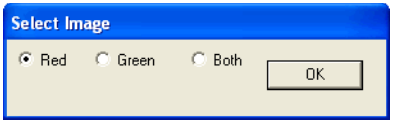
You can also change the number and color of the channels displayed when you set default image settings.

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 Select **Window > New "Filename" View Window**.

"Filename" represents the path and name of the image that you loaded. The Select Image dialog box appears.



2 Select **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.

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, select **Color > "color mode"**.

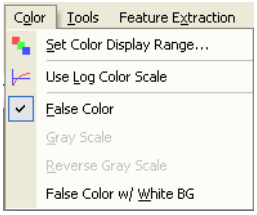


Figure 72 Color menu

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.

- Select **Color > False Color w/ White Bgd**.

  To return to False Color, select **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.

- Select 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, select **Color > False Color**.

Examples of images with different color selections

The following images are intentionally not aligned.

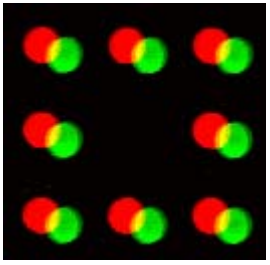


Figure 73 False Color for 2-color image

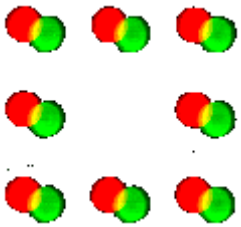


Figure 74 False Color with white BG for 2-color image

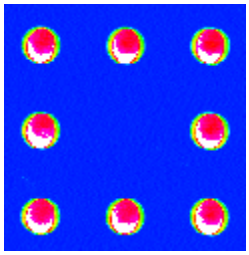


Figure 75 False Color for 1-color image

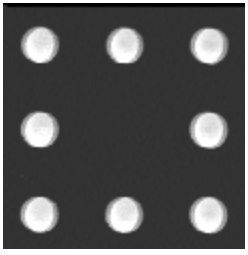


Figure 76 Gray Scale

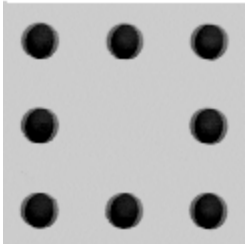


Figure 77 Reverse Gray Scale

Change between log and linear color scale

- Select **Color > Use Log Color Scale** to toggle the color scale from linear to $\log_{10}$ (intensity).

  To return to the linear scale, select **Color > Use Log Color Scale** again to remove the checkmark.

  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. 

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 78 Color scaling segment of the image viewing toolbar

1 Click the down arrow next to the color scaling menu, and select **Auto Color Scaling** if necessary.

2 Click the **CDR Low Mode** icon, and move the lower bar cursor to the minimum percentage for the display range.

3 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.

4 To fine tune the color display ranges manually, select either the **Red Channel** or the **Green Channel** from the color scaling menu.

5 Click the **CDR Low Mode** icon, and move the lower bar cursor to the minimum pixel number for the display range.

6 Click the **CDR High Mode** icon, and move the higher bar cursor to the maximum pixel number for the display range.

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 Select **Color > Set Color Display Range**, or select **Set Color Range** on the toolbar.  .

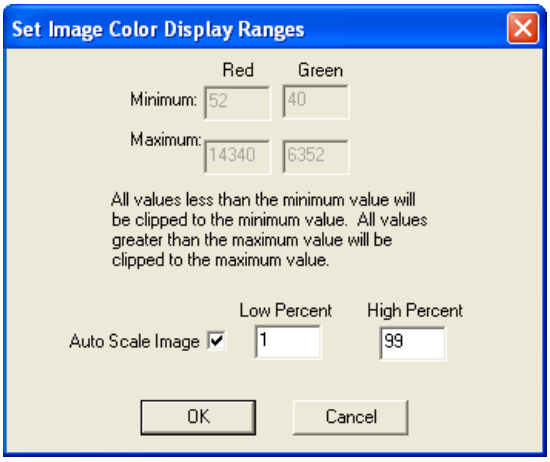


Figure 79 Set Image Color Display Ranges dialog box

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.

Change the magnification (zooming)

To set the default % zoom settings for the loaded image and the grid mode images, see "Change default image viewing settings" on page 233 of this guide.

Magnification can only be set to one of the % zoom levels shown in Figure 80 . 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.

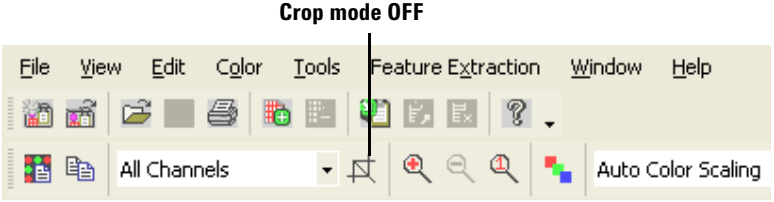
You can control zooming in one of three ways:

- 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.

1 Unclick the **Crop** 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.

Table 7 Tool Bar and Mouse Zooming Buttons

Tool Bar Button	Mouse Short Cut	Description
	Hold down the Ctrl key and double-click with the left mouse button on the image.	Increases the magnification to the next higher level.
	Hold down the Ctrl key and double-click with the right mouse button on the image.	Decreases the magnification to the next lower level.
	None.	Changes magnification to 100%.

Use the Zoom Menu

You can use the Zoom menu with Crop Mode ON or OFF.

- Select **View > Zoom** to jump to any zoom level directly.

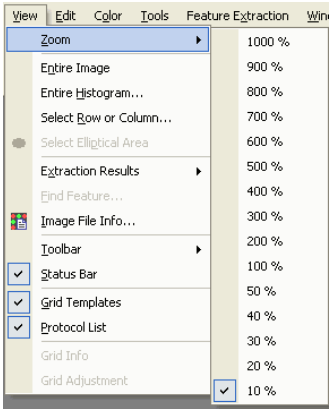


Figure 80 Zoom menu

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.

- Select **View > Entire Image** to resize the window.

Change the extraction results display

Use this option to control the portions of the feature extraction results that are displayed.

- Select **View > Extraction Results**.

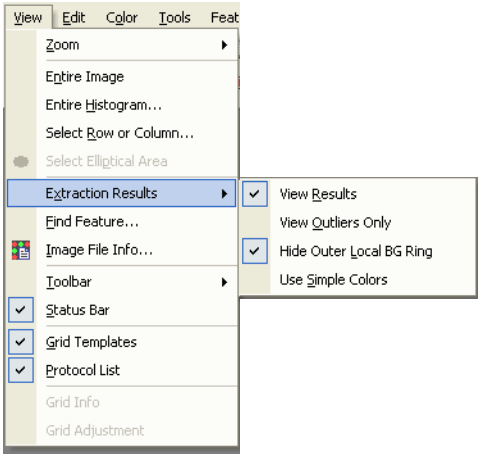


Figure 81 Extraction Results View selection

View Results Shows or hides all visual results

View Outliers Only Shows or hides outlier features

Hide Outer Local BG Ring Shows or hides the local background ring. See "Select a spot statistics method to define features" on page 159 of this guide for a definition of the local background.

Use Simple Colors Toggles the color display of outlier features and background.

Change and save histograms and line graphs

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. Select **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. Select **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.

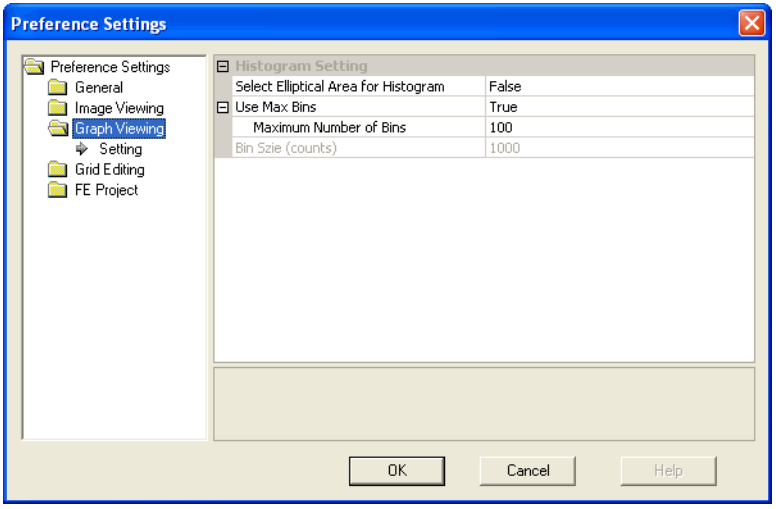


Figure 82 Histogram Preference Settings dialog box

4 Make sure that cropping is OFF.

5 Click  .

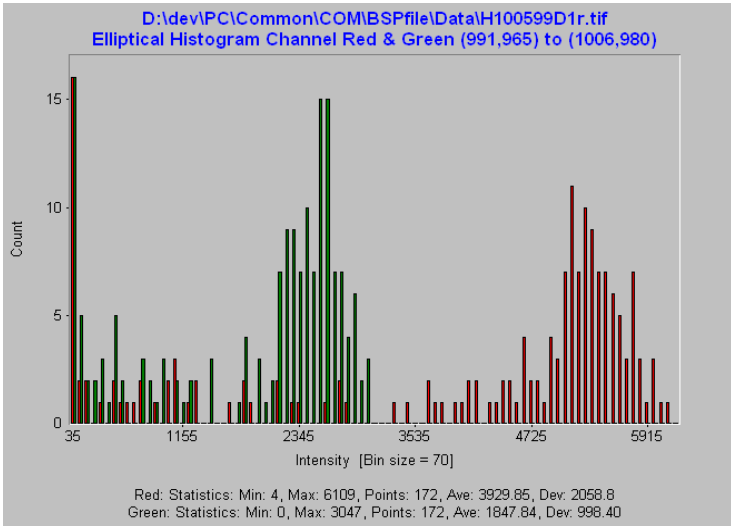
Or, select **View > Select Elliptical Area**.

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, select **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 Select **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 Select **View > Select Row or Column**, or click  or  .

The Select Row/Col dialog box appears.

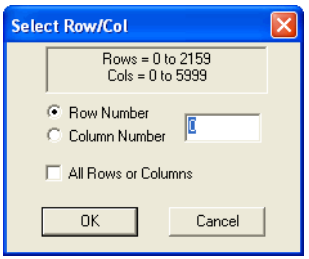


Figure 83 Select Row/Col dialog box

2 Select **Row Number** for showing signals in a row, or select **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**.

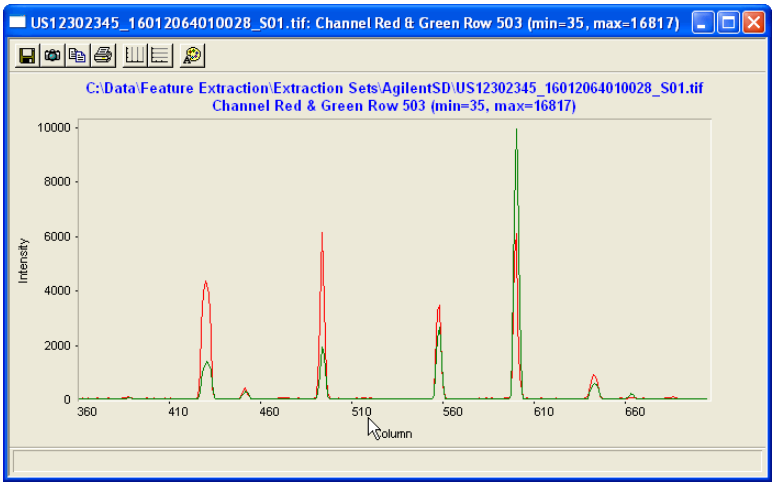


Figure 84 Line graph example

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.

Table 8 Graph Toolbar

Tool Bar Button	Description
	Save a copy of the data in the graph to a tab delimited text file.
	Put a copy of the graphical image into the window's clipboard. This allows you to paste the image into another program, such as a word processor.
	Put a copy of the image data as text into the window's clipboard for export to a spread sheet.

Table 8 Graph Toolbar

Tool Bar Button	Description
	Print the graph.
 	Add horizontal or vertical grid lines.
	Change the fonts used for various titles and scales. This is sometimes helpful when printing or copying the image to another program.

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 "Change the display color" on page 218, "Change between log and linear color scale" on page 221, and "Change the intensity scale (color display range – CDR)" on page 221 in this chapter for details on the selections in the first two sections of the Image Viewing Preferences window.

1 Select **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.

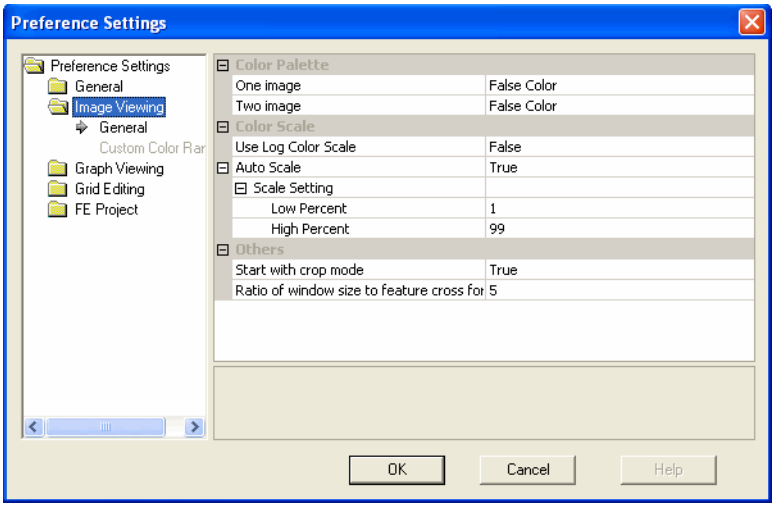


Figure 85 Image View Default Options dialog box

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).

Change the print settings

The Feature Extraction application supports printing to both color and black and white printers.

Change page settings

- Select **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.

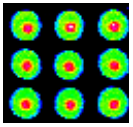


Figure 86 False Color image

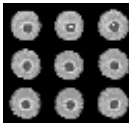


Figure 87 Incorrect image generated by blank and white printer

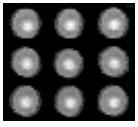
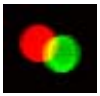


Figure 88 Correct gray scale generated by the Image Analysis Tool

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.



These images are intentionally not aligned.

Figure 89 2-color image

Figure 90 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 9 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
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.

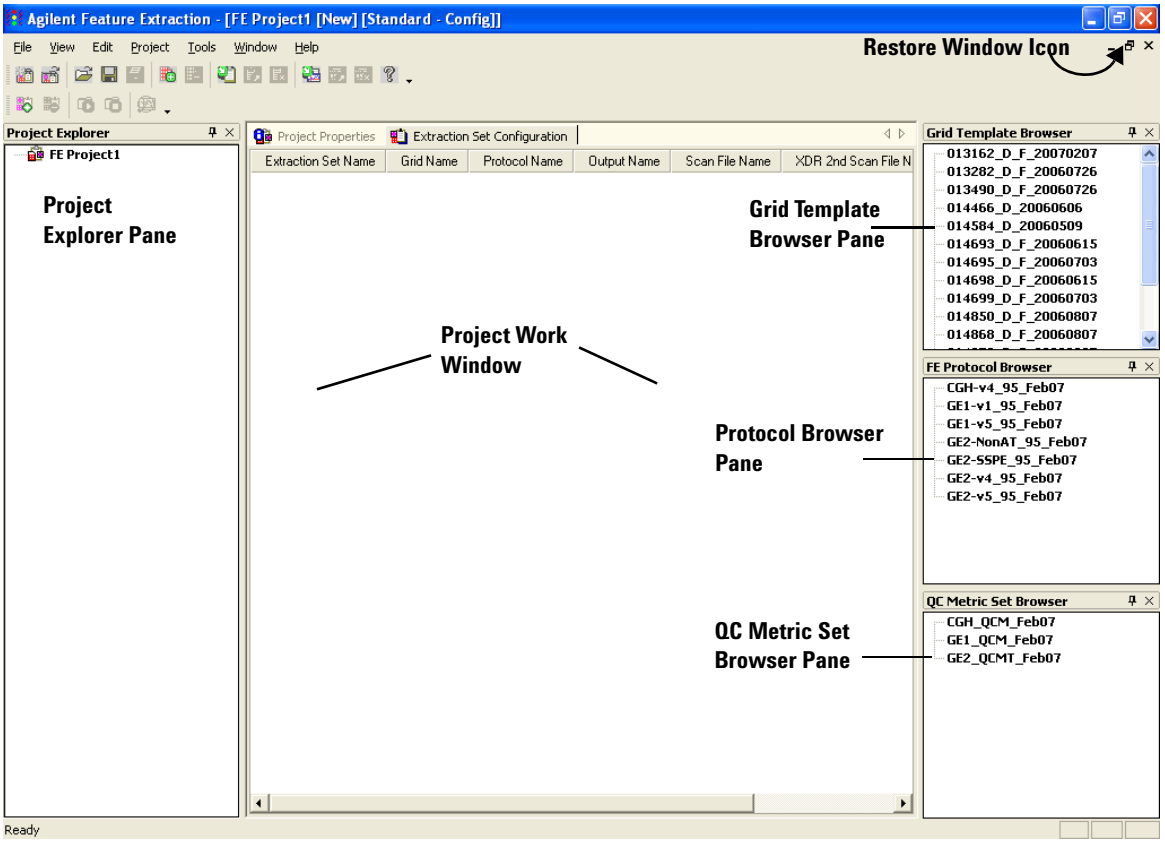


Figure 91 FE Standard Project Configuration window – default

Below is another layout of the FE Standard Project Configuration window. 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.

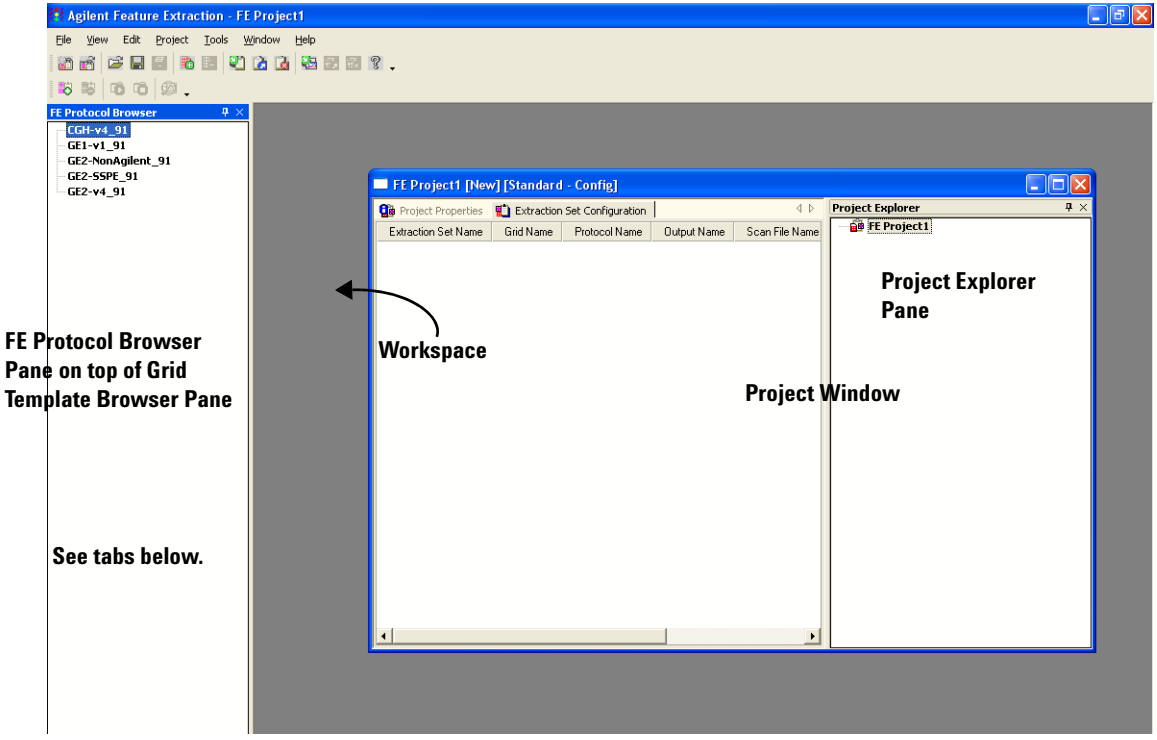


Figure 92 FE Standard Project Configuration window – new layout

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. See Figure 91.

  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.
5. 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.

   Your screen should now look like Figure 92.

   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.

Select the layout to use when software starts

You set the default layout for the Feature Extraction display in the Preferences dialog box.

1 Select **Tools > Preferences** from the top menu.

2 Change the following settings to produce the layout you want upon software 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 FE and reopen the software.

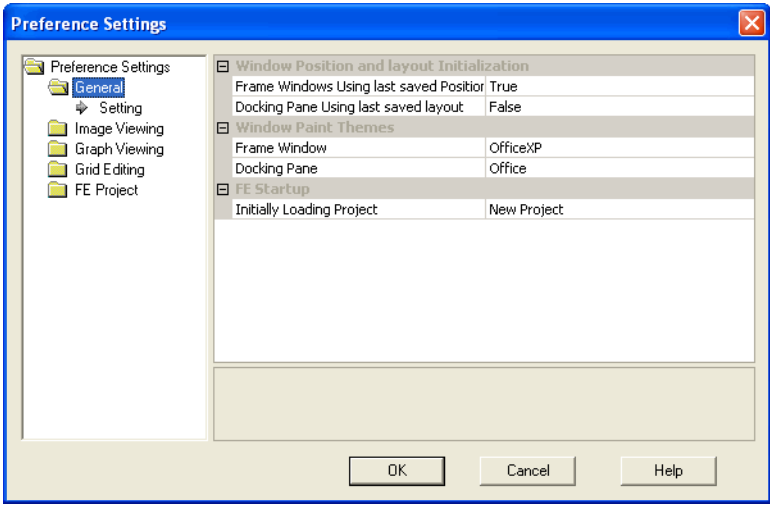


Figure 93 Preference Settings dialog box

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. To have Figure 92 appear upon startup, keep the setting as True.

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. Select True to produce Figure 92.

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. Keep New Project to reproduce Figure 92.

Index

F

G

H

I

L

M

N

O

P

www.agilent.com

In This Book

The *User Guide* presents instructions for using the the Agilent Feature Extraction software (v9.5).

This guide provides instructions for:

- setting up and running Feature Extraction on batches of image files
- creating grid files and templates 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

Printed in USA
Fourth Edition, February 2007

G2566-90012

Agilent Technologies